Washington Center for Equitable Growth

1500 K Street NW, Suite 850
Washington, DC 20005

Working paper series

Short and Medium Run Impacts of Preschool Education: Evidence from State Pre-K Programs

Mariana Zerpa

December 2018

https://equitablegrowth.org/working-papers/short-and-medium-run-impacts-of-preschool-education-evidence-from-state-pre-k-programs

Short and Medium Run Impacts of Preschool Education: Evidence from State Pre-K Programs

Mariana Zerpa*

October 3, 2018

Abstract

I study the effects of state preschool programs on child development, health, and academic progress from ages 5 to 12. I leverage variation in the timing of implementation of pre-K programs across states to look at the effects of a large and representative group of programs, using data from state legislatures and two national household surveys. I find that availability of pre-K programs improves developmental outcomes and causes persistent reductions in grade repetition. My results also suggest increased health problems only in the short run. I provide bounds for policy-relevant local average treatment effects that imply large effects.

JEL Codes: H75, I12, I21, I28, J13, J24.
Keywords: early education, pre-K, child development, health.

*Department of Economics, University of Leuven, Naamsestraat 69, Leuven 3000, Belgium (e-mail: mariana.zerpa@kuleuven.be). I am grateful to Jessamyn Schaller, Tiemen Woutersen, Gary Solon, Ron Oaxaca and Price Fishback for their advice and guidance. I thank Bart Cockx, Marco Manacorda, Costas Meghir, and participants at seminars at the University of Arizona, University of Leuven, Bank of Spain, Universidad Católica de Chile, FGV Sao Paulo, University of Antwerp, Université Libre de Bruxelles, and University of Luxembourg, for helpful comments. Part of the research in this paper was conducted at the National Center for Health Statistics Data Center, and I thank Pat Barnes for her support in providing access to data. Funding from the Washington Center for Equitable Growth, and from the University of Arizona Graduate and Professional Student Council, is gratefully acknowledged. The findings and conclusions expressed are solely those of the author and do not represent the views of Equitable Growth or the National Center for Health Statistics. I declare that I have no relevant or material financial interests that relate to the research described in this paper.

1 Introduction

Evidence on the technology of skill formation suggests that early education has the largest potential to mitigate the effects of socioeconomic disparities on the development of abilities and skills.[1] In the United States, state governments as a group are the largest provider of preschool education, accounting for more than half of the total (public and private) preschool enrollment of 4-year-olds. The number of states offering pre-kindergarten (pre-K) programs has grown dramatically since the 1990s and currently 43 states and the District of Columbia offer pre-K programs, serving 36% of all 4-year-olds.[2] Despite the potential for high social returns of high-quality early education programs (Belfield et al., 2006; Heckman et al., 2010), a third of the population of 4-year-olds are not enrolled in any preprimary education program (McFarland et al., 2018). Further investments in expanding preschool programs are on the agenda in many states, underscoring the need for causal evidence on the scope of benefits of large-scale public early education interventions.

A central topic in the debate about the effectiveness of public preschool education has been whether its impacts fade out over time. This debate was ignited by a randomized study of Head Start that finds small positive effects on test scores immediately after preschool that fade out during the following years (Puma et al., 2010, 2012),[3] and has resurfaced in light of recent evaluations of some universal state-funded pre-K programs on academic achievement. While evaluations of very short-run effects have mostly found positive effects on school readiness test scores,[4] the few papers that have evaluated the effects of these programs on test scores some years after preschool age have found mixed results on the persistence of early positive impacts.[5] The scarce evidence of persistent

[1] Cunha et al. (2006) and Heckman and Mosso (2014) provide extensive reviews of the literature on skill formation and the impacts of interventions in different stages of life.

[2] By contrast, the federal Head Start program, which targets children in poor families, only serves 11% of the population of 4-year-olds. This information corresponds to the 2016-2017 school year. Enrollment rate of 4-year-olds in state pre-K was obtained from Friedman-Krauss et al. (2018). Enrollment rate in Head Start was calculated using total enrollment reported by US Department of Health and Human Services (2018) and estimates of the total population of 4-year-olds from the U.S. Census Bureau.

[3] This conclusion has been questioned on the basis that a large part of the control group attended other public preschool programs. Kline and Walters (2016) show that the findings on fade-out are too imprecise to be conclusive after adjusting for non-compliance. In addition, observational studies have found persistent effects on test scores on some demographic groups but not others (Currie and Thomas, 1995, 1999; Deming, 2009).

[4] These short-run evaluations exploit birth date eligibility rules in regression discontinuity designs that compare children who have just finished with children who are about to start pre-K, and provide evidence on very short-run impacts on test scores (Gormley and Gayer, 2005; Gormley, Phillips and Gayer, 2008; Wong et al., 2008; Weiland and Yoshikawa, 2013). In a recent working paper, Cascio (2017) uses a similar strategy to evaluate the relative effectiveness of 14 universal and targeted programs, and finds that only the universal programs produce improvements in test scores for poor children.

[5] Lipsey, Farran and Durkin (2018) present results for the randomized controlled trial of the Tennessee Voluntary Prekindergarten Program, and find positive effects on achievement tests at the end of pre-K, but the differences with the control group fade out by the end of kindergarten, and even become negative by third grade for some outcomes. Using difference-in-difference strategies, Cascio and Schanzenbach (2013) find that the introduction universal pre-K

effects of preschool programs on test scores contrasts with the findings of large positive long-run impacts of some preschool programs in experimental and non-experimental studies.[6] This contrast raises the question of whether preschool programs affect other short and medium run outcomes that might contribute to long-run impacts.

In this paper, I study the effects of expansions of state-funded pre-K programs on academic, developmental and health outcomes of children. One of the main challenges in addressing this question has been the lack of data on attendance to state pre-K programs in national individual-level datasets. I overcome this challenge by collecting information on the timing of the introduction of state pre-K programs between 1997 and 2005. I identify 15 state pre-K programs that were introduced or significantly expanded during this period. These programs mainly expanded access to preschool education at age four, thus providing one year of education before school entry. I combine the information of pre-K availability by state and year with individual-level data from the Current Population Survey (CPS) and the National Health Interview Survey (NHIS) on the cohorts of children who were 4 years old during that period and are observed at ages 5 to 12. This allows me to estimate intention-to-treat effects of the introduction of pre-K programs on child outcomes by exploiting the differential timing of expansions across states as an exogenous source of variation in the availability of pre-K education.

This is first study to compile evidence on such a large and representative group of universal and targeted state preschool programs. The differential timing of policy changes across states provides two key advantages. First, it allows me to estimate the average causal effects of the introduction of a state pre-K program on child outcomes while controlling for permanent differences in child outcomes across states, changes over time at the national level, individual demographic characteristics, and time-varying characteristics of each state. Second, by combining state-level information on policy rollout with repeated cross-sections from large national samples of the CPS and NHIS, I am able

programs in Georgia and Oklahoma had positive effects on children's test scores as late as eighth grade for children with low-education mothers, while Fitzpatrick (2008) finds positive effects of Georgia's universal pre-K program in fourth grade only for disadvantaged children living in rural areas. Hill, Gormley and Adelstein (2015) study two cohorts of Tulsa's pre-K program using propensity score matching, and find evidence of persistence of early gains in test scores through the third-grade of school only for one of the cohorts, in math but not reading, and for boys but not for girls.

[6]The strongest evidence on long-run impacts of preschool comes from studies of the experimental Perry Preschool Program, which have found that this small, high-quality targeted intervention had long-run positive effects on outcomes such as educational attainment and earnings (Currie, 2001; Heckman et al., 2010). In addition, non-experimental evaluations of Head Start using sibling comparisons show long-term improvements in educational attainment, earnings, crime, and self-reported health (Garces, Thomas and Currie, 2002; Deming, 2009; Ludwig and Miller, 2007). Havnes and Mogstad (2011) also find positive effects of a large-scale expansion of subsidized preschool in Norway (for children ages 3-6) on educational attainment and labor market participation.

to provide a more complete picture of the effects of preschool education, both in terms of a broader set outcomes and the timing of the effects.

In addition, I go beyond intention-to-treat (ITT) effects and discuss and provide bounds for relevant local average treatment effects parameters. While ITT effects are of particular relevance because they capture the full effects of policy implementation on the affected cohorts, it is difficult to interpret their magnitude without an idea of how many children are affected by the policy. Providing estimates of treatment-on-the-treated effects has been particularly challenging in the context of preschool or child care expansions. In addition to the frequent lack of information on individual program attendance, the diversity of alternative child care arrangements complicates the definition of the relevant "treatment" and thus the estimation of the probability of treatment. I circumvent this problem by combining information on aggregate pre-K enrollment rates and individual-level preschool attendance from the CPS with my ITT estimates in Bloom-type estimators (Bloom, 1984). I provide a formal discussion of two relevant LATE parameters of interest in the presence of different counterfactual childcare arrangements and under the assumption of no externalities. First, I estimate the causal effect of adding Pre-K to the available childcare choice set, which has been the most frequently reported type of treatment-on-the-treated estimate.[7] Second, I present bounds for the causal effect of attending Pre-K relative to not attending any preschool, under the additional assumption of pre-K being at least as good as the next best alternative preschool program.

My results show that the presence of a pre-K program in a child's state at age four reduces the likelihood of repeating a grade at ages 6 to 8 by 2.5 percentage points (30%). I find no evidence of this effect being reversed in the following four years, implying sustained reductions in grade retention at least until age 12. I also find that pre-K expansions reduce the incidence of development and behavior problems by 9% of a standard deviation in the first four years after preschool age. In the following four years, the effect on developmental and behavioral problems becomes statistically insignificant but remains negative and about 60% as large as the short run effect. These results are in line with findings of sustained medium-run impacts of Head Start, such as reduced grade repetition and learning disability diagnosis at ages 7-14 (Deming, 2009), and a lower incidence of behavior problems and special education placement for boys at ages 12-13 (Carneiro and Ginja, 2014). They are also consistent with the results from evaluations of the impacts of universal public preschool expansions on grade retention and other academic outcomes

[7] See, e.g., Ludwig and Miller (2007) and Havnes and Mogstad (2011), and the discussions of this issue in Baker, Gruber and Milligan (2008) and García, Heckman and Ziff (2018).

later in childhood in Uruguay (Berlinski, Galiani and Manacorda, 2008), Argentina (Berlinski, Galiani and Gertler, 2009) and Spain (Felfe, Nollenberger and Rodríguez-Planas, 2015).

The estimated effects on health outcomes suggest that pre-K expansions may be associated with worse health in the short-run, but these impacts are not persistent. Pre-K expansions are associated with an increase of 0.6 days (19%) in the number of illness-related missed days of school in the first four years after preschool age. I also find weak evidence of an increase in the likelihood of the health status of a child being reported as fair or poor. These effects show a clear fading-out pattern and disappear by the fourth year after preschool. This contributes to the still incipient research on health effects of preschool and child care program expansions. The studies that have addressed this question find conflicting evidence, with some studies of Head Start finding evidence of positive effects in the medium run, and studies of child care subsidies finding negative effects on child health status and the incidence of illness in the short run.[8]

All the findings are robust to controlling for state-specific linear time trends, region-by-year fixed effects, or region-by-cohort fixed effects. The results are not sensitive to my choice of control states, which include 18 states without any pre-K programs before 2005, and are not driven by any one particular treatment or control state. I also show that there is no evidence of correlation between the timing of the implementation of pre-K programs and other potentially confounding state-level characteristics and policies, including K-12 expenditures, Medicaid/CHIP eligibility expansions, and other early childhood programs. There is also no evidence of pre-existing differential trends in the outcome variables between treatment and control states. The significance of the estimated effects is robust to multiple hypothesis testing adjustment using the Holm-Bonferroni method for controlling the family-wise error rate (Holm, 1979).

I estimate that this group of pre-K programs expanded state pre-K enrollment by 18 percentage points, while it expanded attendance to any preschool program by 7.3 percentage points. The difference between these two figures could be caused by crowding-out of other preschools, and/or by an overlap of enrollment in part-day pre-K and other preschool programs. These estimates, combined

[8]The Head Start Impact Study (Puma et al., 2010) found positive impacts on reported health status and health insurance coverage during Kindergarten that fade out by first grade. Ludwig and Miller (2007) show that Head Start reduced childhood mortality during its implementation in the 1960s, which is likely explained by the increased access to immunizations that the program provided (Currie and Thomas, 1995). For more recent cohorts, Carneiro and Ginja (2014) find that participation in Head Start causes improvements in mental health screenings, reductions in obesity prevalence, and some indication of reductions in disability, but they find no significant effects on reported health limitations, health status, or risky behaviors, for boys at ages 12-13 and 16-17. In contrast, Baker, Gruber and Milligan (2008) find that the Quebec child care subsidy had negative impacts on child health in the short run, and Herbst and Tekin (2012) find that access to subsidized child care is associated with increases in the prevalence of overweight and obesity among low-income children in the U.S.

with the ITT effects, suggest that treatment-on-the-treated effects are large. For example, I find that attending pre-K can reduce both grade repetition and developmental and behavior outcomes by between approximately 0.5 and 1.25 standard deviations for treated individuals, compared to the alternative of not attending any center-based preschool program.

The findings in this paper suggest that increased access to preschool education can have large positive impacts on child development and behavior outcomes in the short run, and sustained effects on academic progression in primary school. Estimated effects of Head Start from previous literature are within my estimated bounds of pre-K treatment-on-the-treated effects, suggesting that the average effectiveness of these programs is not very different from that of Head Start. In addition, the large magnitude of the upper bounds of the effects suggest that pre-K expansions may have impacts not only children who would otherwise not have attended any preschool, but also on children that are drawn from other preschool programs,[9] and potentially even externalities on non-attending children (across siblings and across children in the same cohort through peer effects in school).

This paper contributes to the literature on the effects of early education on child human capital by providing the first evidence on both short and medium-run impacts of a large group of state-funded pre-K programs. A key advantage of my research design is that it allows me to evaluate the effects of pre-K programs at different stages of childhood, up to eight years after preschool age. While the growing literature that examines the impacts of state preschool programs has mainly focused on short-term effects of a handful of (mostly universal) programs, my approach allows me to study a large group of programs that is representative of the diversity of universal and targeted state programs that children currently attend. The simultaneous study of short and medium run effects on a broad set of outcomes helps build a bridge between the conflicting evidence on the fade-out of short-run effects of preschool education on test scores, and the positive long-lasting effects on adult outcomes of Head Start and other preschool programs.

The paper proceeds as follows. Section 2 discusses the policy background and conceptual framework. Section 3 outlines the empirical strategy and presents the data and outcome variables used. Section 4 presents the estimated intention-to-treat effects of pre-K expansions, and discusses the robustness of these results. Section 5 discusses treatment-on-the-treated parameters of interests

[9] Black et al. (2014) find positive effects of child care subsidies in Norway on medium run academic outcomes, despite a small and insignificant effect of the subsidies on child care attendance. They suggest that this result may be explained by a positive shock to disposable income. In line with this explanation, previous research has shown positive causal impacts of family income on test scores and academic outcomes, especially in the lower parts of the income distribution (Dahl and Lochner, 2012; Løken, Mogstad and Wiswall, 2012).

and presents bounds for the magnitude of these these parameters. Section 6 concludes.

2 Background and conceptual framework

2.1 Institutional background

The programs studied in this paper are 15 state-funded pre-kindergarten programs for 4-year-olds that were first implemented or scaled up between 1997 and 2005, and that are qualified as state preschool programs by the National Institute for Early Education Research (NIEER). An initiative is considered to be a state preschool program by NIEER if it meets the following criteria: a) the initiative is funded, controlled, and directed by the state; b) it serves 3- and/or 4-year-old children; c) early childhood education is the primary focus of the initiative; d) it offers a group learning experience at least two days per week; e) it is distinct from the state's system for subsidized child care; f) the initiative is not primarily designed to serve children with disabilities, although it may include children with disabilities; and g) state supplements to Head Start are considered to constitute state preschool programs if they substantially expand the number of children served and the state assumed some administrative responsibility for the program.

The first map in Figure 1 shows the states that implemented pre-K programs in this period (*treatment states*) and the states without programs (*control states*). The control states are those that by 2005 had not yet implemented a state-wide pre-K policy, or which only had a very small scale pre-existing program whose impacts are unlikely to be observed at a state level. Most of these states (12 out of 18) did not have a state preschool program by 2005, and those that had one had low and stable enrollment rates of at most 6% of 4-year-olds by 2005.[10] The remaining states and the District of Columbia are excluded from the main analysis because they had pre-existing programs with enrollment rates of 10% or more during the period, and/or had a program with significant variation in enrollment during the period (*excluded states* hereafter). As discussed in the empirical strategy section, the reason for excluding these states from the main analysis is that they can have changes in enrollment and/or funding during the period that would make them an inadequate counterfactual for the changes in outcomes that would have happened in treatment states in the absence of the implementation of a program. A more detailed discussion of enrollment rates in each group of states is provided in Section 4.1.

[10] The *treatment states* are Arkansas, Florida, Kansas, Louisiana, Missouri, Nebraska, New Jersey, New Mexico, New York, North Carolina, Oklahoma, Pennsylvania, Tennessee, Vermont, and West Virginia. The *control states* are Alabama, Alaska, Arizona, Hawaii, Idaho, Indiana, Iowa, Minnesota, Mississippi, Montana, Nevada, New Hampshire, North Dakota, Rhode Island, South Dakota, Utah, Washington, and Wyoming.

2.1.1 Implementation of state-funded pre-K programs

States began to implement preschool programs for children in low-income or "at-risk" families in the mid-1960s, following the lead of the federal Head Start program. However, only a small number of states had implemented preschool programs for children below age 5 until the 1980s (Barnett et al., 2003). The 1990's and early 2000's was a period of increasing investment of states in preschool education. Georgia was the first state to implement a voluntary universal program, where children from all economic backgrounds are eligible, in 1995. By 1997, 17 states and the District of Columbia had pre-K programs with enrollment rates of at least 10% of 4-year-olds (these are the states excluded from my main analysis sample).

In the period between 1997 and 2005, 15 additional states created new programs or substantially expanded pre-existing ones. In most cases they already had one or more small-scale preschool programs, either to test as a pilot or to target at-risk children and/or children with disabilities. In these cases, the new policy created an additional program or expanded the scale of a pre-existing one. Figure 2 shows preschool enrollment rates by state group in 1997-2005 (left panel) and population-weighted aggregate pre-K enrollment rates in treatment states between 2001, the first year of available data, and 2005 (right panel). The increases in preschool enrollment rates in treatment states during this period follow closely the introduction of pre-K programs and the associated increases in pre-K enrollment of 4-year-olds.

Most of the pre-K programs implemented during this period were the result of a bill passed by the state congress, except for Florida (constitutional amendment approved by voters) and New Jersey (supreme court ruling). The Pre-K programs were implemented by legislatures controlled by both major political parties, although there is a larger proportion of programs passed while Democrats were in control of the state legislature. Of the 15 states programs, 7 were passed by legislatures controlled by the Democratic Party, and another 3 were passed by split legislatures with a majority of Democratic legislators.

In most cases the legislation concerned Pre-K exclusively and there were no other provisions related to K-12 education passed during the same year. The exceptions are states that also passed legislation–at or around the time of the introduction of pre-K–that provided funding for expanding full-day kindergarten (New York and Pennsylvania), and increases in per-student spending or reductions in class-size (New York, North Carolina and Oklahoma).[11] It should be noted, however,

[11] Kansas also had legislation that increased per-student spending, but it took effect four years before the pre-K program was implemented. Another especial case is New Jersey, where supreme court rulings required the state

that full-day kindergarten expansions affect only 5-year-olds, and would therefore also affect children who are in the cohort just before the first pre-K cohort. Expansions in K-12 funding would also affect older cohorts (although admittedly with less intensity the older the children are, due to fewer years of exposure to the increased funding). In Section 4.4 I discuss the robustness of my results to the potentially confounding effects of changes in school spending and expansions in full-day kindergarten, as well as other state policies.

2.1.2 Program characteristics and quality indicators

Table 1 presents a summary of the main characteristics of the programs included in this study as of 2005 (Barnett et al., 2006). Details about each program are provided in Appendix Table B1. Pre-K programs are most often provided in a combination of public schools and private preschool centers, including Head Start providers, through grants or other arrangements with the school districts. This has allowed states to expand enrollment more rapidly while maximizing the use of existing resources. I provide here a brief discussion of these characteristics, comparing whenever possible to Head Start, which is the main publicly provided alternative.

States offer a combination of full-day (5 to 7.5 hours/day) and half-day (2.5 to 3 hours/day) programs, and some states offer both types of programs depending on the decision of the school district. There is similar variation in instruction time across Head Start centers; for example, 63% of the centers in the Head Start Impact Study (conducted in 2002) provided full-day service (Walters, 2015). The heterogeneity in hours-per-day served is mirrored by a similar heterogeneity in total spending per student (including all sources of funding): the average spending in states that offer only full-day programs is \$6,118, compared to \$2,929 in states that only offer half-day programs. The average spending per student across all treatment states is \$4,848, which represents about two thirds of Head Start's average spending per child (\$7,276 as reported by Barnett et al., 2006).

There is large variation across treatment states in pre-K enrollment rates. The second map in Figure 1 shows enrollment rates in the last year of the sample period (school year 2005-2006). Five of the 15 states (Florida, New York, Oklahoma, Vermont, and West Virginia) offer voluntary universal pre-K programs, although not all of them were at the time sufficiently funded to meet demand. Enrollment rates of 4-year-olds in these universal programs range from 29% in New York to

to also implement other improvements in Abbott districts, but most of these took impulse three years after the implementation of the Abbott Preschool Program.

70% in Oklahoma, with an average of 46.6%. The rest of the states offer programs targeted towards children from low-income families or who have other risk factors, with enrollment rates that range from 4% to 22% of 4-year-olds in the state (12.8% on average). Some states use income thresholds to determine eligibility of individual children, while others determine eligibility of a school or provider by requiring a minimum percentage of children served to be below the income threshold. The most commonly used income threshold is 185% of the federal poverty line (FPL), which determines eligibility for reduced-price lunch in schools, and there are two states that use the same income threshold as Head Start (100% of the FPL). The average enrollment rate across all programs in treatment states is close to 24%. During the period of study, federally-funded Head Start enrollment for four-year-olds is relative stable around 12% of the four-year-old population, which does not meet the demand from eligible families. Given Head Start's and state pre-K programs' income eligibility thresholds, it is likely that targeted state programs serve children from Head Start's eligible families' unmet demand as well as children from marginally better-off families, while universal state programs serve children from families that belong to a broader range of the income distribution.

State pre-K programs generally have educational goals that resemble those of Head Start. All but two require providers to follow comprehensive early education standards (Kansas and New York adopted comprehensive early education standards after 2005), and over 70% of the states require programs to offer basic health screenings, referral and support services. Most states (13 out of 15) impose a limit of no more than 20 children per class and at least one staff every 10 children. This is somewhat larger than in Head Start centers, where the average child:staff ratio is 6.8 with a cross-center standard deviation of 1.7 (Walters, 2015).

There is more variation in terms of teacher and assistant teacher degree and specialization requirements. 9 out of the 15 programs require teachers to have a bachelor's degree, and 10 require some specialization in early education, with only 3 programs not requiring either. In comparison, Head Start also has large variation in teacher qualification standards across centers during this period, with only 35% of Head Start teachers having bachelors degrees and 11% having teaching licenses (Walters, 2015).

The heterogeneity in the above input quality indicators is reflected in the variation in the program scores assigned by NIEER. This score is the count of the number of benchmarks met by the program, out of a total of 10.[12] Among this group of pre-K programs, NIEER scores range

[12]The benchmarks are the following: comprehensive early learning standards; teachers have a BA degree; teachers are specialized in pre-K education; assistant teachers have CDA degree or equivalent; teacher in-service at least 15 hours/year; class sizes of 20 or lower; staff-child ratio of 1:10 or better; screening/referral for vision, hearing and

from a minimum of 3 to a maximum of 10 out of 10, with a median of 7 (Barnett et al., 2006). Within this group of programs, NIEER's score has a correlation coefficient of 0.59 with spending per student.

Besides Head Start, the main alternative sources of care to pre-K programs are private preschool programs, daycare centers, family daycare, and other non-center based child care arrangements (parental, relative and non-relative care at home). Children from poor families are more likely to be enrolled in Head Start, while enrollment in preschool centers other than Head Start increases with income.[13] Conditional on income, hispanics are less likely to attend preschool and more likely to be cared for by a relative. It should be noted, however, that the expansion of pre-K provision is not necessarily a substitute for other preschool programs. In many cases private preschool providers pool funds from different sources, including state pre-K, Head Start, and vouchers from child care subsidies, not only to increase the number of slots offered but also to improve quality and/or extend the amount of hours per day of service provided (Magnuson, Meyers and Waldfogel, 2007). Therefore, pre-K expansions may not only expand the number of children that attend preschool centers (who would otherwise be in other relative or non-relative child care arrangements), but can also affect the intensive margin (number of hours of preschool attendance and/or quality).

2.2 Conceptual framework

The conceptual framework for the empirical analysis is inspired by the economic models of child human capital production and parental decisions developed in Cunha and Heckman (2007), Cunha, Heckman and Schennach (2010) and Heckman and Mosso (2014). I propose a simple model of child human capital (or "child quality") production that incorporates the public provision of preschool education, and discuss the implications of public preschool provision for parental decisions and the accumulation of human capital of a child at different stages.

Suppose a household is composed by a parent and her child. The child's life is simplified to three stages: (1) early childhood, when investments in preschool education and other child quality investments are made, (2) a first stage of childhood, when further investments in child quality are made, and (3) a final stage of childhood when the final quality of child human capital is realized.

---

health, and at least 1 support service; at least 1 meal a day; and monitoring site visits.

[13]The federal and state governments also provide child care assistance to low income families through three federal block grants to the states: The Child Care and Development Fund (CCDF), transfers to the CCDF from the Temporary Assistance for Needy Families (TANF) block grant, and the smaller Social Services Block Grant (SSBG). These funds are used to assist low income families in paying for private childcare services, usually through vouchers, which can include private preschool programs and daycare (Magnuson, Meyers and Waldfogel, 2007).

Child quality in each of these stages is indicated by $\mathbf{k}_t$, with $t = 1, 2, 3$. Child human capital is highly multidimensional and is constantly changing as children grow up. Since the child outcomes considered in this paper relate to academic progress, child development, behavior and health, it is useful to think of child quality as a set of three interrelated vectors: cognitive skills, noncognitive skills, and health stocks (Heckman and Mosso, 2014).

In the first period, early childhood, I assume that the initial child quality is exogenously given by $\mathbf{k}_1$. In the second and third periods, child quality is endogenously determined by the previous period's child quality production function and its inputs. These inputs include the previous period's level of child skills, parental skills (s), and parental investments into the production of child human capital. In both periods of investments (periods 1 and 2), parents can make material investments in child quality production, denoted by I_1 and I_2. In addition, in early childhood parents can also invest in parental childcare time (P_1), time spent by the child at (private) center-based preschool education (C_1), time spent by the child at publicly provided preschool education (G_1), other informal (non center-based) childcare arrangements (N_1).

$$\mathbf{k}_2 = f_1(s, \mathbf{k}_1, I_1, P_1, C_1, G_1, N_1)$$
$$\mathbf{k}_3 = f_2(s, \mathbf{k}_2, I_2) = g_2(s, \mathbf{k}_1, I_1, I_2, P_1, C_1, G_1, N_1)$$

I assume that parents are altruistic and maximize the present net value of the stream of future child income as a function of child quality in the last stage $R(\mathbf{k}_3)$, subject to three types of constraints: time constraints, the public provision of preschool, and the budget constraint of the parent. Assuming no leisure for simplicity, the parent's time constraints in the two periods of investments impose that she works full-time in the second period, and must distribute her total time between work hours and child care in the first period:

$$T = h_1 + P_1$$

An additional time constraint is given by the fact that during early childhood the child must be supervised while the parent works, either by attending private center-based preschool or publicly provided preschool:

$$h_1 = C_1 + G_1 + N_1$$

The parent's choice of the time spent by the child in publicly provided preschool is restricted by the availability of public preschool provision. Let $\mathcal{G}$ be the fixed number of hours offered by the government of preschool education, at no cost to the parent ($\mathcal{G} \geq 0$). The parent can choose between sending the child to public preschool for this amount or zero hours:

$$G_1 = \{0, \mathcal{G}\}$$

In this framework, state pre-K programs enter in this model by relaxing this last constraint, i.e. increasing $\mathcal{G}$. This paper aims to assess the impacts of increasing the public provision of preschool education on outcomes that are associated to different dimensions of child quality, at different stages in childhood. I am interested in assessing $\partial \mathbf{k}_2 / \partial \mathcal{G}$ and $\partial \mathbf{k}_3 / \partial \mathcal{G}$, which are the intention-to-treat effects of expanding the supply of public preschool, as well as $\partial \mathbf{k}_2 / \partial G_1$ and $\partial \mathbf{k}_3 / \partial G_1$, the effects of attending public preschool.

Relaxing the public provision constraint should increase attendance to public preschool programs. For some children, this can increase the total hours they spend in a preschool programs ($G_1 + C_1$), by decreasing the time spent in other informal childcare arrangements (N_1), or by decreasing the time spent with the parent (P_1), since the public provision of preschool decreases the opportunity cost of working. The increased public provision may also crowd out attendance to more costly private preschool (increasing G_1 and reducing C_1).

There are several pathways for a causal effect of expansion of public preschool provision on child outcomes. First, public preschool is a direct input in the production of child skills. Pre-K programs are designed to prepare children for school, encouraging the development of cognitive and non-cognitive skills, potentially having a direct impact on child development. The effect of the increase in attendance to publicly provided preschool will depend on the relative productivity (quality) and substitution and complementarity patterns between the different inputs in these production functions. For example, it will depend on the relative quality of pre-K programs compared to the alternative childcare options available, as well as compared to the quality of parental childcare. Pre-K may also increase the productivity of parental childcare, for example by providing the parent with information about how to better care for her child.

An increase in the public provision of preschool also lowers the parent's opportunity cost of working. This can have ambiguous effects, depending on the relative productivity of the marginal hours of parental childcare relative to public preschool in the production of child skills. In addition, an increase in parental labor supply will increase family income, which can increase other child investments. Even for parents who would in any case be working and sending their children to preschool, a switch from private to publicly provided preschool can increase disposable income.

The relative impacts at different stages of childhood depend on the dynamic complementarity of skills ($\partial \mathbf{k}_3 / \partial \mathbf{k}_2$), on how parental investments respond to changes in child quality, and on the productivity of other inputs in the production of skills in the last period (e.g. the quality of schools). There could be a fading out of positive initial impacts if investments at different childhood stages are complementary in the production of skills and require follow-up to be effective (Cunha and Heckman, 2007), or if later parental and/or school investments are compensatory of differences in skills across children, although the evidence appears to point towards parental investments being reinforcing rather than compensatory (Aizer and Cunha, 2012).

Although the health production technology can have some differences with respect to cognitive and noncognitive skills, in general all the same potential channels also apply for health outcomes. Preschool programs can be a direct input in health production, since they usually have an explicit goal of preparing children for school in terms their physical health. This involves for example offering health checkups and requiring immunizations, teaching children and their parents healthy habits, and improving access to preventive care and to health information for parents. However, expanding access to preschool can also have a negative effect on child health in the very short term if it implies that the child's exposure to illness is increased through contact with other children and teachers (e.g. colds, ear and gastro-intestinal infections). It is not clear how this would affect longer-term health outcomes, as the evidence for positive effects on the immune system of earlier exposure to infections is limited to group childcare attendance of children in the first two years of life (e.g. Ball et al., 2000). On the other hand, Pre-K availability can also change parental behaviors and time use, potentially affecting health. If preschool expansions induce increases in maternal work, it can negatively affect children's health through reducing the time shared with the mother and increased maternal stress (Anderson, Butcher and Levine, 2003; Gennetian et al., 2010; Morrill, 2011). This could be reflected for example in an increase of stress-related symptoms such as frequent diarrhea and headaches.

So far, we have considered the process of skill and health production of each child as independent of other children, without any externalities across children. Externalities in the production of child skills can emerge for example from peer effects in school. This could mean that positive effects on the skills of children who attend public preschool could spill over to other children in their cohorts. In the case of health, if children attending preschool are more exposed to contagious illness, the contagion effects may affect not only the children in the same cohort but also siblings and their classmates.

3 Empirical strategy and data

3.1 Empirical Strategy

I take advantage of increases in the supply of preschool education caused by the introduction of state pre-K programs to identify its effect on short and medium run outcomes of children. Between 1997 and 2005, 15 states implemented or substantially expanded state-funded pre-K education programs for four-year-olds. I evaluate the reduced-form (intention-to-treat) effects of these preschool education supply expansions using individual-level data on child developmental and health outcomes from the National Health Interview Survey (NHIS) and the Current Population Survey (CPS), supplemented with state-level information on the implementation of pre-K policies and other state characteristics and policy variables. I estimate regressions that take the form of a generalized difference-in-difference specification with state and cohort fixed effects:

$$Y^a_{isc} = \beta^a_{RF}\text{Post\_Pre-K}_{sc} + \gamma^a X^a_{isc} + \delta^a_c + \delta^a_s + \varepsilon^a_{isc} \tag{1}$$

The subscript i represents a child, c is the child's pre-K cohort (the *reference year* for attending preschool, i.e. the year when the child was 4 by the beginning of the school year), and s is the state where the child lives. The superscript a indicates that the model is separately estimated for child outcomes evaluated at different number of years after preschool age. In my main specifications I group children in two age groups: 1-4 and 5-8 years after pre-K, i.e. observed around ages 5 to 8 in one group and 9 to 12 in the other. This allows me to discern short and medium run effects, while maintaining large enough samples. Post_Pre-K_{sc} is an indicator variable for whether state s had implemented a pre-K program by year c. X^a_{isc} is a vector of control variables whose components vary across specifications, but in the more general case includes individual time-invariant characteristics of the child and her family (the child's gender, race/ethnicity, and mother's educational attainment), dummies for the number of years after pre-K age when the child is observed, and state policy and economic control variables that vary by state and cohort, including state characteristics when the child was 4 (year c) and current characteristics in the year when the child outcomes are observed. The state fixed effects, δ^a_s, control for unobserved differences across states (e.g. permanent differences in the quality of education or health care), and cohort fixed effects, δ_c^a, control for any unobserved changes across cohorts that are common to all states (e.g. general changes in female labor supply, parents' valuation of preschool education, national changes in health outcomes).

The estimated effects that result from the estimation of equation 1 can be interpreted as estimates of the intention-to-treat (ITT) effects. They represent estimates of the reduced-form impacts of the introduction of a pre-K program in a state on all the cohorts of children who live in the state and are age-eligible for pre-K (age 4) after the introduction of the program. This interpretation of the reduced-form estimates is similar to the ITT interpretation of the estimates of impacts of child care policy expansions of Baker, Gruber and Milligan (2008) and Havnes and Mogstad (2011). The main advantage of the estimation of ITT effects is that it allows us to capture the full effects of the policy on a cohort, taking into account for example the size of the actual expansion and any take up issues, the full impact of changes in child care arrangements, and any externalities that the participation of a child in a pre-K program may have on non-participant children. However, it also has the disadvantage that it is very sensitive to the size of the expansions, making it hard to recover treatment-on-the-treated parameters. In section 5, I discuss the strategies I use to approximate local average treatment effects.

If we consider as the "treatment" the availability of a pre-K program in a child's state, this empirical design can be interpreted as a sharp difference-in-differences (DID) design, where there is more than one treated group with treatment beginning at different time periods. Identification relies on the assumption that the timing of the implementation of state pre-K programs is not correlated with other factors that may affect child outcomes. While the exogeneity of the policy variable cannot be directly tested, Section 4.4 presents evidence showing that the implementation of state-funded pre-K programs is not significantly correlated with other state-level characteristics and policies that may affect child health and development. I also discuss the sensitivity of the results to the inclusion of different sets of state-level control variables, adding state-specific linear time trends, and alternative choices of the states included in the control group. In addition, the time frame of the study begins in 1997 to avoid the potentially confounding effects of welfare reform.[14]

We may also think of attending pre-K as the "treatment," and consider the introduction or expansion of a pre-K program as a shock that increases the share of treated children in a state. In

[14]Between 1993 and 1996, 43 states received welfare waivers to requirements of the Aid to Families with Dependent Children (AFDC) program, as a first stage of the welfare reform (Falk, 2013). In 1996 the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA) instituted the Temporary Assistance for Needy Families (TANF) program, which replaced AFDC and became effective in July, 1997.

this context, this empirical setting can be interpreted as a fuzzy DID design, where the share of treated units (children attending pre-K) increases more in treatment states than in control states. In a fuzzy DID setting, identification also requires a stable percentage of treated units in the control group (De Chaisemartin and D'HaultfŒuille, 2018). This assumption guides my choice of control states in my main specification, which exclude those states that had pre-existing pre-K programs by 1997. Figure 2 (left panel) shows preschool enrollment rates in treatment, control, and excluded states. During the sample period, preschool enrollment was fluctuating in excluded states, with a general increasing trend. Increases in preschool enrollment during this period have been mostly associated to the implementation and expansion of pre-K programs. States without pre-K programs (control states) had a relatively stable preschool enrollment rate of 4-year-olds throughout the period.

Because pre-K policies have different characteristics across treatment states, each ITT effect that I estimate constitutes a population-weighted average across the marginal effects in the different treatment states. In particular, it averages across the different increases in the coverage of public pre-K that each program introduction causes, and across phase-in and full implementation years of these expansions. In addition, it averages across the marginal effects for the different populations that are affected by these programs in different states, for example in terms of socioeconomic status and alternative child care arrangements. An advantage of this empirical strategy is that the policy effects are estimated for a large group of pre-K policies that is representative of the diversity of programs currently being implemented across the U.S.

All estimated parameters are presented with their corresponding robust standard errors clustered at the state level. This allows errors to be arbitrarily correlated across children living in the same state, across the different years and cohorts.[15] In addition, and because I estimate effects on four main outcome variables, I also present p-values that are adjusted to reflect the inference problem caused by multiple hypothesis testing using Bonferroni-Holm method (Holm, 1979). I use this method because of its simplicity and because it provides strong control of the family-wise error rate, although it is more conservative that other (e.g. bootstrap or re-sampling) alternatives (Romano and Wolf, 2005). For simplicity, I present adjusted p-values that can be compared to the significance level (0.05).[16]

[15]To account for the fact that standard errors are clustered at the state level and the number of states is relatively small, all significance tests and confidence intervals are computed using a t-student distribution with $G-1$ degrees of freedom, where G is the number of states in the sample. In my main estimation sample there are 33 states.

[16]Since I test four hypothesis for each sample, I adjust the first (lowest) p-value by multiplying it by 4. If the corresponding null hypothesis is rejected (i.e. if the adjusted p-value is ≤ 0.05), I multiply the second-lowest p-value

3.2 Data and measurement

I construct two separate samples using individual-level CPS October Supplement data. The first sample consists of 4-year-olds observed in the October CPS from 1997 to 2005. I use this sample to evaluate the effect of the implementation of pre-K programs on preschool attendance. Because the CPS does not include information about date of birth, I assume that all children who are 4 years old when observed in the October questionnaire would satisfy the age eligibility requirement for state pre-K. I code a child as attending preschool if she is reported to be attending school at the nursery level. Table 2 shows summary statistics for the CPS sample of 4-year-olds. 60.4% of the main sample of 4-year-olds (in treatment and control states) are attending preschool at the time of the survey. Those not attending preschool are either not attending school (33.2%) or are already enrolled in kindergarten (6.4%).[17] The full 4-year-old sample has 15,541 observations, while the main 4-year-old sample (after dropping observations from excluded states) has 8,880 observations.

The second sample from the CPS consists of children observed at ages 5 to 12. I construct this sample by selecting the children ages 5 to 12 in the October Supplements of the CPS Surveys of 1998 to 2014 who would have been 4 years old at some October between the years 1997 and 2005. Because date of birth information is not available, I define eligibility for pre-K as the year when they are interviewed (in October) minus the difference between the child's current age and 4. I use the these data to construct one of the four outcome variables, grade repetition. I construct an indicator for whether a child is currently repeating a grade, based on questions about what grade a child is currently attending and what grade she attended in the previous year in the CPS October Supplement. I only define this outcome variable for children observed 2 to 8 years after preschool age, thus omitting the first year after preschool age when children are likely in kindergarten and haven't had the opportunity to be repeating a grade. In the full sample of the CPS children ages 6-12, 4.5% are observed repeating a grade. Because grade repetition is most common in the first years of primary school, this outcome variable is particularly relevant in the short run after preschool. In the full sample, the mean of this outcome for children observed 2-4 years after preschool age is

by 3, and so on. If a hypothesis cannot be rejected, I multiply all the following p-values by the same number as the first non-rejected p-value. Finally, I enforce that p-values are not greater than 1.

[17] Magnuson, Meyers and Waldfogel (2007) compare the 1999 CPS measure of school attendance of 3- and 4-year-olds to the more detailed data on child care arrangements from other surveys (NHES 1999, the ECLS-K 1998, and NSAF 1999), and their findings indicate that the measure of school enrollment in the October CPS is similar to measures in other studies that include center-based care, Head Start, nursery school, and pre-kindergarten. From this comparison it seems that parents do not identify informal child care and family day care as 'school', even if the latter is a licensed child care provider. Therefore, the alternative modes of care to preschool attendance according to the question in the CPS October survey would include being cared for at home or through informal child care arrangements, attending family day care, or enrolling in kindergarten.

8.4%, while it is only 1.6% for children observed 5-8 years after preschool age. Summary statistics for the CPS data on 5 to 12-year-olds are reported in Table 3. After dropping observations with missing values in the outcome variable and controls, the full sample has 127,070 observations, while the main estimation sample (in treatment and control states) has 70,644 observations.

I use information on developmental and health outcomes of children from repeated cross-sections of the National Health Interview Survey (NHIS) from 1997 to 2014.[18] The sample includes children in the "Sample Child" files, supplemented with information for the same sample from the "Person Level" files, for children of ages 5 to 12. I impute the year in which each child would have been eligible for pre-K using the information on the month and year of birth, and I keep in my sample the children who are observed between 1 and 8 years after pre-K age. I define eligibility for pre-K based on the year when a child would have been 4 years old by September 1, which is the most frequent birthday eligibility thresholds used by states.[19] The number of years since they were 4 years old determines the age group to which they belong. Summary statistics for the NHIS data are reported in Table 3. After dropping the observations with missing data in the main outcome variables and individual controls, the full sample of children whose age is determined to correspond to 1 to 8 years after pre-K has 36,385 observations. The main estimation sample consists of children who live in treatment and control states, which has 16,911 observations.

I use the NHIS data to construct three outcome variables: a Development and Behavior Problems Index (DBPI), and two measures of physical health. The DBPI is constructed using a Principal Component Analysis (PCA), as an underlying factor that is obtained from multiple observed measures of child functional limitations related to developmental and mental, emotional or behavioral disabilities.[20] To perform PCA analyses I use five measures of limitations in daily life related to disability that can be related to developmental problems and have the potential to affect a child's academic achievement[21] : "speech problem," "intellectual disability," "other developmental prob-

[18]This part of the empirical analysis was conducted at a National Center for Health Statistics (NCHS) Research Data Center because the state of residence is restricted access information.

[19]I have also explored the sensitivity of my results to using October 1 as the threshold for all states, and to using each state's legal threshold. All of the results using NHIS data are robust to both alternatives (available upon request).

[20]The rich information available in the NHIS would allow me to explore the impacts of pre-K on a large set of child outcomes, but this would imply at least two problems. First, each outcome would be an imperfect measure of child development. Second, testing effects on multiple outcomes comes at the cost of a multiple inference problem: as the number of outcome variables grows, so does the probability of incorrectly rejecting a true null hypothesis of no causal effects. Recent papers in the literature have dealt with the high-dimensionality of child skills, health, or wellbeing by using similar data-reduction methods, such as the computation of summary indices based on the variance-covariance matrix of observed measures (Anderson, 2008; Deming, 2009; Carneiro and Ginja, 2014; Schaller and Zerpa, forthcoming), or through factor analysis (Aizer and Cunha, 2012; Attanasio et al., 2017).

[21]Within children 3 to 21 years old served under Individuals with Disabilities Education Act (IDEA), the largest

lem," "other mental, emotional or behavioral problem," and "learning disability." The frequency of these conditions associated with limitations in daily life in the full sample of children in the NHIS are 2.5% for learning disabilities, 2.1% for speech problems, 1.5% for other mental, emotional or behavioral problems, 0.9% for other developmental problems, and 0.3% for intellectual disability. The PCA results are described in the Data Appendix (A). I start with no assumptions about the number of principal components that should be retained, and based on the results I decide to retain one principal component. The resulting projected principal component is standardized so that it has a mean of 0 and a standard deviation of 1 in the full sample.

The health measures include an indicator variable for whether the health status is reported as fair or poor, and the number of days of school missed due to illness or injury in the past 12 months. 2% of children in the full sample are reported as having fair or poor health, and they missed an average of 3.2 days of school. To explore health effects in more detail, I also explore the effects on four indicators of specific health conditions. These are conditions that are frequently observed in children and that allow me to explore different pathways through which preschool education may affect child physical health. More details about these variables are provided in A.

I supplement the NHIS and CPS samples with state-level control variables collected from various sources. Tables 2 and 3 include summary statistics of the state-level variables merged to the CPS sample of 4-year-olds, and to the CPS and NHIS samples of children ages 5-12. The state-level variables are merged to the CPS 4-year-olds dataset by state and year, for the years 1997 to 2005, and they are merged to the NHIS and CPS datasets of children ages 5-12 in different ways. First, I merge all the state-level control variables relevant to young children and their families for the years 1997 to 2005 using the year when the child was 4 (*reference year*). Second, I merge the covariates for the years 1998 to 2014, using the year when the child health outcomes are observed (i.e. at the current age). More details about the data sources and variables are provided in A.

The main explanatory variable is a dummy variable that indicates whether there was a State Pre-K program available in a child's state when the child was 4 years old. The primary source of information on the availability and characteristics of pre-K programs are NIEER State of Preschool reports. NIEER began collecting and reporting information on state-funded pre-K programs in 2003, with data corresponding to the 2001-2002 school year, but each report includes background information about the programs described, including brief information about the history of program

category of disability in the 2005-2006 school year was *specific learning disability* (41%), followed by *speech or language impairments* (22%), and *other health impairments*, *mental retardation*, and *emotional disturbance* (each in the order of 7-8%) (Snyder and Dillow, 2010).

or significant recent changes. After identifying the existing programs, I establish the school year in which each initiative is effectively established or expanded as the school year in which funding is allocated, based on the corresponding state legislation collected through the State Policy Database of the Education Comission of the States (2015) and the states' legislature online databases. I create the variable that indicates if a child was exposed to a pre-K program (*Post Pre-K*) by combining the state-level information on the year of implementation of a pre-K program in treatment states with individual-level information on each child's state of residence and preschool age cohort (year when the child was four). For children in the NHIS and CPS samples of 5 to 12-year-olds, the state of residence at age 4 is not observed, so I use the current state of residence.[22]

4 Results

4.1 Four-year-olds' preschool attendance and maternal employment

This section documents the increase in attendance to state pre-K programs, attendance to any preschool, and changes in maternal employment, associated to the pre-K expansions included in this study. This serves four purposes. First, looking at the change in enrollment to pre-K programs allows me to check the relevance of the main explanatory variable. Second, comparing the changes in pre-K and overall preschool attendance is useful to understand the degree of crowding-out of other preschool sources implied by these expansions. Third, these estimated impacts on enrollment are also useful to calculate, under certain assumptions, approximations to treatment-on-the-treated effects, which are presented in section 5. Finally, looking at the changes in maternal employment associated with pre-K expansions allows me to explore one of the potential channels for impacts on child outcomes.

I begin by discussing the effects of the pre-K expansions on pre-K enrollment. Because information on pre-K enrollment at the individual level is not available, I use information on the aggregate enrollment of 4-year-olds in pre-K programs by state reported by Barnett et al. (2006) for the school years 2001-2002 to 2005-2006, to construct a proxy for the change in pre-K enrollment rates in treatment states before and after a pre-K expansion. The average enrollment rates in the years before and after pre-K expansions are presented in Appendix Table B2.[23] I compute a weighted

[22]If some children move across states between the ages of 4 and 12, this can introduced some measurement error in my estimates. Unless they are selectively moving in response to the introduction of Pre-K programs, this constitutes classical measurement error that can introduce a downward bias to my estimates. This bias is likely to increase with child age as the likelihood of moving states increments with the distance between the time when child is observed and when they were 4 years old.

[23]I compute the *Post Pre-K Average* enrollment as the simple average of the enrollment rates for the years after

average of the expansion of pre-K enrollment across all treatment states using the population of 4-year-olds in 2000 as weights. This results in an approximated increase of 18.4 percentage points in the enrollment of 4-year-olds in pre-K programs. The states with enrollment increases above the average are, in descending order, Florida, Vermont, Oklahoma and New York, with increases that range from 47 to 24 percentage points, which are four of the five states with universal programs. Among the states with targeted programs, the two with the largest expansions were New Jersey and West Virginia, while four states had increases in enrollment of 6 percentage points or less: Nebraska, Pennsylvania, Missouri and New Mexico.[24]

Using the CPS sample of 4-year-olds, I estimate the impacts of pre-K expansions on individual attendance to any preschool program. I estimate the following equation:

$$P_{isc} = \beta \text{Post\_Pre-K}_{sc} + \gamma X_{isc} + \delta_c + \delta_s + \varepsilon_{isc} \tag{2}$$

where the outcome variable P_{isc} is an indicator of preschool attendance for child i, living in state s in year c (preschool cohort). The regressions include state and year fixed effects as well as individual and state level controls, and are weighted using CPS weights. The results are presented in the first column of Table 4. The results for the full sample indicate that after the introduction of a pre-K program, attendance to preschool increases by 7.3 percentage points, which implies a 13% increase relative to the control states' average preschool enrollment throughout the period. The results are similar for children with mothers of low (high-school or less) and high (more than high-school) educational attainment.[25] The estimated increase in preschool attendance is about 60% smaller than the estimates of pre-K enrollment presented above, which suggests the presence of substantial crowding-out of other preschool programs, and/or attendance to pre-K in addition to other preschool programs (an intensive-margin effect).

Given the heterogeneity of the programs studied, I explore whether there are heterogeneous impacts on enrollment across targeted and voluntary universal pre-K programs. Column 2 of Table

a state pre-K policy was implemented or expanded, if implemented in or after 2001, or the simple average of years 2001-2005, if implemented before 2001. *Pre Pre-K Average* enrollment is computed as the simple average of the enrollment rates for the years before a state pre-K policy was implemented or expanded with available information, if implemented after 2001. For the five states with pre-K programs implemented in 2001 or before, I use additional information to approximate the pre-expansion enrollment in pre-K in the state. The details for how this is done for each of the five states are provided in the footnote of Table B2.

[24] In a robustness analysis (Section 4.4) I explore how my estimated results change if the treatment variable is switched off for these four states.

[25] The estimations by gender, available upon request, are not statistically significantly different from one another, although the point estimate is larger for girls.

4 shows the results of adding the interaction of the Post Pre-K expansion variable with an indicator for whether the program is universal. I cannot reject the hypothesis that the coefficient for the interaction of *Post Pre-K* and the indicator for a universal program is zero, meaning that there is no statistically significant difference in the impact of the two types of programs on overall preschool enrollment rates. Given that enrollment in universal programs is higher, this suggests that there is larger crowding out from other preschool programs in states with voluntary universal pre-K. Since I do not have data on the specific type of program that each child attends, I cannot directly test the hypothesis of crowding-out or of simultaneous attendance to multiple programs. The CPS asks whether the child attends public or private school, but it is not clear whether this identifies the source of funding.[26] In Appendix Table B3, I present estimates of the effects of the two types of programs on reported public and private preschool attendance, as suggestive evidence. Universal pre-K expansions are associated with larger increases in public preschool attendance, and with *decreases* in private preschool attendance, suggesting the presence of crowding-out.[27]

In columns 3-6 of Table 4, I explore the effects of pre-K expansions on maternal employment and full-time employment. In this case, there are interesting differences by type of program and maternal educational attainment. Overall, pre-K expansions have no significant effects on maternal employment status for the full sample or for each of the educational attainment samples. However, for the sample of low education mothers, there are asymmetric effects by type of program; targeted pre-K programs have a significant negative effect on maternal employment (of 8 percentage points), while universal programs have a positive but not statistically significant effect. Turning to the effects on full-time employment, the results of universal pre-K are stronger, with a significant increase in the likelihood of working full-time of 11.5 percentage points among mothers of 4-year-olds with high-school education or less.

[26] Head Start programs and many state pre-K programs have public funding but are carried out by private providers, and it is not clear that parents can identify whether a particular preschool center is publicly or privately funded unless it is actually located in a public school.

[27] My estimate of 60% crowding out of other preschool programs is very close to the estimates by Cascio and Schanzenbach (2013) for the universal pre-K programs of Georgia and Oklahoma, and by Cascio (2017) for low-income children attending universal and targeted pre-K programs. However, Cascio finds similar rates of crowding-out in the universal and targeted programs she studies, while my estimates suggest a much lower level of crowding-out in the targeted programs included in my sample.

4.2 Intention-to-treat effects of pre-K expansions on child outcomes

Table 5 presents reduced-form estimates of the effects of a state pre-K expansion on academic, development and health outcomes by age-group (1-4 and 5-8 years after pre-K age).[28] The results show significant beneficial effects on academic and developmental outcomes in the first four years after preschool age. Children who live in a state with a Pre-K program at age 4 are 2.5 percentage points less likely to be repeating a grade in school (at ages 6-8), and have improved developmental outcomes as measured by a decrease in the development and behavior problems index (DBPI) of 9% of a standard deviation. Both effects are significant at a 5% significance level before adjusting for multiple hypothesis testing. After adjusting the p-values, the effect on grade repetition is still significant at a 5% level, while the effect on DBPI is only significant at a 10% level.

The estimated effect on the likelihood of currently repeating a grade is also negative five to eight years after preschool age, but close to zero and not statistically significant. Although I cannot construct a good cumulative measure of grade repetition using the CPS, looking at current grade repetition can also inform us about whether the effects of preschool education fade out over time.[29] The results show that there are decreases in grade repetition at the beginning of primary school and show no evidence of increased repetition later on, which implies long-lasting effects on children's school progression up to eight years after preschool age. This is a sign that pre-K programs improve the maturity with which children enter school. In addition, given that grade repetition can have permanent negative effects on long term outcomes such as educational attainment,[30] this could be a pathway to long-run effects of pre-K education on adult outcomes.

The effect on the index of development and behavior problems is also negative 5-8 after preschool

[28] For all outcome variables presented here, a negatively signed effect can be interpreted as beneficial. All regressions are estimated separately for each age-group, and include state and cohort fixed effects, individual controls (gender, race/ethnicity, and maternal education), state-level controls measured when the child was four years old (federally-funded Head Start enrollment rates, SCHIP/Medicaid eligibility income-to-poverty ratio thresholds for children under 5, annual unemployment rate, and annual state median income), and state-level controls measured in the year when the child is in the sample (SCHIP/Medicaid eligibility income-to-poverty ratio thresholds for children above 5, annual unemployment rate, and annual state median income).

[29] The CPS October questionnaire provides information about grade repetition in the current year, but no retrospective information on grade repetition. It is possible to construct an imprecise indicator of grade progression for age, using the age reported in October (birth dates are not reported). However, its measurement error is likely correlated with measurement error in the main explanatory variable, which is also constructed using the reported age to identify each child's pre-K cohort. The results of regressions using this measure of grade progression as an outcome show similar positive effects in the short and medium run, but they are both very imprecise and not statistically significant.

[30] Grade retention is associated with a higher probability of dropping out of school before finishing high school. Recent papers have attempted to disentangle the causal effects of grade retention as a policy on the children on the margin of grade promotion, and have found that it increases the likelihood of dropping out (Jacob and Lefgren, 2009; Manacorda, 2012; Eren, Depew and Barnes, 2017).

age and large in magnitude (5.3% of a standard deviation), but it is not statistically significantly different from zero. While this medium-run estimated effect is smaller than the short-run effect 1-4 years after preschool, the difference between the two is not statistically significant. Thus, I do not find conclusive evidence that the effect on this outcome fades out over time. To provide more evidence on the dynamics of the effects, Figure 3 shows graphs of estimated effects by 3-year overlapping age groups, in the spirit of a moving average. The top-left graph shows the estimates for grade repetition, which confirm that the estimate converges to zero over time but does not become positive. The estimates for the DBPI index, on the top-right, do not show a clear convergence to zero. The size of the estimated effect decreases in the first three years and is stable and negative afterwards, but the estimates are imprecise.

The first panel of Table 6 presents an exploratory analysis of the estimated effects of pre-K expansions on each individual indicator that is used to construct the DBPI index, for the pooled sample of both genders. These estimates should be taken with caution because of the increased multiple inference problem of looking at effects on more outcomes, and because of the small statistical power implied by the low mean percentage of children who are reported as having each of the specific problems indicated by these outcomes. With these caveats in mind, looking at the effects on the individual components of the index provides suggestive evidence on what types of developmental outcomes are affected by the pre-K expansions. The results suggests that the effects on the DBPI index are driven by negative effects on limitations related to intellectual disability, mental, emotional and behavioral problems, and learning disability, with significant effects only on the first two in the first four years after preschool.

In regards to health outcomes, the results in table 5 indicate that the implementation of pre-K programs has negative effects on health in the short run, as evidenced by a significant increase of 0.6 days in the number of missed days of school due to sickness in the first four years after preschool age. Consistent with this, respondents are more likely to report that the child's health status is fair or poor, although this effect is only significant at 10% significance level and is not robust to multiple inference adjustment. By the time children are 5-8 years after preschool age, however, there are no significant effects on health.[31] The short-term nature of these negative health effects is corroborated by the 3-year "moving-average" estimates presented in the bottom panel of Figure

[31]A potential concern when looking at reported health conditions is that incidence and/or awareness of some health problems may be sensitive to changes in access to health care. In analyses available upon request, I find that there are no statistically significant changes in the probability of having health insurance, or in an indicator for health care access problems. Not only are the estimates not significantly different from zero for these two variables, but the point estimates are also very small and quite precisely estimated.

3. The graph on the right shows the estimated effects for overlapping 3-year age groups on the number of missed days of school, which clearly converge to zero from the 4-6 age group onward. The left graph shows the estimates for the effects on the indicator of fair or poor reported health, which are only significant at 2-4 and 3-5 years after preschool, and are smaller and not statistically significant for older age groups.

To better understand the short-term effects on health, the second panel of Table 6 presents an exploratory analysis of the effects of pre-K programs on a group of health conditions commonly experienced by children, as well as having been hospitalized and having been to the emergency department due to an asthma attack in the past year. The results indicate no effects on most conditions, except for an increase in colitis/frequent diarrhea in the short run.[32] Given these results, we should also consider the possibility that part of the effect on missed days of school could be explained by an increase in the number of days missed *conditional* on having a health problem, for example due to changes in children's school performance or parents' work schedule. Whether the increase in the number of missed days of school is associated to worse health or not, it is a relevant outcome in and of itself, since it may lead to worse school performance (Goodman, 2014; Cattan et al., 2017) and potentially mitigate the improvements induced by pre-K education.

4.3 Heterogeneity of effects

In Table 7, I explore the heterogeneity of effects according to the characteristics of the pre-K programs and demographic characteristics of children. In general, there is no conclusive evidence of differential effects by program type and quality indicators. Panel 1 presents the results of adding to equation 1 the interaction of the Post Pre-K dummy with an indicator for universal programs. The estimated effects of the interaction are not statistically significant for any outcome, and although they are imprecisely estimated the signs of the coefficients show no indication of larger treatment effects for universal programs.[33]

In panel 2, I include the interaction of Post Pre-K with the quality score assigned to each

[32] The presence of frequent diarrhea could be associated to acute conditions (such as gastro-intestinal infections), or to chronic conditions (such as irritable bowel syndrome). Thus, the channels through which pre-K could be causing this may include not only the direct contagion effect of exposure to illness during pre-K, but also, for example, increased maternal labor force attachment or other changes in behavior and nutrition.

[33] The only other paper that contrasts the impacts of universal and targeted pre-K programs is Cascio (2017), which provides evidence of larger effects of universal programs on test scores at age 5 for low income children. While the universal programs in that study overlap substantially with those included in my sample, the targeted programs almost do not overlap at all, due to the different criteria we use for selecting the programs, and the states in our control groups also differ.

program in the school year 2005-2006 by NIEER.[34] The coefficient for the interaction indicates the marginal effect of having one more score point, assuming the effect is linear. For the sample of children observed 1-4 years after preschool age, none of the interaction effects are significant, although they have the expected sign. For children observed 5-8 years after preschool age, the estimates have a positive sign (indicating a deleterious effect) for the DBPI index and missed days of school. The absence of positive effects of this type of input quality indicators is perhaps not surprising. Walters (2015) finds that variation in some of the inputs included in the NIEER score (teacher qualifications and class size) have no effect on the effectiveness of Head Start centers, while providing full-day care and home visits do. Unfortunately I do not have detailed information on the last two, which in many cases vary locally. Full-day services are likely to be highly correlated with spending per student, although higher spending may also be related to increases in other inputs (such as reducing class size, teacher salaries, or providing additional services). Panel 3 presents the results of interacting the Post Pre-K indicator with the per-student state spending in pre-K in the 2005-2006 school year (Barnett et al., 2006). The estimates are not statistically significant for any of the outcomes.

The panels 4 and 5 of Table 7 explore the heterogeneity of effects by maternal education and child race/ethnicity groups. I do not find evidence of differential effects for children whose mothers have educational attainment of high-school or less. I do find, however, some evidence of differential effects on grade repetition by race/ethnicity. While the effects in the short run are not statistically different across groups, the estimated coefficient for the interaction of Post Pre-K and black is negative and statistically significant. Given that the estimated effect of Post Pre-K for the omitted group is also negative, this implies a statistically significant negative effect of pre-K on grade repetition for children of black ethnicity even in the years 5 to 8 after preschool age (of about 2 percentage points), indicating stronger long run effects for this group. On the other hand, the interaction of pre-K with hispanic is positive and statistically significant. Taken together with the negative main effect of pre-K for the omitted group, this could imply an overall increase in grade repetition in the years 5-8 after preschool that could mitigate the initial gains in school progression for this group. However, the sum of the two coefficients is not statistically significant so the evidence is not conclusive.[35]

[34] For a description of how the NIEER score is constructed, see Section 2.1.2. Among the group of programs in treatment states, the NIEER score is most correlated with the following indicators: at least one meal served (correlation of 0.79), teachers are specialized in pre-K education (0.7), teacher in service at least 15 hours per year (0.63), comprehensive learning standards (0.58), class size (0.53), staff:children ratio (0.53).

[35] The finding of sustained effects on grade repetition for blacks is consistent with the evidence presented by Deming

The last two panels of Table 7 show that there is no evidence of heterogeneous effects by gender when I estimate separate regressions for boys and girls. The effects on grade repetition and the DBPI index in the short run are not statistically different for boys and girls, although the point estimates and statistical significance are stronger for boys in grade repetition and for girls in the DBPI. Like in the pooled sample, the effects in the medium run are not statistically significant for either boys or girls. The estimated effects on health outcomes are also not statistically significantly different for boys and girls.

4.4 Robustness checks

This section explores the robustness of the results to potential threats to identification such as pre-existing trends, sensitivity to model specification and states included in the sample, and the correlation of the timing of implementation of pre-K programs with potential confounders.

The first four panels of Table 8 explore the possibility of differential trends across treatment and control states that may be confounding with the impacts of pre-K expansions. The first panel of shows that the results are robust to the inclusion of state-specific linear time trends. The second and third panels explore the sensitivity of the results to the inclusion of region-by-cohort and region-by-year fixed effects, respectively. The effects on grade repetition, the DBPI index, and the fair/poor health indicator are if anything stronger than in the main specification, while the effect on missed school days is attenuated at most by one third of the main specification estimate.

To explore the presence of pre-existing trends and heterogeneous effects in the years after the implementation, the fourth panel shows estimates of the following model:

$$Y^a_{isc} = \beta^a_{-1}\mathbb{1}(t = t_s - 1)_{sc} + \beta^a_0\mathbb{1}(t = t_s)_{sc} + \beta^a_{1+}\mathbb{1}(t > t_s)_{sc} + \gamma^a X^a_{isc} + \delta^a_c + \delta^a_s + \varepsilon^a_{isc} \tag{3}$$

where t_s is the year when a pre-K program was implemented in state s; $\mathbb{1}(t = t_s - 1)_{sc}$, $\mathbb{1}(t = t_s)_{sc}$, and $\mathbb{1}(t > t_s)_{sc}$ equal one if a child lives in a treatment state and belongs to the cohort that was age 4 in the year immediately preceding, the first year of, and one or more years after a pre-K program expansion was implemented, respectively. The estimates of β_{-1} are not statistically significantly different from zero for any of the outcomes, which supports the identification assumption of no pre-

(2009) and Ludwig and Miller (2007) for Head Start. Deming (2009) finds that Head Start reduced grade repetition and increases high school graduation for blacks, despite a fade out in effects on test scores for this group, and Ludwig and Miller (2007) find evidence of positive effects on educational attainment for both whites/hispanics and blacks. In contrast, Currie and Thomas (1995) find no effect on blacks' educational attainment, while Currie and Thomas (1999) find that Head Start reduces cumulative grade repetition by age 10+ for hispanics.

existing diverging trends between treatment and control groups. The estimates for β_0 and β_{1+} are not statistically significantly different, which goes against the hypothesis that the implementation of pre-K is a gradual process. If anything, the point estimates would indicate the effects are of a slightly larger magnitude in the first year of implementation. However, differences between these coefficients could also be caused by heterogeneity in pre-K effects across treatment states, since the effects after the first year of implementation are identified only from changes in the outcomes in states that implemented programs before the last year of the sample period.

Panels 5 and 6 of Table 8 explore the sensitivity of the results to control variables. In my main specification I include controls for the generosity of state public health insurance for children and for the enrollment in Head Start program, as well as state economic conditions and individual controls for demographic characteristics. In panel 4, I present the estimates of regressions that include no controls other than the state and cohort fixed effects, and show that the results are not sensitive to the inclusion of the control variables. Another policy that may be simultaneously affecting the preschool attendance and development and health outcomes of children is the Child Care Development Fund (CCDF).[36] Previous literature has found that receiving a childcare subsidy increases maternal supply but has negative effects on children's cognitive and behavioral outcomes, and increases the prevalence of child obesity.[37] Panel 5 shows the results when the model includes a control variable indicating the number of 4-year-olds served by the CCDF in each state. There is no information on the CCDF for the first year of my sample period, so these regressions are estimated with a smaller sample. However, the main results are all in line with the main estimates, with small differences in the sizes of the effects on the academic and development outcomes (of larger magnitude).

The last two panels show the sensitivity of my results to changes in the states included in the sample. In panel 6, I "switch off" the treatment variable for the four states that had the smaller enrollment increases after the implementation of their pre-K programs (6 percentage points or less). The results indicate that this group of states are not driving the conclusions. On the contrary, as

[36]The CCDF is a childcare subsidy program targeted towards low-income working families, implemented through a federal block grant. Although the federal government establishes some eligibility requirements, it gives states freedom to decide how to implement the subsidies. States can allocate TANF funds to the program, they can establish family income eligibility limits below the federal maximum, and they are responsible for determining eligibility controls, payment rates, and requirements for child care providers.

[37]Blau and Tekin (2007) show that child care subsidy receipt increases the labor supply of single mothers. Herbst and Tekin (2010) find that receiving the subsidy in the year before kindergarten is associated with lower reading and math test scores and greater behavior problems at kindergarten entry, for children from single mothers. Herbst and Tekin (2012) find that subsidized child care leads to increases in the prevalence of overweight and obesity among low-income children.

could be expected given the small size of these programs, omitting them from the treatment group increases the magnitude of the estimated effects, especially for grade repetition and missed days of school. In panel 7, I explore the robustness of the main results to the inclusion of all *excluded states* in the sample. Because these are states by the beginning of the sample period already have state pre-K programs, and enrollment rates of 4-year-olds in those states was increasing during the period, their inclusion in the control group should cause a bias toward finding no effects of pre-K programs. The results show that some of the estimated treatment effects are slightly attenuated, but none of the main conclusions are changed. Finally, to show that the results are not driven by any one particular outlier state, I re-estimate the main specification taking one state out of the sample at a time. The results from this exercise for short run outcomes are shown in Figure 4.[38] The point estimates for all the outcomes are fairly unchanged when each state is taken away from the sample. The only exception is the effect on health fair/poor, which is quite stable but has a smaller magnitude when North Carolina is excluded.

In Appendix C, I explore the correlation of the implementation of pre-K programs with other time-variant characteristics of the states, to address the concern of the possibility of other changes at the state level that may be correlated with pre-K expansions. In Table C1, I present the estimation results of regression models at the state-year level, where I predict the timing of the implementation of a pre-K program as a function of observable state characteristics, controlling for state and year fixed effects. The results indicate that the political parties in control of the state government and legislature do not play a significant role in the timing of the implementation of a pre-K program, and that investments in pre-K are not responding to the economic cycle. The pre-K expansions are not significantly associated with changes in expenditures in K-12 education or eligibility requirements for children's public health insurance programs. The year of implementation of pre-K is positively correlated with federally funded Head Start enrollment during the same year, but this association is throughout the years after pre-K implementation, mitigating the concern of confounding effects.

In Table C2, I explore the correlation of the implementation of pre-K programs with changes in state characteristics by running regressions at the individual level. I estimate regression models similar to Equation 2, but where the individual demographic controls and the state-level economic and policy variables outcome variables. Using individual-level information allows me to include the demographic characteristics of the children and their families, and to measure the state environment not only when children are four years old but also at the time when they are observed in the CPS

[38]Similar graphs for the effects 5-8 years after Pre-K available upon request.

sample. The results suggest that there are no diverging demographic, economic, or policy trends across treatment and control states after the implementation of pre-K. In particular, the estimated effect on Head Start enrollment at age 4 is not significantly different from zero, and the negative point estimate implies a negligible effect of 0.26 percentage points.

5 Approximating local average treatment effects

To answer policy-relevant questions in the context of the expansion of public provision of preschool education, one needs to take into account that children can be drawn into the new or expanded program from different alternative childcare environments, imposing different local average treatment effects. In this section, I present and discuss how to estimate or bound two parameters of interest under different alternative child care arrangements.[39] The intention-to-treat effects discussed before represent the average effect of introducing a typical pre-K program on the affected cohorts of children in the state. However, not all children in those cohorts actually have access to pre-K, not only because of non-compliance, but also because of the limited availability of spots in the programs and eligibility criteri. In this context, we may be interested in estimating the size of the effects of the implementation of a pre-K policy relative to the size of the expansion in pre-K coverage. In this sense, we are interested in the treatment-on-the-treated (TOT) effect, the average effect on the children actually affected by the pre-K expansion.

In broad terms, we can think of each 4-year-old as participating in one of three possible "treatments": State Pre-K, which I label k, other center-based preschool programs, denoted by c, and no preschool (i.e. home care or other forms of informal care), labeled n. Let $Z_i \in \{0, 1\}$ indicate whether child i lives in a state that has a pre-K program when she is four years old, and $D_i(z) \in \{k, c, n\}$ represent each child's potential treatment status as a function of the availability of a pre-K program in the child's state at age 4. To pin down treatment-on-the-treated effects departing from the reduced-from estimate, I need to make three assumptions. First, I assume there are no externalities on children who are not treated. Second, I make the assumption that anyone who changes their behavior as a response to the pre-K expansion does so to attend pre-K.[40] Fi-

[39] I follow a notation close to that used by Kline and Walters (2016). A similar discussion of relevant parameters of interest is present in García, Heckman and Ziff (2018).

[40] This implies that preferences across other modes of childcare are not changed because of the increased availability of pre-K. It also implies assuming that other preschool programs are not rationed. If there is excess demand for other preschool programs, it is possible that when children switch from the other programs to state pre-K, they open up slots for other children who would otherwise not be attending preschool. Rationing is likely to occur in Head Start programs. This would be a minor issue if the proportion of children drawn from Head Start to pre-K is relatively small, or if the treatment effects of pre-K are similar to those of Head Start.

nally, I assume that the only differences in treatment effects across individuals are related to the alternative mode of childcare, and not other individual differences.

We can partition the population of children into the following groups: n-compliers, c-compliers, n-never takers, c-never takers, and always takers. n-compliers are children who switch to state pre-K if a pre-K expansion is implemented in their states, but would otherwise not be enrolled in preschool ($D_i(1) = k$, $D_i(0) = n$). Similarly, the c-compliers are children who switch to state pre-K if a pre-K expansion is implemented in their states, but would otherwise be enrolled in other (private or public) preschool programs ($D_i(1) = k$, $D_i(0) = c$). Never takers are children who do not switch to state pre-K even if a pre-K expansion is implemented in their states, and instead stay enrolled in other preschool (c-never takers, $D_i(1) = D_i(0) = c$) or not enrolled in any program (n-never takers, $D_i(1) = D_i(0) = n$).[41] Finally, always takers are children who would be enrolled in state pre-K even in the absence of a pre-K expansion ($D_i(1) = D_i(0) = k$).

One parameter of interest is the local average treatment effect (LATE) considering as "treated" all of the compliers, with their corresponding mix of alternative child care arrangements. Under the assumption of no externalities, this LATE is equivalent to the ratio of the reduced-form effect of receiving the treatment on an outcome variable, and the first-stage effect of receiving the treatment on participation in pre-K (β_{LATE} hereafter):

$$\beta_{LATE} = \frac{\mathrm{E}[Y_i|Z_i = 1] - \mathrm{E}[Y_i|Z_i = 0]}{\mathrm{E}[1\{D_i = k\}|Z_i = 1] - \mathrm{E}[1\{D_i \neq k\}|Z_i = 0]} \tag{4}$$

A Wald estimator for β_{LATE} (for each outcome variable and age group) would be the ratio of the reduced-form estimator ($\hat{\beta}_{RF}$) and a first-stage estimation of the effect of the pre-K expansion on pre-K enrollment. I approximate this first-stage effect with the proxy for the change in enrollment rates in pre-K in treatment states before and after the expansions presented in Section 4.1. I estimate the following as an approximation of β_{LATE}:[42]

$$\hat{\beta}_{LATE} = \frac{\hat{\beta}_{RF}}{\Delta \text{Pre-K}} \tag{5}$$

[41]Note that in this setting, the definition of never takers does not imply that a child has the possibility of attending pre-K and does not take it, because the treatment here is not actually being offered a pre-K spot but instead just the implementation of a pre-K expansion in the state (which does not guarantee a spot in the pre-K program). Thus, the groups of n-never takers and c-never takers also include children who live in a state with a pre-K policy but who do not have the possibility of enrolling in the program.

[42]The use of this proxy as a substitute for the first-stage estimate introduces additional uncertainty to the LATE estimate, but I do not have a measure of this uncertainty. Thus, the reported standard errors for $\hat{\beta}_{LATE}$ are only based on the uncertainty of the reduced-form estimate $\hat{\beta}_{RF}$. In particular, $\hat{se}_{\hat{\beta}_{LATE}} = \sqrt{(\hat{se}_{\hat{\beta}_{RF}}/\Delta \text{ Pre-K Enrollment})^2}$.

The LATE can be decomposed, as pointed out by Kline and Walters (2016), into a weighted average of local average treatment effects for the two groups of compliers:

$$\beta_{LATE} = S_c \beta_{LATE_{ck}} + (1 - S_c)\beta_{LATE_{nk}} \tag{6}$$

where $\beta_{LATE_{ck}} \equiv \mathrm{E}[Y_i(k) - Y_i(c)|D_i(1) = k, D_i(0) = c]$ is the average treatment effect of attending pre-K on c-compliers (those drawn from other preschool programs); $\beta_{LATE_{nk}} \equiv \mathrm{E}[Y_i(k) - Y_i(n)|D_i(1) = k, D_i(0) = n]$ is the average treatment effect of attending pre-K on n-compliers (those drawn from non-preschool childcare arrangements); and $S_c \equiv \frac{\Pr(D_i(1)=k, D_i(0)=c)}{\Pr(D_i(1)=k, D_i(0)\neq k)}$ represents the fraction of compliers drawn from other preschool programs.

The parameter $\beta_{LATE_{nk}}$ is also of interest, as it represents the average treatment effect of attending pre-K compared to not attending preschool at all. This parameter cannot be directly estimated with the available data, but it can be bounded under the assumptions stated before and the additional assumption that pre-K impacts are of the same sign and at least as large as the impacts of attending other preschools compared to not attending preschool. From equation 6, it follows that $\beta_{LATE_{nk}} = \beta_{LATE} + S_c(\beta_{LATE_{nk}} - \beta_{LATE_{ck}})$. Thus, $\beta_{LATE_{nk}}$ can be bounded by bounding the difference between $\beta_{LATE_{nk}}$ and $\beta_{LATE_{ck}}$.

For simplicity of exposition, assume both of these LATEs are positive.[43] Under the assumption that the effect of pre-K is at least as large as the effect of attending the next best preschool, $\beta_{LATE_{ck}}$ is bounded between 0, in the case that the two programs are of identical quality, and $LATE_{nk}$, in the case that the next best alternative preschool has no effect relative to home care. Therefore, when $\beta_{LATE_{ck}} = \beta_{LATE_{nk}}$, β_{LATE} provides a lower bound for $\beta_{LATE_{nk}}$:

$$\beta^{Lower}_{LATE_{nk}} = \beta_{LATE} \tag{7}$$

On the other extreme, when $\beta_{LATE_{ck}} = 0$, we have the following upper bound: $\beta^{Upper}_{LATE_{nk}} = \frac{1}{1-S_c} \times \beta_{LATE}$, which can be re-written as:[44]

$$\beta^{Upper}_{LATE_{nk}} = \frac{\mathrm{E}[Y_i|Z_i = 1] - \mathrm{E}[Y_i|Z_i = 0]}{\mathrm{E}[1\{D_i \neq n\}|Z_i = 1] - \mathrm{E}[1\{D_i = n\}|Z_i = 0]} \tag{8}$$

[43] An analogous argument can be made when both pre-K and the next best alternative preschool have negative effects relative to home care.

[44] Re-writing $\frac{1}{1-S_c}$ as $\frac{\mathrm{E}[1\{D_i=k\}|Z_i=1]-\mathrm{E}[1\{D_i\neq k\}|Z_i=0]}{\mathrm{E}[1\{D_i=k\}|Z_i=1]-\mathrm{E}[1\{D_i=n\}|Z_i=0]}$, and substituting β_{LATE} for the expression in Equation 4, we have that $\beta^{Upper}_{LATE_{nk}} = \frac{\mathrm{E}[Y_i|Z_i=1]-\mathrm{E}[Y_i|Z_i=0]}{\mathrm{E}[1\{D_i=k\}|Z_i=1]-\mathrm{E}[1\{D_i=n\}|Z_i=0]}$. Under the assumption that the only changes in childcare arrangements induced by a pre-K expansion are changes to enroll in pre-K (no switches between home (n) and other preschools (c)), $\mathrm{E}[1\{D_i = k\}|Z_i = 1] = \mathrm{E}[1\{D_i \neq n\}|Z_i = 1]$, and Equation 8 follows.

The denominator in this expression represents the change in preschool enrollment (including pre-K and other preschool programs) caused by the introduction of a pre-K expansion, which is larger than the denominator in Equation 4 if there is some degree of crowding-out. Note that if there is no crowding-out from other preschools, then the upper and lower bounds are identical.[45] The intuition behind the upper bound is that, if we think that pre-K and other preschool programs are equivalent ($\beta_{LATE_{ck}} = 0$), $\beta_{LATE_{nk}}$ would represent the treatment effect of preschool. If instead state pre-K constitutes a higher quality option than the average alternative preschool program, then $\beta_{LATE_{ck}} > 0$, and $\beta^{Upper}_{LATE_{nk}}$ overestimates the effect of attending pre-K because it does not include in the denominator those children who are switching from other preschools to pre-K, when making this change has an impact on them.

I estimate $\beta^{Upper}_{LATE_{nk}}$ using a Two-Sample Two-Stage Least Squares strategy (TS-2SLS), using as first stage the effect of pre-K expansions on preschool enrollment (Equation 2). Because the model is exactly identified, this TS-2SLS estimator is the ratio of the reduced-form estimator ($\hat{\beta}_{RF}$) and the first-stage estimator ($\hat{\beta}_{FS}$, the estimate of β in Equation 2):[46]

$$\hat{\beta}^{Upper}_{LATE_{nk}} = \frac{\hat{\beta}_{RF}}{\hat{\beta}_{FS}} \tag{9}$$

Table 9 presents the results. Each panel shows the results for a different outcome variable, for the sub-samples of children observed 1-4 and 5-8 years after pre-K age. The first column repeats the ITT estimates presented in Table 5. The second column, presents the estimates of β_{LATE}, where the ITT estimates are re-scaled by the increase in pre-K enrollment induced by the pre-K expansions. The last column shows the estimates of $\beta^{Upper}_{LATE_{nk}}$, where the ITT estimates are re-scaled by the increase in any preschool attendance induced by the pre-K expansions.

As explained above, to interpret any of these estimates as approximation of different TOT effect parameters, we need to assume that there are no externalities or general equilibrium effects. This assumption is particularly problematic for the health outcomes examined, where contagion across

[45] As discussed in Section 4.1, the difference between the observed increase in total preschool enrollment and in pre-K enrollment is suggestive of the presence of crowding-out and/or simultaneous participation in pre-K and other preschool programs.

[46] I compute the first-stage estimates using CPS data on 4-year-olds, and the reduced-form estimates using NHIS data separately for each age group. To estimate $LATE^{upper}_{nk}$, I re-estimate the second stage only including control variables that are constant over time, as well as cohort and state fixed effects, because all control variables must be the same in both stages of the model. In practice this means that I do not include in the second stage any of the variables evaluated in the year of observation. Following Dee and Evans (2003), I compute standard errors using the delta method, assuming there is zero covariance between the first-stage and reduced form-estimates. Under this assumption, the delta method implies that the standard errors of the TS-2SLS estimator can be approximated by: $\hat{se}_{\hat{\beta}_{TOT2}} = \sqrt{(\hat{\beta}^2_{RF}/\hat{\beta}^2_{FS}) * [(\hat{se}_{\hat{\beta}_{FS}}/\hat{\beta}_{FS})^2 + (\hat{se}_{\hat{\beta}_{RF}}/\hat{\beta}_{RF})^2]}$

children that attended and did not attend pre-K (and even children of different cohorts) is very likely to play a role. For this reason, while I present estimates for all outcomes, I only focus the discussion of the estimates for the academic and development outcomes. If anything, the estimates for missed days of school could be taken as an upper bound of the individual effects. Nevertheless, the estimates for the academic and development outcomes should also be interpreted with caution, as it is also possible that there may be peer effects for grade repetition and the DBPI index as well.[47]

The estimates of β_{LATE} indicate that attending pre-K, compared to the actual mix of alternative childcare arrangements, reduces the likelihood of repeating a grade at ages 6-8 by 13.6 percentage points. This is a very large effect; it is more than 1.5 times the average likelihood of repeating a grade for the full sample, although it is possible that compliers could have a larger likelihood of repeating than the average, and it is 49% of its standard deviation. It should also be noted that the 95% confidence interval is large, and I cannot reject an effect as low as 2.6 percentage points. The estimated effect for the DBPI index implies that attending pre-K reduces development and behavior problems in the next four years by 49% of a standard deviation of the full sample. In this case the 95% confidence interval is also very large due to the imprecision of the estimates, so I cannot rule out treatment effects very close to zero.

These estimates also provide lower bounds for the treatment effects of pre-K relative to home care under the additional assumptions discussed above. The estimated upper bounds are presented in the last column. These estimates imply that attending a Pre-K program could reduce grade repetition up to 124% of a standard deviation (34.2 percentage points), and reduce development and behavioral problems up to 123% of a standard deviation of the DBPI, during the first four years after preschool age, compared to not attending any preschool program. These estimated effects are large, but are not far from previous estimates in the literature for Head Start.[48]

[47] For example, the DBPI index is constructed with indicators of reported daily life limitations related to different development and behavior problems. It is possible that the perception of these limitations may be shaped by the behavior that parents observe in other children of their child's cohort. Similarly, a teacher's recommendation that a child should repeat a grade may be based on the child's academic progress compared to other children in the same cohort. It is possible that when children that attended pre-K do better in terms of academic progress or behavior, other similar children are perceived as doing worse in relative terms, even if they objective behavior has not changed. On the other hand, it is also possible that other children's outcomes are actually improved by being in school with children whose academic progress and behavior is improved by having attended pre-K, through positive peer effects.

[48] For example, Currie and Thomas (1995) estimate that participation in Head Start reduces the likelihood of ever having repeated a grade by 47 percentage points for white children at age 10 or higher, relative to not attending any preschool. Deming (2009) finds that Head Start participation improves an index of cognitive outcomes by 39% of a standard deviation for boys 3 to 10 years after preschool age, again relative to not attending preschool. Carneiro and Ginja (2014) estimate that Head Start participation leads to a 60% of a standard deviation decrease in a behavior problems index for boys at ages 12-13. Finally, Kline and Walters (2016) find that re-scaling the HSIS reduced-form

Finally, it should be noted that this interpretation ignores any intensive margin or income effects on children who would attend preschool in any case. For families that substitute public pre-K for private child care, the public provision of pre-K may imply an increase in disposable income, which may in turn positively affect children. In addition, children may be simultaneously enrolled in more than one preschool program, as many programs are only part-day. Consequently, pre-K expansions may increase the amount of hours per week that children are exposed to preschool education (intensive margin), rather than just the actual attendance to preschool (extensive margin). The group of children potentially affected by pre-K expansions through the intensive margin and income effects includes children who would otherwise be attending some preschool program anyway. These are additional reasons why we would expect the treatment effect of attending Pre-K relative to no preschool ($\beta_{LATE_{nk}}$) to be below $\hat{\beta}^{Upper}_{LATE_{nk}}$. Even when the imprecision of the estimates does not allow me to rule out smaller effect sizes, the generally large magnitude of the point estimates suggest that the relevant "treatment" is not just preschool attendance, and that children who otherwise would be enrolled in other preschool programs may also benefit from the expansion of state pre-K education.

6 Conclusion

In this paper, I study the impacts of expansions of state-funded pre-K programs on grade repetition, development and behavior problems, and child health, one to eight years after preschool age. I leverage the differential timing of introduction of pre-K programs across states, paired with individual-level repeated cross-sectional data from two national surveys and state-level data from various sources, to estimate effects of pre-K programs in regression models with state and cohort fixed effects. I find very robust evidence of positive short-run impacts on children's academic and developmental outcomes. The beneficial short-run effects on grade progression are not reversed by age 12, suggesting a potential pathway for long-term impacts on educational achievement. I also find suggestive evidence of a negative impact on child health in the short run, as evidenced by an increase in the number of missed days of school due to illness in the first one to four years after preschool age.

These results complement the previous literature in several ways. First, this is the first evidence of positive short-run impacts of pre-K on academic and developmental outcomes other than

estimates of the effects of Head Start by the first-stage effect of random assignment on participation implies a TOT effect of 25% of a standard deviation in test scores in the first year, compared to the compliers' alternative child care arrangements. Their estimate of the LATE of moving from home care to Head Start is around 37% of a s.d.

standardized achievement tests, allowing to have a more complete picture of these programs' impacts. Furthermore, my results do not provide support for the hypothesis that these effects fade out after the first couple of years after preschool, which is consistent with previous literature that has found sustained impacts throughout childhood on similar child outcomes for Head Start and for universal preschool programs in other countries. Additionally, this is the first study to estimate health effects of state pre-K programs, finding deleterious short-run effects that are in line with the findings in papers that study the effects of child care subsidies, but against some of the findings of the literature on the federal Head Start program. This raises questions for future research in terms of the channels that may explain these effects and the role that specific program characteristics can play to prevent these effects.

In addition to providing intention-to-treat estimates on full cohorts of children, I also provide a discussion of two local average treatment effects under different counterfactual child care arrangements. I find implied effects that are very large, although not very different from the estimates found in the literature on Head Start. This finding highlights the relevance of estimating intention-to-treat effects on the full affected cohorts of children, which do not rely on any assumptions regarding who are the affected children, and particularly on whether there are any externalities on children that do not attend the programs. The relevance of externalities and peer effects in early childhood experience is an open question that is part of a promising research agenda. For example, recent research by Carrell, Hoekstra and Kuka (2018) shows long-run impacts of having disruptive peers on test scores, educational attainment and adult earnings. To the extent that early education experiences can affect children's academic and behavioral outcomes in the first years of school, there is potential for propagation of these positive effects on their peers.

References

Aizer, Anna, and Flavio Cunha. 2012. "The production of human capital: Endowments, investments and fertility." Working Paper 18429, National Bureau of Economic Research.

Anderson, Patricia M, Kristin F Butcher, and Phillip B Levine. 2003. "Maternal employment and overweight children." *Journal of health economics*, 22(3): 477–504.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. "Universal child care, maternal labor supply, and family well-being." *Journal of Political Economy*, 116(4): 709–745.

Ball, Thomas M, Jose A Castro-Rodriguez, Kent A Griffith, Catharine J Holberg, Fernando D Martinez, and Anne L Wright. 2000. "Siblings, day-care attendance, and the risk of asthma and wheezing during childhood." *New England Journal of Medicine*, 343(8): 538–543.

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

Barnett, W Steven, Kenneth B. Robin, Jason T. Hustedt, and Karen L. Schulman. 2003. *The State of Preschool: 2003 State Preschool Yearbook.*: National Institute for Early Education Research.

Belfield, Clive R, Milagros Nores, Steve Barnett, and Lawrence Schweinhart. 2006. "The High/Scope Perry preschool program: Cost-benefit analysis using data from the age-40 followup." *Journal of Human Resources*, 41(1): 162–190.

Berlinski, Samuel, Sebastian Galiani, and Paul Gertler. 2009. "The effect of pre-primary education on primary school performance." *Journal of Public Economics*, 93(1-2): 219–234.

Berlinski, Samuel, Sebastian Galiani, and Marco Manacorda. 2008. "Giving children a better start: Preschool attendance and school-age profiles." *Journal of Public Economics*, 92(5-6): 1416–1440.

Black, Sandra E, Paul J Devereux, Katrine V Løken, and Kjell G Salvanes. 2014. "Care or cash? The effect of child care subsidies on student performance." *Review of Economics and Statistics*, 96(5): 824–837.

Blau, David, and Erdal Tekin. 2007. "The determinants and consequences of child care subsidies for single mothers in the USA." *Journal of Population Economics*, 20(4): 719–741.

Bloom, Howard S. 1984. "Accounting for no-shows in experimental evaluation designs." *Evaluation review*, 8(2): 225–246.

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): 135–73.

Carrell, Scott E, Mark Hoekstra, and Elira Kuka. 2018. "The long-run effects of disruptive peers." *American Economic Review*, in press.

Cascio, Elizabeth U. 2017. "Does Universal Preschool Hit the Target? Program Access and Preschool Impacts." Working Paper 23215, National Bureau of Economic Research.

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), Brookings Institution.

Cattan, Sarah, Daniel A Kamhöfer, Martin Karlsson, and Therese Nilsson. 2017. "The short-and long-term effects of student absence: evidence from Sweden." Working Paper W17/21, Institute for Fiscal Studies.

Cunha, Flavio, and James Heckman. 2007. "The technology of skill formation." *American Economic Review*, 97(2): 31–47.

Cunha, Flavio, James J Heckman, Lance Lochner, and Dimitriy V Masterov. 2006. "Interpreting the evidence on life cycle skill formation." *Handbook of the Economics of Education*, 1 697–812.

Cunha, Flavio, James J Heckman, and Susanne M Schennach. 2010. "Estimating the technology of cognitive and noncognitive skill formation." *Econometrica*, 78(3): 883–931.

Currie, Janet. 2001. "Early childhood education programs." *Journal of Economic Perspectives*, 15(2): 213–238.

Currie, Janet, and Duncan Thomas. 1995. "Does Head Start make a difference?" *The American Economic Review*, 85(3): , p. 341.

Currie, Janet, and Duncan Thomas. 1999. "Does Head Start help Hispanic children?" *Journal of Public Economics*, 74(2): 235–262.

Dahl, Gordon B, and Lance Lochner. 2012. "The impact of family income on child achievement: Evidence from the earned income tax credit." *American Economic Review*, 102(5): 1927–56.

De Chaisemartin, Clément, and Xavier D'HaultfŒuille. 2018. "Fuzzy differences-in-differences." *The Review of Economic Studies*, 85(2): 999–1028.

Dee, Thomas S, and William N Evans. 2003. "Teen drinking and educational attainment: evidence from two-sample instrumental variables estimates." *Journal of Labor Economics*, 21(1): 178–209.

Deming, David. 2009. "Early childhood intervention and life-cycle skill development: Evidence from Head Start." *American Economic Journal: Applied Economics*, 1(3): 111–34.

Education Comission of the States. 2015. *Education Commission of the States (ECS) State Policy Database*. https://www.ecs.org/state-legislation-by-state/, Accessed February 3, 2015.

Eren, Ozkan, Briggs Depew, and Stephen Barnes. 2017. "Test-based promotion policies, dropping out, and juvenile crime." *Journal of Public Economics*, 153 9–31.

Falk, Gene. 2013. "Temporary Assistance for Needy Families: Welfare Waivers." CRS Report for Congress R42627, Congressional Research Service.

Felfe, Christina, Natalia Nollenberger, and Núria Rodríguez-Planas. 2015. "Can't buy mommy's love? Universal childcare and children's long-term cognitive development." *Journal of Population Economics*, 28(2): 393–422.

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

Friedman-Krauss, Allison H, W Steven Barnett, GG Weisenfeld, Richard Kasmin, Nicole DiCrecchio, and Michelle Horowitz. 2018. *The State of Preschool 2017: State Preschool Yearbook.*: National Institute for Early Education Research.

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

García, Jorge Luis, James J Heckman, and Anna L Ziff. 2018. "Gender differences in the benefits of an influential early childhood program." *European Economic Review*, in press.

Gennetian, Lisa A, Heather D Hill, Andrew S London, and Leonard M Lopoo. 2010. "Maternal employment and the health of low-income young children." *Journal of Health Economics*, 29(3): 353–363.

Goodman, Joshua. 2014. "Flaking out: Student absences and snow days as disruptions of instructional time." Working Paper 20221, National Bureau of Economic Research.

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.

Gormley, William T, Deborah Phillips, and Ted Gayer. 2008. "Preschool Programs Can Boost School Readiness." *Science*, 320(5884): 1723–1724.

Havnes, Tarjei, and Magne Mogstad. 2011. "No child left behind: Subsidized child care and children's long-run outcomes." *American Economic Journal: Economic Policy*, 3(2): 97–129.

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

Heckman, James J, and Stefano Mosso. 2014. "The economics of human development and social mobility." *Annual Review of Economics*, 6(1): 689–733.

Herbst, Chris M, and Erdal Tekin. 2010. "Child care subsidies and child development." *Economics of Education Review*, 29(4): 618–638.

Herbst, Chris M, and Erdal Tekin. 2012. "The geographic accessibility of child care subsidies and evidence on the impact of subsidy receipt on childhood obesity." *Journal of Urban Economics*, 71(1): 37–52.

Hill, Carolyn J, William T Gormley, and Shirley Adelstein. 2015. "Do the short-term effects of a high-quality preschool program persist?" *Early Childhood Research Quarterly*, 32 60–79.

Holm, Sture. 1979. "A simple sequentially rejective multiple test procedure." *Scandinavian Journal of Statistics* 65–70.

Jacob, Brian A, and Lars Lefgren. 2009. "The effect of grade retention on high school completion." *American Economic Journal: Applied Economics*, 1(3): 33–58.

Kaiser Commission on Medicaid and the Uninsured. 2006-2014. "A 50 State Update on Eligibility Rules, Enrollment and Renewal Procedures, and Cost-sharing Practices in Medicaid and SCHIP." Henry J. Kaiser Family Foundation. https://www.kff.org/medicaid/report/annual-updates-on-eligibility-rules-enrollment-and/.

Kline, Patrick, and Christopher R Walters. 2016. "Evaluating public programs with close substitutes: The case of Head Start." *The Quarterly Journal of Economics*, 131(4): 1795–1848.

Lipsey, Mark W, Dale C Farran, and Kelley Durkin. 2018. "Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade." *Early Childhood Research Quarterly*, in press.

Løken, Katrine V, Magne Mogstad, and Matthew Wiswall. 2012. "What linear estimators miss: The effects of family income on child outcomes." *American Economic Journal: Applied Economics*, 4(2): 1–35.

Ludwig, Jens, and Douglas L Miller. 2007. "Does Head Start improve children's life chances? Evidence from a regression discontinuity design." *The Quarterly Journal of Economics*, 122(1): 159–208.

Magnuson, Katherine A, Marcia K Meyers, and Jane Waldfogel. 2007. "Public funding and enrollment in formal child care in the 1990s." *Social Service Review*, 81(1): 47–83.

Manacorda, Marco. 2012. "The cost of grade retention." *Review of Economics and Statistics*, 94(2): 596–606.

McFarland, J., B. Hussar, X. Wang, J. Zhang, K. Wang, A. Rathbun, A. Barmer, E. Forrest Cataldi, and F Bullock Mann. 2018. "The Condition of Education 2018." (NCES 2018-144) U.S. Department of Education. Washington, DC: National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018144, Accessed July 12, 2018.

Morrill, Melinda Sandler. 2011. "The effects of maternal employment on the health of school-age children." *Journal of Health Economics*, 30(2): 240–257.

NGA Center for Best Practices. 1997-2011. "*Maternal and Child Health (MCH) Update*." National Governors Association. http://www.nga.org.

Puma, Michael, Stephen Bell, Ronna Cook, Camilla Heid, Gary Shapiro, Pam Broene, Frank Jenkins, Philip Fletcher, Liz Quinn, Janet Friedman et al. 2010. "Head Start Impact Study. Final Report." Research Report, U.S. Department of Health and Human Services, Administration for Children and Families.

Puma, Mike, Stephen Bell, Ronna Cook, Camilla Heid, Pam Broene, Frank Jenkins, Andrew Mashburn, and Jason Downer. 2012. "Third Grade Follow-Up to the Head Start Impact Study: Final Report." OPRE Report 2012-45, U.S. Department of Health and Human Services, Administration for Children and Families.

Romano, Joseph P, and Michael Wolf. 2005. "Exact and approximate stepdown methods for multiple hypothesis testing." *Journal of the American Statistical Association*, 100(469): 94–108.

University of Kentucky Center for Poverty Research. 2017. "UKCPR National Welfare Data, 1980-2016." Gatton College of Business and Economics, University of Kentucky, Lexington, KY. http://www.ukcpr.org/data. Accessed May 3, 2018.

US Department of Health and Human Services. 2018. *Head Start program facts: Fiscal year 2017.*: Administration for Children and Families, Office of Head Start..

Walters, Christopher R. 2015. "Inputs in the production of early childhood human capital: Evidence from Head Start." *American Economic Journal: Applied Economics*, 7(4): 76–102.

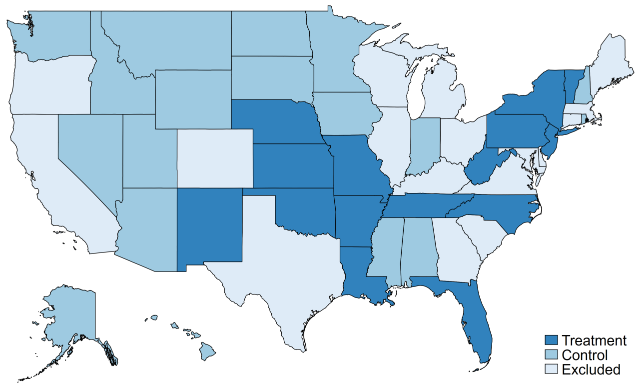
Weiland, Christina, and Hirokazu Yoshikawa. 2013. "Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills." *Child Development*, 84(6): 2112–2130.

Wong, Vivian C, Thomas D Cook, W Steven Barnett, and Kwanghee Jung. 2008. "An effectiveness-based evaluation of five state pre-kindergarten programs." *Journal of Policy Analysis and Management*, 27(1): 122–154.

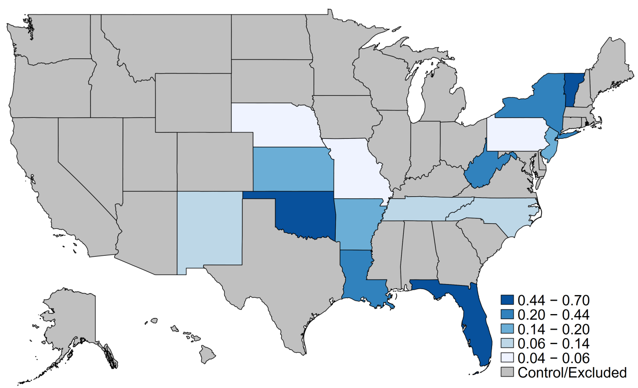
Figures and Tables

Figure 1

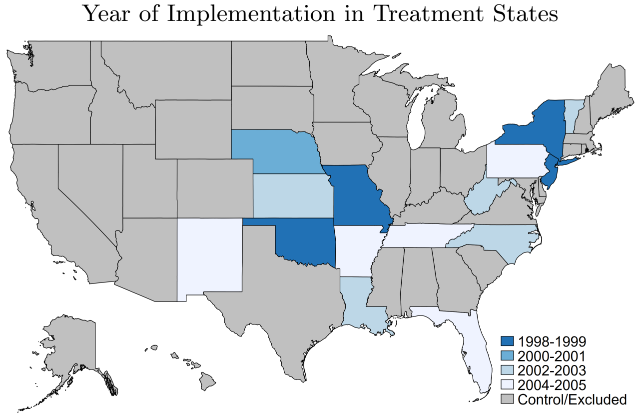
Treatment, Control, and Excluded States



Pre-K Enrollment in Treatment States (2005)

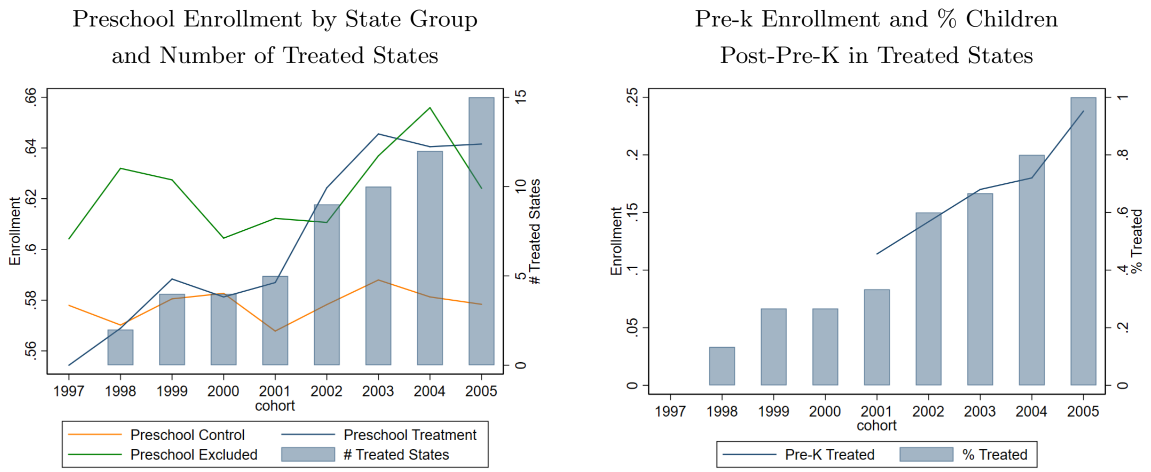


Year of Implementation in Treatment States



Notes: Treatment indicates that a state implemented or significantly expanded pre-K between 1998 and 2005. Control indicates that a state had not yet implemented or had only a very small state pre-K program by 2005. *Excluded* indicates that a state is excluded from the main sample because it already had a pre-K program by 1997. Pre-K enrollment rates in 2005-2006 school year were obtained from NIEER (2006). Year of implementation indicates the first 4-year-old cohort exposed to a Pre-K program in treatment states.

Figure 2: Preschool Enrollment and Pre-K Availability

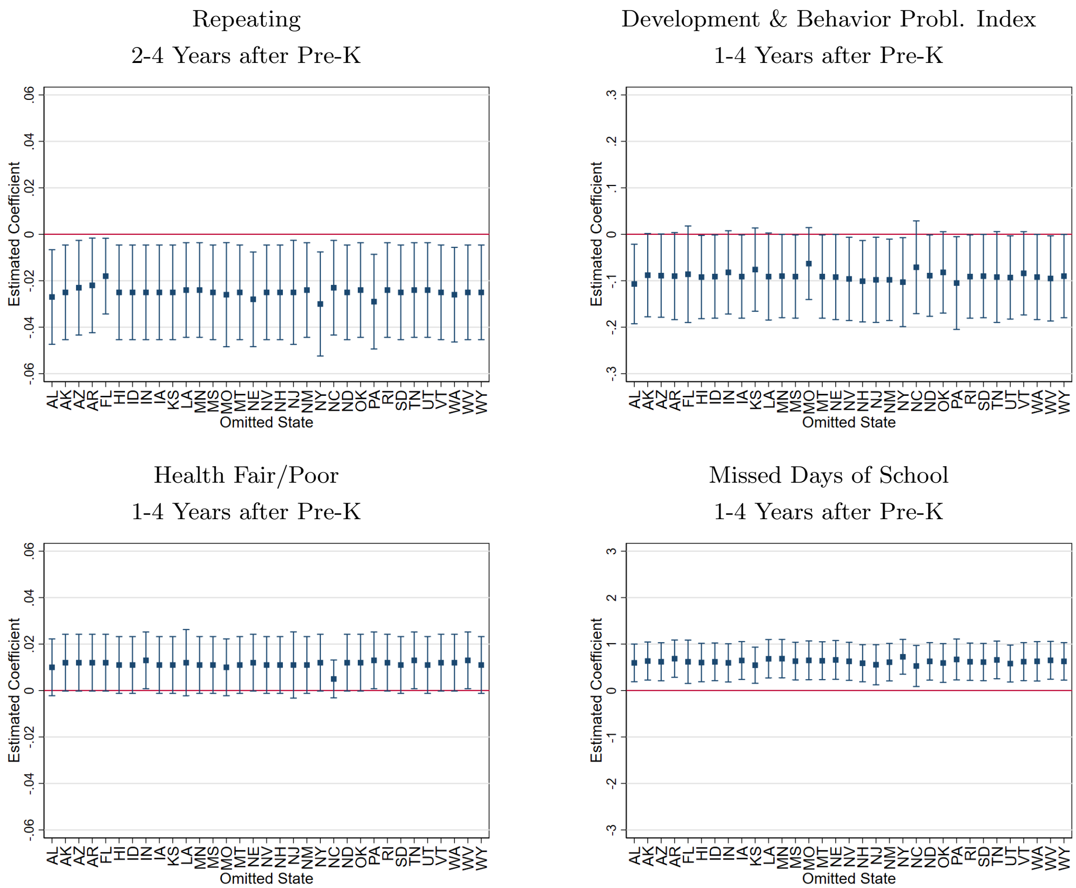


Notes: The graph on the left plots 2-year moving averages of preschool enrollment of 4-year-olds in the CPS October supplement 1997-2005 in treatment, control and excluded states (left axis). The first data point is estimated using only 1997. The bars indicate the number of treated states (states that had started a pre-K program) among treated states, by cohort (right axis). The graph on the right presents aggregate enrollment in state pre-K programs from NIEER reports, weighted using CPS sampling weights of 4-year-olds in the October supplement (left axis). The bars indicate the percentage of 4-year-olds treated (i.e. living in a state that has started a pre-K program) by cohort of the CPS October samples of 4-year olds (right axis).

Figure 3: Reduced-Form Effects for Moving Averages of Number of Years after Pre-K

Notes: The figure plots the reduced-form estimates of the effects of a pre-K expansion from separate regressions for samples of 3-year moving averages of number of years after preschool age, for pooled samples of both genders. The shaded region is the 95% confidence interval based on standard errors clustered at state level and Student's t-distribution with 32 degrees of freedom.

Figure 4: Sensitivity of Reduced-Form Results to Each State in Sample



Notes: Reduced-form estimate and 95% confidence interval of the effects of a pre-K expansion when excluding one state at a time from the main sample. Each point estimate is obtained in a separate regression, for the pooled sample of both genders, for the outcome indicated in the column headings, observed 1-4 years after pre-K (or 2-4 years after pre-K for grade repetition).

Table 1: Descriptive statistics for program characteristics

	Pop.-Weighted	Unweighted		
	Mean	Mean	Median	Std. Dev.
Universal (1=yes)	0.42	0.33	0.00	0.49
Enrollment (%)	24.34	23.80	18.00	19.46
Universal programs	40.01	46.60	47.00	15.01
Targeted programs	12.78	12.40	11.50	7.47
Full day (1=yes)	0.32	0.33	0.00	0.49
Full/part day det. locally (1=yes)	0.32	0.27	0.00	0.46
Per student spending	4442.06	4848.07	4061.00	2412.41
NIEER Score	6.29	6.87	7.00	2.23
Comprehensive learning standards	0.76	0.87	1.00	0.35
Teacher degree BA	0.45	0.60	1.00	0.51
Teacher specialized in pre-K	0.53	0.67	1.00	0.49
Assistant teacher CDA or equiv.	0.17	0.27	0.00	0.46
Teacher in-service training $\geq$ 15 hs/yr	0.65	0.67	1.00	0.49
Max. class size $\leq$ 20	0.97	0.93	1.00	0.26
Staff-child ratio $\geq$ 1:10	0.97	0.93	1.00	0.26
Vision, hearing, health checkups; support services	0.76	0.80	1.00	0.41
At least 1 meal per day	0.50	0.53	1.00	0.52
Required monitoring (site visits)	0.78	0.73	1.00	0.46

Notes: All indicators are based on information by state program reported by NIEER's The State of Preschool 2006 (Barnett et al., 2006) and refer to the 2005-2006 school year. Population-weighted averages are computed using the population of 4-year-olds by state from the U.S. Census. Universal is an indicator for whether the program does not have any income eligibility requirements. Enrollment rate statistics are reported for all treatment states and separately for universal and targeted programs. *Full day* indicates that all centers in the state program provide at least 5 hours per day of services. *Full/part day det. locally* indicates that centers may operate full day or not depending on local decisions. Per student spending is the total spending per enrollee in the (including state funding, local matching funds, and federal grants administered by the state). NIEER Score is the number of NIEER quality standards met by the program in 2005-2006 (out of 10). All of the indicators used to compute the score are listed thereafter, and they take a value of 1 when the condition described is a requirement of the state pre-K policy. The last item on the list is an indicator of whether states are taking steps to monitor programs' implementation of the quality standards. CDA stands for Child Development Associate.

Table 2: Descriptive Statistics for the CPS Sample of 4-year-olds in Treatment, Control and Excluded States

	Treatment States		Control States		Excluded States	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Individual Characteristics						
Female	0.48	0.50	0.48	0.50	0.50	0.50
Black	0.17	0.37	0.10	0.30	0.14	0.35
Hispanic	0.14	0.34	0.12	0.33	0.24	0.43
Other Race/Ethnicity	0.05	0.22	0.07	0.26	0.06	0.24
Mom High-School Graduate	0.31	0.46	0.29	0.45	0.28	0.45
Mom Some College	0.18	0.38	0.20	0.40	0.19	0.39
Mom College Graduate	0.38	0.49	0.39	0.49	0.35	0.48
Attends Preschool	0.61	0.49	0.58	0.49	0.61	0.49
Attends Kindergarten	0.07	0.26	0.05	0.22	0.08	0.27
State Characteristics						
Federal Head Start Enrollment (%)	11.91	3.24	12.72	7.15	11.81	2.68
SCHIP/Medicaid Income-to-Pov Ratio (Age 1-5)	2.19	0.59	2.03	0.43	2.03	0.40
Annual Unemployment Rate	4.90	0.85	4.64	1.14	5.17	1.10
Median Income ($1,000s)	53.11	7.14	56.12	7.86	58.04	5.73
Governor is Democrat	0.36	0.48	0.47	0.50	0.41	0.49
Democratic Party Controls Legislative	0.32	0.46	0.26	0.44	0.40	0.49
Observations	4765		4115		6661	

Notes: Summary statistics for the CPS samples of children living in treatment, control, and excluded states, observed at 4 years of age, 1997 to 2005 (weighted using sample weights). All state characteristics are merged to the CPS from other data sources (see Data Section for more details).

Table 3: Descriptive Statistics for Children Ages 5 to 12 in Treatment, Control and Excluded States

	NHIS Sample						CPS Sample					
	Treatment		Control		Excluded		Treatment		Control		Excluded	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Individual Characteristics												
Female	0.49	0.50	0.49	0.50	0.49	0.50	0.49	0.50	0.49	0.50	0.49	0.50
Black	0.18	0.38	0.10	0.29	0.15	0.35	0.17	0.37	0.09	0.29	0.14	0.34
Hispanic	0.15	0.36	0.13	0.34	0.26	0.44	0.15	0.35	0.14	0.34	0.25	0.43
Other Race/Ethnicity	0.04	0.21	0.07	0.26	0.06	0.23	0.06	0.24	0.08	0.27	0.08	0.27
Age (years after pre-K age)	4.49	2.30	4.55	2.29	4.55	2.29	4.52	2.29	4.51	2.30	4.50	2.30
Mom High-School Graduate	0.27	0.44	0.24	0.43	0.23	0.42	0.30	0.46	0.26	0.44	0.27	0.44
Mom Some College	0.19	0.39	0.21	0.41	0.20	0.40	0.18	0.38	0.21	0.41	0.19	0.39
Mom College Graduate	0.40	0.49	0.42	0.49	0.39	0.49	0.41	0.49	0.40	0.49	0.38	0.49
State Characteristics at Pre-K Age												
Federal Head Start Enrollment (%)	12.01	3.31	12.39	6.52	11.77	2.79	11.96	3.27	12.80	7.38	11.75	2.70
SCHIP/Medicaid Income-to-Pov Ratio (Age 1-5)	2.18	0.60	2.05	0.44	2.03	0.39	2.18	0.59	2.03	0.44	2.04	0.40
Annual Unemployment Rate	4.88	0.85	4.59	1.12	5.12	1.11	4.89	0.85	4.63	1.16	5.14	1.10
Median Income ($1,000s)	52.97	7.25	56.50	7.87	58.01	5.94	52.92	7.08	56.14	7.92	58.15	5.79
State Characteristics at Current Age												
SCHIP/Medicaid Income-to-Pov (Age 6-15)	2.44	0.66	2.24	0.41	2.21	0.35	2.43	0.66	2.22	0.41	2.22	0.35
Annual Unemployment Rate	5.88	1.84	5.69	2.07	6.46	2.19	5.90	1.88	5.74	2.12	6.44	2.15
Median Income ($1,000s)	52.63	7.61	55.81	8.00	57.31	6.66	52.59	7.37	55.53	8.06	57.52	6.53
Observations	11247		5664		19474		36149		34495		56426	

Notes: Summary statistics for the NHIS and CPS samples of children living in treatment, control, and excluded states, for the pre-K cohorts of 1997 to 2005, observed between 1998 and 2014 at 1 to 8 years after pre-K age in the NHIS and 2 to 8 years after pre-K age in the CPS. Weighted using sample weights. State characteristics at pre-K age are imputed according to the state where the child currently lives and the estimated year when the child would have been 4 years old (based on month/year of birth and month/year of interview). State characteristics at current age correspond to the state of residence and the interview year. All state characteristics are merged to the NHIS samples from other data sources (see Data section for more details).

Table 4: Effects of Pre-K Expansions on 4-Year-Olds' Preschool Attendance and Maternal Employment

	Preschool	Preschool	Mom Employed	Mom Employed	Employed Full-time	Employed Full-time
Full Sample						
Post Pre-K	0.073***	0.080***	-0.021	-0.033*	-0.003	-0.031
	(0.020)	(0.026)	(0.017)	(0.019)	(0.026)	(0.021)
Post Pre-K*Universal		-0.024		0.045		0.100***
		(0.032)		(0.035)		(0.031)
Outcome mean	0.602	0.602	0.626	0.626	0.453	0.453
N	8880	8880	8868	8868	8880	8880
P-value(Universal)		0.010		0.711		0.047
Mom with High-School or Less						
Post Pre-K	0.066**	0.060*	-0.042	-0.080**	-0.018	-0.076***
	(0.028)	(0.033)	(0.031)	(0.038)	(0.029)	(0.026)
Post Pre-K*Universal		0.020		0.125**		0.191***
		(0.038)		(0.047)		(0.053)
Outcome mean	0.510	0.510	0.553	0.553	0.434	0.434
N	3704	3704	3702	3702	3704	3704
P-value(Universal)		0.017		0.188		0.014
Mom with More than High-School						
Post Pre-K	0.075***	0.088***	-0.003	0.005	0.007	-0.000
	(0.025)	(0.032)	(0.035)	(0.045)	(0.040)	(0.044)
Post Pre-K*Universal		-0.050		-0.030		0.026
		(0.037)		(0.060)		(0.048)
Outcome mean	0.668	0.668	0.678	0.678	0.466	0.466
N	5176	5176	5166	5166	5176	5176
P-value(Universal)		0.154		0.561		0.569

Notes: The first two lines of each panel show estimates for separate regressions, for the outcome variable indicated in each column heading, for the sample of 4-year-olds in the CPS 1997-2005. All regressions include state and cohort fixed effects, individual-level control variables for maternal education, race/ethnicity, gender, age dummies, and state-level control variables. Robust standard errors (clustered by state) in parentheses. P-value(Universal) is the p-value of testing if the sum of the coefficients for Post Pre-K and Post Pre-K*Universal equals 0. * $p<0.1$, ** $p<0.05$, *** $p<0.01$ based on naïve p-values and Student's t-distribution with 32 degrees of freedom.

Table 5: Reduced-Form Effects on Child Outcomes 1-8 Years after Pre-K Age

	Sample: 1-4 years after pre-K				Sample: 5-8 years after pre-K			
	Repeating #	DBPI	Health Fair/Poor	Missed Days	Repeating	DBPI	Health Fair/Poor	Missed Days
Post Pre-K	-0.025**	-0.090**	0.011*	0.626***	-0.006*	-0.053	0.007	0.130
	(0.010)	(0.044)	(0.006)	(0.198)	(0.003)	(0.045)	(0.005)	(0.260)
Naïve p-value	0.016	0.047	0.076	0.003	0.070	0.249	0.148	0.620
Adj. p-value	0.048	0.094	0.152	0.014	0.279	0.996	0.592	1.000
Outcome mean	0.084	-0.023	0.018	3.298	0.016	0.020	0.021	3.166
N	25813	7841	7841	7671	33216	9070	9070	8980

Notes: Each cell shows results for separate regressions, for the age-group (number of years after pre-K age) indicated at the top heading, and the outcome variable indicated in each column heading (development problems index or health problems index). All regressions include state and cohort fixed effects, individual-level control variables for maternal education, race/ethnicity, gender, age dummies, and state-level control variables. Robust standard errors (clustered by state) in parentheses. Naïve p-values are not adjusted for multiple hypothesis testing. Adjusted p-values are calculated using Holm's method, considering each pair of hypotheses tested for each age group. #Children observed 1 year after preschool age are excluded from the sample. Outcome means indicated for the full sample of all states. * p<0.1, ** p<0.05, *** p<0.01 based on naïve p-values and Student's t-distribution with 32 degrees of freedom.

Table 6: Effects on Additional Outcomes 1-8 Years after Pre-K Age

	Sample: 1-4 years after pre-K			Sample: 5-8 years after pre-K		
Outcome variable	Post Pre-K Coefficient	N	Outcome Mean	Post Pre-K Coefficient	N	Outcome Mean
1. Individual Components of DBPI						
Limitations, Speech	0.003	7841	0.026	-0.004	9070	0.018
	(0.009)			(0.004)		
Limitations, Intellectual Disability	-0.007***	7841	0.002	-0.004	9070	0.003
	(0.002)			(0.003)		
Limitations, Learning Disability	-0.008	7841	0.019	-0.004	9070	0.030
	(0.009)			(0.008)		
Limit., Mental, Emotional, Behavioral	-0.008**	7841	0.010	-0.007	9070	0.019
	(0.003)			(0.006)		
Limit., Other Developmental Problem	0.001	7841	0.009	0.005	9070	0.009
	(0.005)			(0.004)		
2. Additional health outcomes						
Frequent ear infections	0.003	7841	0.061	-0.004	9070	0.028
	(0.016)			(0.007)		
Any asthma episode	0.018	7841	0.063	0.016	9070	0.059
	(0.013)			(0.011)		
Frequent headaches	-0.004	7841	0.038	0.006	9070	0.070
	(0.012)			(0.015)		
Colitis/Frequent diarrhea	0.015***	7839	0.011	0.001	9070	0.011
	(0.005)			(0.005)		
Any hospitalization	0.007	9065	0.019	0.005	9003	0.015
	(0.005)			(0.006)		
Asthma-related ER visit	0.005	8376	0.022	-0.004	8497	0.017
	(0.008)			(0.007)		

Notes: Each cell in the coefficients column shows results for separate regressions, for the age-group (number of years after pre-K age) indicated at the top heading, and the outcome variable indicated in the left column. All regressions include state and cohort fixed effects, individual-level control variables for maternal education, race/ethnicity, gender, age dummies, and state-level control variables. Robust standard errors (clustered by state) in parentheses. Outcome means indicated for the full sample of all states. * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Heterogeneity of Effects by Program and Individual Characteristics

	Sample: 1-4 years after pre-K				Sample: 5-8 years after pre-K			
	Repeating #	DBPI	Health Fair/Poor	Missed Days	Repeating	DBPI	Health Fair/Poor	Missed Days
1. Universal and targeted programs								
Post Pre-K	-0.026**	-0.102*	0.013*	0.547**	-0.006	-0.078	0.009	0.138
	(0.010)	(0.055)	(0.007)	(0.236)	(0.004)	(0.052)	(0.006)	(0.298)
Post Pre-K*Universal	0.003	0.050	-0.007	0.353	0.002	0.105	-0.006	-0.036
	(0.016)	(0.079)	(0.009)	(0.329)	(0.005)	(0.086)	(0.008)	(0.313)
2. NIEER Quality Score								
Post Pre-K	-0.003	-0.029	0.003	0.810	0.002	-0.228**	-0.001	-1.008**
	(0.021)	(0.132)	(0.008)	(0.479)	(0.006)	(0.100)	(0.011)	(0.452)
Post Pre-K*NIEERS	-0.003	-0.009	0.001	-0.027	-0.001	0.026**	0.001	0.168***
	(0.002)	(0.016)	(0.001)	(0.069)	(0.001)	(0.012)	(0.002)	(0.056)
3. Per Student Spending								
Post Pre-K	-0.024*	-0.208**	0.023**	0.843*	-0.007*	-0.052	0.001	-0.301
	(0.014)	(0.093)	(0.010)	(0.430)	(0.004)	(0.069)	(0.007)	(0.360)
Post Pre-K*Spending	-0.000	0.025*	-0.002	-0.046	0.000	-0.000	0.001	0.090*
	(0.002)	(0.014)	(0.001)	(0.085)	(0.001)	(0.012)	(0.001)	(0.044)
4. Maternal Education								
Post Pre-K	-0.024**	-0.107**	0.005	0.604**	-0.006*	-0.065	0.010*	0.079
	(0.009)	(0.052)	(0.006)	(0.227)	(0.003)	(0.063)	(0.006)	(0.299)
Post Pre-K*Low Edu.	-0.001	0.043	0.016*	0.056	-0.000	0.031	-0.008	0.132
	(0.009)	(0.040)	(0.008)	(0.133)	(0.003)	(0.069)	(0.007)	(0.327)
5. Race/Ethnicity								
Post Pre-K	-0.023**	-0.096*	0.008	0.596***	-0.005*	-0.074	0.011*	0.071
	(0.009)	(0.053)	(0.008)	(0.216)	(0.003)	(0.062)	(0.006)	(0.267)
Post Pre-K*Black	-0.003	0.016	0.005	-0.087	-0.015***	0.034	-0.011	-0.096
	(0.016)	(0.078)	(0.009)	(0.281)	(0.005)	(0.087)	(0.011)	(0.293)
Post Pre-K*Hispanic	-0.008	0.025	0.017	0.403	0.013***	0.110	-0.012	0.545
	(0.010)	(0.057)	(0.021)	(0.335)	(0.004)	(0.087)	(0.009)	(0.520)
6. Boys Only Sample								
Post Pre-K	-0.031***	-0.081	0.009	0.595**	-0.003	-0.116	0.006	0.257
	(0.011)	(0.062)	(0.007)	(0.279)	(0.003)	(0.075)	(0.007)	(0.439)
N	13243	4045	4045	3947	16993	4615	4615	4574
7. Girls Only Sample								
Post Pre-K	-0.017	-0.118**	0.017	0.639*	-0.009	0.049	0.011	0.027
	(0.012)	(0.051)	(0.011)	(0.374)	(0.006)	(0.076)	(0.008)	(0.283)
N	12570	3796	3796	3724	16223	4455	4455	4406

Notes: Each cell shows results for separate regressions, for the outcome variable indicated in the column heading, and the sample (number of years after pre-K age) indicated at the top. All regressions include state and cohort fixed effects, individual-level control variables for maternal education, race/ethnicity, gender, age dummies, and state-level control variables. Sample sizes are the same as in Table 5 unless otherwise indicated. Robust standard errors (clustered by state) in parentheses. #Children observed 1 year after preschool age are excluded from the sample. * p<0.1, ** p<0.05, *** p<0.01.

Table 8: Specification Checks

	Sample: 1-4 years after pre-K				Sample: 5-8 years after pre-K			
	Repeating #	DBPI	Health Fair/Poor	Missed Days	Repeating	DBPI	Health Fair/Poor	Missed Days
1. Add state-specific linear trends								
Post Pre-K	-0.024*	-0.133**	0.017**	0.404*	0.000	-0.044	0.013*	0.072
	(0.014)	(0.060)	(0.008)	(0.204)	(0.004)	(0.060)	(0.007)	(0.337)
*2. Add region*cohort fixed effects*								
Post Pre-K	-0.026**	-0.125**	0.012**	0.501*	0.000	-0.076	0.011*	0.090
	(0.012)	(0.058)	(0.006)	(0.250)	(0.003)	(0.048)	(0.006)	(0.184)
*3. Add region*year fixed effects*								
Post Pre-K	-0.030**	-0.124**	0.014**	0.460**	0.000	-0.064	0.009*	0.093
	(0.013)	(0.060)	(0.005)	(0.222)	(0.003)	(0.051)	(0.005)	(0.246)
4. Timing of effects								
-1 yrs. after Pre-K	0.008	-0.016	0.010	0.300	-0.005	-0.069	-0.005	0.005
	(0.015)	(0.060)	(0.006)	(0.179)	(0.004)	(0.046)	(0.005)	(0.202)
0 yrs. after Pre-K	-0.025**	-0.116***	0.018**	0.871***	-0.007**	-0.120*	0.003	-0.047
	(0.010)	(0.035)	(0.008)	(0.261)	(0.003)	(0.063)	(0.005)	(0.327)
1+ yrs. after Pre-K	-0.017	-0.073	0.013**	0.594*	-0.009**	-0.050	0.007	0.306
	(0.014)	(0.069)	(0.006)	(0.338)	(0.004)	(0.041)	(0.006)	(0.341)
5. Only cohort & state FE								
Post Pre-K	-0.021*	-0.108**	0.012**	0.507**	-0.006*	-0.061	0.006	0.081
	(0.010)	(0.044)	(0.005)	(0.216)	(0.003)	(0.043)	(0.005)	(0.240)
6. Add control for 4-year-olds served by CCDF (1998-2005)								
Post Pre-K	-0.027**	-0.111**	0.012	0.608***	-0.005	-0.071	0.010*	0.037
	(0.011)	(0.047)	(0.008)	(0.219)	(0.004)	(0.057)	(0.006)	(0.292)
N	22890	7368	7368	7207	29076	8017	8017	7941
7. Switch off treatment for states with small Pre-K expansions								
Post Pre-K	-0.034***	-0.102**	0.012*	0.787***	-0.008***	0.023	0.002	0.338*
	(0.010)	(0.045)	(0.006)	(0.232)	(0.003)	(0.032)	(0.006)	(0.199)
8. Include all states in sample								
Post Pre-K	-0.022***	-0.089**	0.009	0.462***	-0.006**	-0.021	0.008*	0.154
	(0.008)	(0.042)	(0.005)	(0.157)	(0.003)	(0.036)	(0.005)	(0.241)
N	46015	16985	16985	16680	60768	19400	19400	19201

Notes: Each cell shows results for separate regressions, for the outcome variable indicated in the column heading, and the sample (number of years after pre-K age) indicated at the top. All regressions include state and cohort fixed effects, individual-level control variables for maternal education, race/ethnicity, gender, age dummies, and state-level control variables. Sample sizes are the same as in Table 5 unless otherwise indicated. Robust standard errors (clustered by state) in parentheses. #Children observed 1 year after preschool age are excluded from the sample. * p<0.1, ** p<0.05, *** p<0.01.

Table 9: Approximations to Local Average Treatment Effects

	$\hat{\beta}_{ITT}$	$\hat{\beta}_{LATE}$	$\hat{\beta}^{Upper}_{LATE_{nk}}$
First stage coef.		0.184	0.073***
First stage s.e.		–	(0.020)
Repeating			
2-4 Years After Pre-K	-0.025 **	-0.136 **	-0.342 **
	(0.010)	(0.054)	(0.166)
5-8 Years After Pre-K	-0.006 *	-0.033 *	-0.082 *
	(0.003)	(0.016)	(0.047)
Development and Behavior Problems Index			
1-4 Years After Pre-K	-0.090 **	-0.490 **	-1.232 *
	(0.044)	(0.239)	(0.690)
5-8 Years After Pre-K	-0.053	-0.288	-0.725
	(0.045)	(0.245)	(0.647)
Health Fair/Poor			
1-4 Years After Pre-K	0.011 *	0.060 *	0.151
	(0.006)	(0.033)	(0.092)
5-8 Years After Pre-K	0.007	0.038	0.096
	(0.005)	(0.027)	(0.073)
Missed Days of School			
1-4 Years After Pre-K	0.626 ***	3.405 ***	8.566 **
	(0.198)	(1.077)	(3.584)
5-8 Years After Pre-K	0.130	0.707	1.779
	(0.260)	(1.414)	(3.591)

Notes: Each cell shows results for separate regressions, for the outcome variable indicated in the left column, and the sample (number of years after pre-K age) indicated at the top heading. All regressions include state and cohort fixed effects. Robust standard errors (clustered by state) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

A Data Appendix For Online Publication

A.1 Development and Behavior Problems Index

The aim of this paper is to estimate the impacts of pre-K programs on children's development and health. I construct four main outcome variables: one that reflects academic outcomes (grade repetition), one that summarizes information on development and behavior problems (Development and Behavior Problems Index), and two measures of physical health. In this section, I provide more details on the construction of the Development and Behavior Problems Index.

As described in section 3.2 of the paper, I begin with five measures of child developmental and behavioral problems. Each of these measure is an indicator of daily life limitations identified by the respondent to be related to one of the following categories of child conditions: speech problem, intellectual disability, other developmental problem, other mental, emotional or behavioral problem, and learning disability. These measures are based on the NHIS Health Status and Limitation of Activity Section of the Family Core. Information on activity and daily life limitations is collected for each family member, and if any limitations are identified the respondent is asked to specify the condition(s) causing the limitation. The responses are classified in thirteen categories of child conditions based on the International Classification of Diseases, Ninth Revision, Clinical Modification.[49]

The goal of the principal component analysis (PCA) is to reduce the dimensions of the information by projecting the data into a lower number of components or factors, while retaining as much as possible of the variance of the data. The principal components are selected so that the first one minimizes the distance between the data and their projection, i.e. maximizes the variance of the projected data, and each subsequent component is chosen similarly with the condition that it is uncorrelated with all the previous ones.

Table A1 shows the results of the principal component analysis of the five measures. I start with no assumptions about the number of principal components that should be retained, so the table shows the estimated factor loadings for all the five possible factors, as well as their Eigenvalues and cummulative variance explained. Using Keiser's eigenvalue rule for selecting the number of factors,

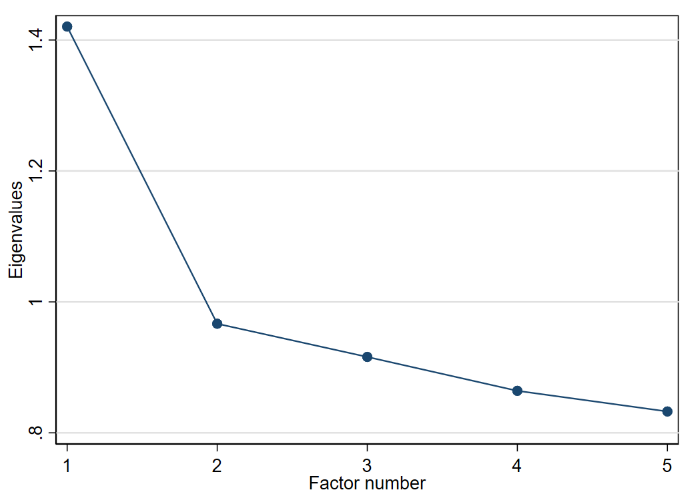
[49]The full list of categories for children includes: "vision/problem seeing," "hearing problem," "speech problem," "asthma/breathing problem," "birth defect," "injury," "intellectual disability (mental retardation)," "other developmental problem (e.g., cerebral palsy)," "other mental, emotional, or behavioral problem," "bone, joint, or muscle problem," "epilepsy or seizures," "learning disability," "attention deficit/hyperactivity disorder," and two instances of "other impairment problem." Respondents could supply a 50-character verbatim response for one or both of the "other impairment problem" categories, which were then back-coded to one of the 13 fixed categories when possible.

I choose to retain only the first factor, which is the only one with an eigenvalue greater than 1. This decision is confirmed by an inspection of the scree plot of the eigenvalues, shown in Figure A1, where it is clear that after the first factor the eigenvalues decrease smoothly and are almost aligned in a straight line. In addition, it is only in the first factor that all of the measures enter with a positive sign, which allows us to interpret this factor as a combination of developmental and behavioral problems faced by children.

Table A1: Principal Component Analysis of Development Outcomes

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor Loadings					
Limitations, Speech Problems	0.519	0.025	-0.759	0.298	-0.255
Limitations, Intellectual Disability	0.590	0.156	0.067	-0.735	-0.289
Limitations, Learning Disability	0.591	-0.404	-0.040	-0.075	0.693
Limitations, Mental, Emotional, Behavioral	0.533	-0.370	0.528	0.391	-0.384
Limitations, Other Developmental Problem	0.413	0.801	0.234	0.277	0.237
Eigenvalues	1.421	0.967	0.916	0.864	0.833
Cummulative variance	28%	48%	66%	83%	100%

Figure A1: Scree Plot of Eigenvalues of Development Principal Components



A.2 Additional health outcomes

To explore health effects in more detail, in addition to the two measures of overall physical health, I also explore the effects on four indicators of specific health conditions. These are conditions that are frequently observed in children and that allow me to explore different pathways through which preschool education may affect child physical health. "Frequent ear infections" is a binary variable that indicates whether the child experienced 3 or more year infections in the past 12 months. This variable has a mean of 4.3% of the full sample, and it is decreasing with age. "Frequent headaches" indicates whether the child had frequent headaches/migraines in the past 12 months. This variable has a mean of 5.5% of the full sample, and it is increasing with age. "Asthma episode" indicates whether the child had an asthma episode in the last 12 months (6.1% of the full sample). "Frequent diarrhea/colitis" indicates whether the child had frequent diarrhea/colitis (also referred to as irritable bowel syndrome) in the past 12 months (1.1% of the full sample). Asthma episodes and frequent diarrhea are both relatively stable with age. These are based on the NHIS Child Conditions, Limitation of Activity and Health Status Section, which questions respondents on a list of specific conditions.[50] Finally, I look at two variables related to health care utilization: an indicator for whether the child had a hospital stay in the past 12 months, and an indicator for any ER visit related to an asthma episode in the past 12 months. Both of these variables indicate utilization of health care for potentially serious health events.

A.3 State-level data sources

I obtained the annual average state unemployment rate from the Bureau of Labor Statistics (BLS), and the state median household income from the U.S. Census Bureau. State family-income-to-poverty ratio requirements for eligibility for Medicaid or SCHIP (whichever is lowest) for children ages 1-5 and 6-15 were obtained from NGA Center for Best Practices (1997-2011) and Kaiser Commission on Medicaid and the Uninsured (2006-2014). I obtained measures of political context (political party in control of the state government and legislature) from University of Kentucky Center for Poverty Research (2017).

To control for other policies affecting early childhood education, I compute the federally-funded enrollment of 4-year-olds in Head Start as a percentage of the state's population of 4-year-olds

[50] The full list of child conditions assessed includes: mental retardation, developmental delays, Attention Deficit Hyperactivity Disorder (ADHD) or Attention Deficit Disorder (ADD), Down's syndrome, cerebral palsy, muscular dystrophy, cystic fibrosis, sickle cell anemia, autism, diabetes, arthritis, congenital and other heart disease, asthma, allergies, colitis, anemia, ear infections, seizures, headaches, stuttering, and stammering.

based on information from the U.S. Department of Health and Human Services and the U.S. Census Bureau. First, I compute the number of federally-funded enrollment of 4-year-olds by multiplying the total federally-funded enrollment (number of children) by the percentage of actual enrollment that corresponds to 4-year-olds, both obtained from Head Start Program Information Reports 1997-2005, from the Office of Head Start, U.S. Department of Health and Human Services. I then divide this number by the population of 4-year-olds by state in 2000, obtained from the U.S. Census Bureau.

I also compute the percent of 4-year-olds in each state served by the Child Care Development Fund (CCDF), which provides child care subsidies to low-income families and is implemented by states through a federal block-grant, using data from the U.S. Department of Health and Human Services and the U.S. Census Bureau. I use data from the Office of Child Care, U.S. Department of Health and Human Services on the number of children served (monthly average by state) from 1998 to 2005, and the percentage of served children by age from 2002 to 2005. For the years 2002-2005 I compute the percentage of 4-year-olds in each state served by the CCDF, by multiplying the total number of children served by the percentage of children served who were 4-year-olds, and then dividing this number by the population of 4-year-olds in each state in the year 2000 (from the U.S. Census Bureau data). For the years 1998-2001 I follow a similar procedure but, since the percentage of children served by age is not available, I use the average percentage by age in the years 2002-2004.

B Supplemental Tables For Online Publication

Table B1: Pre-K Programs Description

State	Year	Program Name (New?)	Targetting	Private providers	NIEERS, Enrollment, Spending	Hrs/day Days/week	Law/Ruling, Motivation and Funding Sources	Party Legis. Control	Party Change	K-12 Provisions
Arkansas	2004	Better Chance for School Success	<200% FPL, Low scoring schools	Yes	9/10, 18%, $7,769	7.5h/d, 5 d/w	Act 1332 (2003) was passed to expand availability of early childhood programs. Participation of eligible school districts is required. The Arkansas Better Chance for School Success program is created to provide funidng for early education programs for 3- and 4-year-olds.	D	No	No
Florida	2005	Voluntary Prekindergarten	Universal	Yes	4/10, 47%, $2,163	3h/d, 5d/w	State Constitution Art. IX (2004). Voters approved ballot initiative proposing an amendment to the Florida Constitution in the November 2002 general election. Required the Legislature to establish a universal free prek program by the 2005 academic year. The legislature negotiated the bill (HB 1-A) that was signed into law in January 2005.	R	No	No
Kansas	2002	Four-Year-Old At-Risk	<100% FPL or other risk factor		3/10, 15%, $2,554	2.5h/d, 4-5d/w	House Bill 2249 (1998) incorporated a small pilot of Four-Year-Old At-Risk pupils in the school finance formula. The pilot was implemented in 2001 and scaled up in 2002.	R	No	No
Louisiana	2002	LA4 and NSECD	<185% FPL (LA4); <200%FPL (NSECD)	Yes	8/10 (LA4), 9/10 (NSECD), 22%, $5,012	6-10h/d, 5d/w	Senate Bill 776 (2001). The goal is to provide high quality early childhood education to 4-year-olds who are at risk of learning difficulties and to improve the children's readiness to begin school. Directed toward the development of cognitive, social, emotional, language and literacy, and motor skills in a manner and at a pace consistent with the needs and capabilities of the individual child.	D	No	No

(Continued...)

Table B1: Pre-K Programs Description (Continued)

State	Year	New Program	Targetting	Private providers	NIEERS, Enrollment, Spending	Hrs/day Days/week	Law/Ruling, Motivation and Funding Sources	Party Legis. Control	Party Change	K-12 Provisions
Missouri	1999	Missouri Preschool Project	<185% FPL or special needs	Yes	6/10, 4%, $2,632	3-6.5h/d, 5d/w	The General Assembly passed House Bill 1519 (1998) that created the Early Childhood Care and Education Fund. Its goal is to promote the growth and quality of early childhood care and education and school readiness for children. The funding comes from Missouri gaming funds.	D	No	No
Nebraska	2001	Early Childhood Education	<185% FPL or other risk factor	Yes	8/10, 4%, $7,418	3-6h/d, 4-5d/w	The legislature passed LB 759 (2001) with the goal of encouraging schools and community-based organizations to provide high-quality early childhood education programs which include family involvement. It also encourages and assists coordination and efforts to improve training opportunities for staff.	–	–	No
New Jersey	1999	Abbott Preschool Program	Poor urban school districts	Yes	9/10, 18%, $9,854	6h/d, 5d/w	State supreme court ruling (1998) determined that the State had not given sufficient evidence that the educational needs of children in poor urban (Abbott) districts were being made. They directed the State to implement a comprehensive set of remedies to improve education in the Abbott districts, including universal preschool beginning in the 1999-2000 school year.	R	No	Yes
New Mexico	2005	New Mexico Pre-K	Low income (Title I) school districts	Yes	5/10, 7%, $2,269	2.5-3h/d, 5d/w	Pre-Kindergarten Act (2005) was passed to advance childhood development. It creates two funds which consist of appropriations, income from investment of the funds, gifts, grants and donations, to be used for reimbursing preschool programs in public schools and eligible private providers).	D	No	No
New York	1998	Universal Prekindergarten	Universal		5/10, 29%, $3,512	2.5-5h/d, 5d/w	Amendment to the NY State Education Law (1997). The budget agreement created new Universal Prekindergarten program.	Split (D)	N.A.	Yes

(Continued...)

Table B1: Pre-K Programs Description (Continued)

State	Year	New Program	Targetting	Private providers	NIEERS, Enrollment, Spending	Hrs/day Days/week	Law/Ruling, Motivation and Funding Sources	Party Legis. Control	Party Change	K-12 Provisions
North Carolina	2002	More at Four	<75% SMI or other risk factor	Yes	10/10, 12%, $3,892	6h/d, 5d/w	House Bill 1531 (2001) earmarked an increase in cigarette tax for public school class size reduction (25%), the More at Four program (25%), and teacher salaries (25%).	D	No	Yes
Oklahoma	1998	Early Childhood Four-Year-Old Program	Universal	Yes	9/10, 70%, $6,167	2.5-6h/d, 5d/w	House Bill 1657 (1998) aimed to maximize the educational return of the budget allocated to education. It removed the income requirement that restricted access to early childhood programs, increased the funding per student allocated to full day and half-day early childhood programs, and eliminated State aid for underage students enrolled in kindergarten and the 1st grade.	D	No	Yes
Pennsylvania	2004	EAGB and Head Start Supplemental Assistance Program	Low income or at risk with thresholds determined locally; <100% FPL	Yes	2-6/10, 6%, $3,090-$6,369	2.5+h/d, 5d/w	State supplements to Head Start through the Head Start Supplemental Assistance Program (HSSAP) were initiated in 2004-2005, and the supplemental funds were primarily used to create new Head Start slots. In the same year, Pennsylvania started the Education Accountability Block Grant (EABG), which provided additional funding to school districts and was primarily used to implement pre-K programs and expand full day kindergarten. These funds were directed to school districts that were identified as low performing, in accordance to the 2000 Education Empowerment Act.	R	No	Yes

(Continued...)

Table B1: Pre-K Programs Description (Continued)

State	Year	New Program	Targetting	Private providers	NIEERS, Enrollment, Spending	Hrs/day Days/week	Law/Ruling, Motivation and Funding Sources	Party Legis. Control	Party Change	K-12 Provisions
Tennessee	2005	Voluntary Pre-K Initiative	<185% FPL or other risk factor	Yes	9/10, 11%, $4,061	5.5h/d, 5d/w	House Bill 2333 (2005) allocated excess lottery funds to implement the Voluntary Pre-K Initiative. The legislative argued that it was based on the success of the existing pilot pre-K programs. In addition to lottery funding, in the 2006-2007 school year the state increased general revenue allocation for prekindergarten to open additional programs.	Split (D)	Yes (D)	No
Vermont	2003	Publicly Funded Prekindergarten	Universal	Yes	7/10, 47%, $2,930	10h/w	While Act 62, which established publicly funded prekindergarten education, became effective in 2007, reforms in school funding and preschool enrollment started before that. Act 60 (1997) reforming the state's distribution of educational taxes and spending, and Act 68 (2003) made some additional changes in educational taxes and funding allocation. A large increase in fuding allocated towards preschool education and enrollment was observed starting in the 2003-2004 school year.	Split (D)	Yes (Split R)	No
West Virginia	2002	Universal Pre-K Program	Universal	Yes	7/10, 40%, $7,758	12+h/w	Code 18-5-44 Amendment (2002) required the state to expand access to preschool education programs in order to make pre-k available to all 4-year-olds by 2012-13. Funding is part of the State School Aid Funding Formula.	D	No	No

Notes: Sources described in Appendix A. Year is the effective school year of implementation. New Program indicates if a new program was created, otherwise it was an expansion of a pre-existing program. Targetting indicates eligibility criteria for students, school districts, schools or programs. Private providers and Head Start centers indicate whether program was implemented in those providers as well as public schools. NIEERS is the number of NIEER quality standards met by the program in 2005 (out of 10). Enrollment is percent of 4-year-olds enrolled in the program in 2005. Spending is the total spending per enrollee in 2005 (including state funding, local matching funds, and federal grants administered by the state. Hrs/day and Days/week indicate the minimum number of hours in classroom that all providers must offer per week. Motives & Funding describes the goals established in the legislation/ruling and the sources of funding (when available). Party Legis. Control indicates the political party with legislative control in the year when the program was introduced. Party Change indicates whether the party with legislative control had recently changed before the introduction. K-12 Provisions indicates whether there were other provisions related to K-12 education introduced in the same year.

Table B2: State Pre-K Enrollment in Treatment States 2001-2005

State	State Pre-K Enrollment					Pre-K	After	Before	Enrollment	4-Year-Old
	2001	2002	2003	2004	2005	Year	Pre-K	Pre-K	Increase	Population
Arkansas	6	6	6	12	18	2004	15	6	9	36,909
Florida	0	0	0	0	47	2005	47	0	47	194,475
Kansas	6	15	15	15	15	2002	15	6	9	38,444
Louisiana	12	21	22	20	22	2002	21	12	9	64,196
Missouri	5	4	4	4	4	1999	4	0[a]	4	75,416
Nebraska	2	3	4	3	4	2001	3	1[b]	2	23,881
New Jersey	20	24	26	26	25	1999	24	9[c]	15	114,766
New Mexico	1	1	1	1	7	2005	7	1	6	26,461
New York	25	30	30	29	29	1998	29	4[d]	25	256,184
North Carolina	1	6	9	10	12	2002	9	1	8	107,107
Oklahoma	56	59	64	68	70	1998	63	35[e]	28	47,075
Pennsylvania	2	2	2	5	6	2004	6	2	4	152,001
Tennessee	2	3	3	3	11	2005	11	3	8	74,575
Vermont	9	10	36	45	47	2003	43	10	33	7,421
West Virginia	24	29	33	35	40	2002	34	24	10	21,141
Total										1,240,052

Notes: Enrollment rates correspond to years 2001-2005 for each treatment state, from NIEER (2006). The population of 4-year-olds by state corresponds to the information from the 2000 Census (US Census Bureau). *Post Pre-K Average* enrollment is computed as the simple average of the enrollment rates for the years after a state pre-K policy was implemented or expanded, if implemented in or after 2001, or the simple average of years 2001-2005, if implemented before 2001. *Pre Pre-K Average* enrollment is computed as the simple average of the enrollment rates for the years before a state pre-K policy was implemented or expanded with available information, if implemented after 2001. For states with pre-K programs implemented in 2001 or before, the pre-Pre-K enrollment rates were computed using information from other sources, as indicated in the note corresponding to each figure. [a]No pre-K program existed before 1999. [b]A pilot program was being implemented before 2001. Because no information is available, I make the (conservative) assumption that the expansion duplicated the size of the pilot program. This yields an pre-expansion average enrollment similar to the average enrollment rates of control states in 2001. [c]Pre-K expansion corresponds to program implemented for Abbot Districts, so pre-expansion enrollment was computed by adding Non-Abbot districts' enrollment in 2002 (NIEER, 2003) and the estimated Abbot districts enrollment of 4-year-olds in 1998, from Frede et al. (2009). [d]EPK enrollment in 2001. [e]Mainly operates in public schools, so assumption that increase is captured by enrollment in public preschools in CPS is reasonable. From CPS, public preschool enrollment in Oklahoma: 0.23 in 1993-1997 and 0.51 in 1998-2002, so increase of 28 p.p.

Table B3: Effects on Pre-K Expansions on Attendance to Preschool Reported as Public/Private

	Public Preschool	Private Preschool	Public Preschool	Private Preschool
Post Pre-K expansion	0.050***	0.023	0.029	0.051*
	(0.018)	(0.024)	(0.018)	(0.028)
Post Universal Pre-K			0.075**	-0.098***
			(0.032)	(0.034)
Observations	8880	8880	8880	8880

Notes: Each cell shows results for separate regressions, for the outcome variable indicated in each column heading (public or private preschool attendance), for the sample of 4-year-olds in the CPS 1997-2005. All regressions include state and cohort fixed effects, individual-level control variables for maternal education, race/ethnicity, gender, age dummies, and state-level control variables. Robust standard errors (clustered by state) in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$ based on naïve p-values and Student's t-distribution with 32 degrees of freedom.

Table B4: Specification Checks, Boys Sample

	Sample: 1-4 years after pre-K				Sample: 5-8 years after pre-K			
	Repeating #	DBPI	Health Fair/Poor	Missed Days	Repeating	DBPI	Health Fair/Poor	Missed Days
Add state-specific linear trends								
Post Pre-K	-0.032*	-0.077	0.015	0.386	-0.001	-0.114	0.015	0.245
	(0.016)	(0.076)	(0.010)	(0.320)	(0.004)	(0.082)	(0.010)	(0.445)
N	13243	4045	4045	3947	16993	4615	4615	4574
*Add region*cohort fixed effects*								
Post Pre-K	-0.033**	-0.119*	0.010	0.557*	0.004	-0.137	0.014*	0.045
	(0.014)	(0.065)	(0.007)	(0.313)	(0.004)	(0.096)	(0.008)	(0.345)
N	13243	4045	4045	3947	16993	4615	4615	4574
*Add region*year fixed effects*								
Post Pre-K	-0.037**	-0.119	0.011	0.464	0.001	-0.116	0.009	0.174
	(0.014)	(0.075)	(0.008)	(0.298)	(0.003)	(0.083)	(0.007)	(0.406)
N	13243	4045	4045	3947	16993	4615	4615	4574
Add control for 4-year-olds served by CCDF (1998-2005)								
Post Pre-K	-0.034**	-0.097	0.011	0.583*	0.000	-0.148	0.009	-0.003
	(0.014)	(0.063)	(0.009)	(0.309)	(0.003)	(0.099)	(0.008)	(0.525)
N	11727	3787	3787	3696	14881	4091	4091	4055
Only cohort & state FE								
Post Pre-K	-0.028**	-0.092	0.010	0.479	-0.003	-0.114	0.004	0.287
	(0.012)	(0.067)	(0.007)	(0.315)	(0.003)	(0.078)	(0.007)	(0.405)
N	13243	4045	4045	3947	16993	4615	4615	4574
Switch off treatment for states with small Pre-K expansions								
Post Pre-K	-0.043***	-0.106	0.004	0.697*	-0.003	-0.017	-0.000	0.499
	(0.012)	(0.063)	(0.008)	(0.353)	(0.004)	(0.048)	(0.010)	(0.470)
N	13243	4045	4045	3947	16993	4615	4615	4574
Include all states in sample								
Post Pre-K	-0.025***	-0.070	0.006	0.391	-0.004	-0.069	0.010	0.292
	(0.009)	(0.065)	(0.006)	(0.299)	(0.003)	(0.055)	(0.007)	(0.405)
N	23651	8767	8767	8602	31025	9929	9929	9831

Notes: Each cell shows results for separate regressions, for the outcome variable indicated in the column heading, and the sample (number of years after pre-K age) indicated at the top. All regressions include state and cohort fixed effects, individual-level control variables for maternal education, race/ethnicity, age dummies, and state-level control variables. Robust standard errors (clustered by state) in parentheses. #Children observed 1 year after preschool age are excluded from the sample. * p<0.1, ** p<0.05, *** p<0.01.

Table B5: Specification Checks, Girls Sample

	Sample: 1-4 years after pre-K				Sample: 5-8 years after pre-K			
	Repeating #	DBPI	Health Fair/Poor	Missed Days	Repeating	DBPI	Health Fair/Poor	Missed Days
Add state-specific linear trends								
Post Pre-K	-0.016	-0.219**	0.021	0.404	0.000	0.071	0.015	-0.126
	(0.017)	(0.085)	(0.015)	(0.481)	(0.006)	(0.102)	(0.010)	(0.436)
N	12570	3796	3796	3724	16223	4455	4455	4406
*Add region*cohort fixed effects*								
Post Pre-K	-0.019	-0.157*	0.016	0.452	-0.003	0.030	0.010	0.212
	(0.014)	(0.085)	(0.012)	(0.485)	(0.005)	(0.089)	(0.009)	(0.292)
N	12570	3796	3796	3724	16223	4455	4455	4406
*Add region*year fixed effects*								
Post Pre-K	-0.022	-0.142*	0.021*	0.426	-0.001	0.025	0.010	0.078
	(0.015)	(0.072)	(0.012)	(0.402)	(0.006)	(0.075)	(0.009)	(0.337)
N	12570	3796	3796	3724	16223	4455	4455	4406
Add control for 4-year-olds served by CCDF (1998-2005)								
Post Pre-K	-0.018	-0.144**	0.016	0.641	-0.011	0.050	0.016	0.132
	(0.014)	(0.061)	(0.013)	(0.423)	(0.008)	(0.081)	(0.009)	(0.302)
N	11163	3581	3581	3511	14195	3926	3926	3886
Only cohort & state FE								
Post Pre-K	-0.014	-0.124**	0.017	0.539	-0.009	0.031	0.010	-0.094
	(0.012)	(0.049)	(0.012)	(0.395)	(0.007)	(0.077)	(0.008)	(0.276)
N	12570	3796	3796	3724	16223	4455	4455	4406
Switch off treatment for states with small Pre-K expansions								
Post Pre-K	-0.023	-0.112*	0.022*	0.857**	-0.014**	0.086	0.007	0.166
	(0.013)	(0.059)	(0.012)	(0.418)	(0.005)	(0.077)	(0.008)	(0.317)
N	12570	3796	3796	3724	16223	4455	4455	4406
Include all states in sample								
Post Pre-K	-0.018*	-0.117***	0.109	0.565*	-0.009	0.059	0.088	0.022
	(0.009)	(0.043)	(0.074)	(0.323)	(0.006)	(0.061)	(0.060)	(0.266)
N	22364	8218	8077	8078	29743	9471	9369	9370

Notes: Each cell shows results for separate regressions, for the outcome variable indicated in the column heading, and the sample (number of years after pre-K age) indicated at the top. All regressions include state and cohort fixed effects, individual-level control variables for maternal education, race/ethnicity, age dummies, and state-level control variables. Robust standard errors (clustered by state) in parentheses. #Children observed 1 year after preschool age are excluded from the sample. * p<0.1, ** p<0.05, *** p<0.01.

C Appendix for Online Publication: Correlation of pre-K programs with other state characteristics

In this appendix I explore the correlation of the implementation of pre-K programs with other time-variant characteristics of the states. This is not a direct test of the exogeneity of the pre-K policies, as there can always remain potentially confounding unobserved factors, but it allows me to mitigate concerns about the potentially confounding effect of a large number of observable state characteristics. In particular, I explore the presence of confounding factors related to states' economic situation, changes in expenditures in K-12 education, implementation of other social programs, and the demographic composition of the states.

Table C1 presents the estimation results of regression models at the state-year level, which predict the timing of the implementation of a pre-K program as a function of observable state characteristics, controlling for state and year fixed effects. In the first two columns the outcome variable is a dummy variable that indicates the year when a pre-K program is implemented, while in the third column the outcome variable indicates all post-pre-K years. All regressions are weighted by the state population of 4-year-olds in 2000.[51] In the first column, I begin by exploring the role of the political party in control of the states' government and legislature, by regressing the indicator for the year of pre-K implementation on two variables that indicate whether the Democratic Party is in control of the state government and of the state legislature. The second column additionally includes per-pupil expenditures in K-12 education, as well as variables that measure the availability of federally-funded spots in Head Start, the generosity of the eligibility requirements for public health insurance, and the economic situation of the state. The third column of Table C1 shows the results of regressing the post-pre-K indicator on the same set of state economic conditions and policy generosity measures. This allows me to explore the economic and policy context in treatment states not only at the year when a pre-K program is implemented, but also in the subsequent years within the sample period.

The estimated coefficients for the democratic control variables are small and not statistically significant, which indicates that the political party in control of the state government and legislature do not play a significant role in the timing of the implementation of a pre-K program, conditional on state and year fixed effects. The estimates for the state unemployment rate and median income are also not significantly different from zero, which suggests that investments in

[51] The results are qualitatively the same when the regressions are estimated without weighting states by population (available upon request).

pre-K are not responding to the economic cycle. Additionally, the introduction of pre-K is not significantly associated with changes in expenditures in K-12 education or eligibility requirements for children's public health insurance programs. There is only a significant negative effect at 10% level of Medicaid/CHIP income eligibility cutoff for children under 5 on the year of implementation of pre-K, which would imply that pre-K implementation is associated to a lower income requirement, but the effect on the indicator for all post-pre-K implementation years is small and not statistically significant.

The year of implementation of pre-K is positively correlated with federally funded Head Start enrollment during the same year. If states invest more in pre-K at the same time as the federal government invest in Head Start, the estimated effects of pre-K (if Head Start enrollment was not included as a control) could confound the effects of the two programs. The estimated effect implies that a one-percentage-point (8%) increase in Head Start enrollment is associated with a 5.9 percentage point increase in the likelihood of introducing a pre-K program in that year. However, when the outcome variable is Post Pre-K, which indicates all post pre-K implementation years, the effect of Head Start enrollment is negative and statistically insignificant. This suggests that the increased Head Start enrollment is not sustained throughout the years after pre-K implementation, mitigating the concern of confounding effects. The effect of per-pupil education expenditures in the years after pre-K implementation is non-negative, which mitigates the concern that pre-K may be competing for resources with K-12 education, and the fact that it is not significantly different from zero also mitigates the concern of confounding effects of simultaneous increases in K-12 expenditures when pre-K is implemented.

In Table C2, I explore the correlation of the implementation of pre-K programs with changes in state characteristics by running regressions at the individual level. I estimate regression models similar to Equation 2 (using the CPS sample of children ages 5 to 12 in treatment and control states), but where the individual demographic controls and the state-level economic and policy variables are used as outcome variables instead of controls. The idea is to check whether the implementation of a pre-K program, conditional on state and cohort fixed effects, is predictive of the economic circumstances of states, their demographic composition, or the generosity of other state social programs that may have confounding effects. Using individual-level information allows me to include the demographic characteristics of the children and their families, and to measure the state environment not only when children are four years old but also at the time when they are observed in the CPS sample.

Each row of Table C2 shows the estimated coefficient and standard error of the Post Pre-K variable on a different outcome variable. The policy variable has no effect on any of the demographic variables, which mitigates any concerns about changing demographic characteristics of the states. Additionally, there are no significant effects on any of the measures of the economic situation of the state when the child is observed or when the child was preschool age. This suggests that there are no diverging economic trends across treatment and control states after the implementation of pre-K. The results for state health insurance policies when children were four years old are generally consistent with the findings of the regressions at the state level presented in the last column of the previous table. In addition, the generosity of public health insurance for children ages 6-15 is not correlated with cohorts affected by pre-K programs. In terms of educational policies, the effect on per-student spending in K-12 education is not significant either. The estimated effect on Head Start enrollment at age 4 is not significantly different from zero, and the negative point estimate implies a negligible effect of 0.26 percentage points.

Table C1: Regressions of Pre-K Implementation on State Political and Policy Variables

	1st Year Pre-K Implemented	1st Year Pre-K Implemented	Post Pre-K
Governor is Democrat	-0.004	0.034	
	(0.054)	(0.043)	
Democratic Control of Legislative	0.003	0.037	
	(0.080)	(0.099)	
Per Pupil Education Expenditures		-0.061	0.021
		(0.054)	(0.072)
Fed. Head Start Enrollment at 4		0.053**	-0.038
		(0.025)	(0.031)
Medicaid/CHIP (1-5)		-0.271*	-0.012
		(0.145)	(0.100)
Medicaid/CHIP (6-15)		0.234	0.110
		(0.159)	(0.126)
Unemployment Rate		-0.021	-0.040
		(0.051)	(0.048)
Median Income		-0.004	-0.007
		(0.007)	(0.010)
N	297	297	297

Notes: Each column shows a separate regression for the outcome variable indicated in the column heading, with observations at the state-year level, weighted by the population of 4-year-olds in each state in the year 2000. The sample includes treatment and control states, between 1997 and 2005. See data section for a description of the variables and data sources. All regressions include state and year fixed effects. Robust standard errors (clustered by state) in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table C2: Prediction of Demographic and State Characteristics as a Function of Pre-K

Outcome variable	Post Pre-K Coefficient	S.E.	N
Individual Characteristics			
Female	-0.007	(0.008)	70644
Black	-0.015	(0.009)	70644
Hispanic	-0.001	(0.005)	70644
Other Race/Ethnicity	0.004	(0.004)	70644
Age	-0.028	(0.042)	70644
Mom's Education High School	0.005	(0.010)	70644
Mom's Education Some College	0.009	(0.012)	70644
Mom's Education College	0.004	(0.008)	70644
Married Mom	-0.004	(0.009)	70644
State Characteristics			
Fed. Head Start Enrollment at 4	-0.264	(0.205)	70644
Medicaid/CHIP (1-5) at 4	0.205	(0.129)	70644
Medicaid/CHIP (6-15)	-0.023	(0.054)	70644
Unemployment Rate at 4	-0.146	(0.162)	70644
Unemployment Rate	-0.154	(0.138)	70644
Median Income at 4	-0.258	(0.637)	70644
Median Income	0.107	(0.353)	70644
Per-Pupil Education Expenditures	0.144	(0.170)	70644

Notes: Each cell shows results for separate regressions using the October CPS sample of children observed 1-8 years after pre-K age, for the outcome variable indicated in the left column. All regressions include state and cohort fixed effects. See data section for a description of the variables and data sources. Robust standard errors (clustered by state) in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.