RAND CORPORATION

Informing Investments in Preschool Quality and Access in Cincinnati

Evidence of Impacts and Economic Returns from National, State, and Local Preschool Programs

Lynn A. Karoly and Anamarie Auger

Key findings

- There are numerous examples of full-scale preschool programs with rigorous evaluations that show improvements in school readiness for participating children.
- Favorable impacts have been demonstrated for part- and full-day preschool programs, as well as one- and two-year programs, but the research is not definitive about the comparative effectiveness of these options.
- High quality is a common element among the preschool programs with the largest effects on school readiness and with sustained effects at older ages.
- Children across the income spectrum may benefit from high-quality preschool, but the impacts tend to be larger for more-disadvantaged children.
- Although differences in achievement scores between preschool program participants and nonparticipants tend to narrow as they advance through the elementary grades, high-quality preschool programs show sustained benefits for other aspects of school performance, such as lower rates of special education use, reduced grade repetition, and higher rates of high school graduation.
- Improving the alignment between preschool and the early elementary grades may help sustain the initial boost in cognitive and noncognitive skills from preschool participation.
- High-quality preschool programs represent a significant investment of resources, but that investment may be paid back through improved outcomes during the school-age years and beyond.

In the past decade, a growing number of U.S. cities have redirected existing funds or identified new funding streams in order to expand access to and raise the quality of preschool programs serving children one or two years before kindergarten entry. Cincinnati, Ohio, is among these cities, motivated by an extensive body of evidence that purports to show that high-quality early learning programs can prepare children to enter kindergarten with the skills required for successful performance in the elementary grades and beyond. Evidence of positive economic returns to preschool programs provides further impetus for such investments on the part of the public sector.

Cincinnati has a long, rich history of investing in children and providing support and resources for early care and education programs, including preschool. For more than a decade, the Success by Six initiative of the United Way of Greater Cincinnati has strived to raise awareness about the importance of early learning programs for school readiness and engaged with multiple stakeholders to ensure the continued support of such programs. 4C for Children, in addition to its role as a resource and referral agency, has been a central hub for professional development offerings and coaching to strengthen the knowledge and competencies of the early care and education professionals in the region. Providers in Cincinnati also participate in the Ohio quality rating and improvement system, Step Up To Quality, which works with programs to identify strengths and areas where quality can be improved.

In Cincinnati, stakeholders in the public and private sector are further compelled to focus on early childhood investments because nearly half of the city's children younger than the age of five live in families with income below the federal poverty level.[1] Children in low-income families are more likely to enter school with fewer skills identified as important for school readiness.[2] The shortfall in school readiness skills is further

compounded as children move through the early elementary grades. If skill begets skill, these children are at risk for being on a lower trajectory throughout their schooling years, with diminished academic achievement and educational attainment relative to what they could have realized if they had entered school on par with their more economically advantaged peers.[3]

While many view this evidence base in favor of preschool investments as compelling, others point to research that casts doubt on the findings, with concerns about the generalizability of favorable impacts from small-scale demonstration programs or evidence that the initial gains for preschool participation eventually fade away. Often it is challenging for nonexperts to sort through the conflicting research evidence. An objective assessment of the available national, state, and local research is critical for stakeholders in Cincinnati and other communities that are seeking to allocate scarce resources toward those strategies that are most likely to produce beneficial and lasting impacts for children.

To inform stakeholders in Cincinnati regarding potential investments designed to expand preschool access and quality, this report seeks to address the following questions:

- Do high-quality preschool programs produce favorable effects for participating children and their families, what are the magnitudes of the impacts, and how long do the beneficial effects last?
- Is there evidence of a positive return on investment when public dollars are used to pay for such programs?

Our answers to these questions are based on a compilation and critical assessment of the most-rigorous evidence regarding the effects of publicly funded preschool programs in the United States at the national, state, and local levels. We provide evidence for specific programs, as well as results from syntheses across multiple preschool program evaluations. We assemble evidence of the impacts of the preschool programs on children's school readiness. In cases where children have been followed beyond the preschool years, we also consider research regarding longer-term effects.

Given the policy focus in Cincinnati on preschool programs for children one or two years before kindergarten entry (often referred to as preschool for three- and four-year-olds), such programs are the focus of our review. Thus, we do not consider the broader range of early childhood interventions that serve children at even earlier ages, such as home visiting programs or center-based programs for infants and toddlers.[4] In focusing on preschool programs, our primary interest is in findings for real-world programs currently implemented in states or local communities (e.g., school districts), supported with public funding. Such programs may be part-day or full-day, delivered during the academic year or year-round, and made available to targeted groups of children or all age-eligible children, regardless of circumstances.[5]

We proceed with two main sections. First, we consider evidence that preschool programs improve school readiness and later outcomes once participants enter school. Second, we examine the evidence regarding preschool program costs, benefits, and economic returns. The final section identifies the implications for decisionmakers in Cincinnati seeking to design an effective and cost-beneficial preschool investment strategy for the city.

EVIDENCE THAT PRESCHOOL WORKS

Rigorous evaluations of center-based preschool programs were conducted as early as the 1960s. The experimental evaluation of Perry Preschool, a demonstration program that began in 1962 in Ypsilanti, Michigan, is perhaps the best-known randomized control trial of a high-quality preschool program. This is, in part, because the children assigned randomly to the program and their control group counterparts were followed through their school-age years and into adulthood to measure longer-term impacts. The published findings of the follow-up to age 40 demonstrate that the Perry Preschool program had lasting favorable effects on the participants, with higher earnings, lower welfare use, and reduced criminal activity relative to the control group, among a host of other measured outcomes.[6] Although the Perry Preschool program findings are compelling, policymakers considering investments in preschool need evidence from real-world programs that are currently operating in school districts, cities, states, or nationwide. We focus on the body of evaluation evidence for such programs in this section. Before presenting those results, we discuss the importance of rigorous evaluation methods for capturing the causal effects of preschool programs on children's outcomes.

Focus on Real-World Programs with Rigorous Evaluations

Our review centers on full-scale national-, state-, or district-level publicly funded preschool programs, currently operating in the United States, which have been the focus of a

rigorous evaluation to assess the effects of participating in the program on kindergarten readiness and/or subsequent educational outcomes. In total, we examine the evaluation findings for one national program (Head Start), 11 state-funded programs (Arkansas, Georgia, Michigan, New Jersey, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, West Virginia, and Washington), and three district-level programs (Boston, Chicago, and Tulsa, Oklahoma).[7] Results from the small-scale Perry Preschool evaluation are included as a point of reference because it is frequently cited in discussions of preschool program effectiveness.

These 15 full-scale programs share a common objective of delivering a high-quality preschool program to children one or two years before entering kindergarten. Programs may be delivered in center-based settings, either in public schools and/or in such community-based settings as private for-profit or nonprofit preschool programs, child care centers, or nursery schools. Beyond this shared aim, there is considerable variation across the programs in their features, such as whether they are universal or targeted and, if so, the specific target population; the duration (one or two years) and intensity of the program (i.e., hours per day and weeks per year); the structural features, such as group sizes and ratios, the education and training requirements of the classroom staff, the nature of the curriculum employed in the classroom, and the professional development support for staff; the resulting quality of the learning environment experienced by participating children; and the provision of other supports to children and families beyond early learning services (e.g., the requirements for Head Start to support referrals to social services and to address the physical, mental, and oral health needs of participating children and families through screenings, health checks, referrals for treatment, and access to health insurance). Several of these features are summarized in Table A.1 in Appendix A. As a result of these and other features, the estimated annual spending to implement these programs ranges from about $3,600 per child to $13,400 per child. In sum, the programs we reviewed represent a diverse mix of publicly funded programs that we collectively refer to as "preschool," although they may also be referred to as "prekindergarten" or "preK" programs.

The long tradition of evaluating preschool programs stems from an interest in understanding whether a given program is having its intended effects in terms of preparing children for school and perhaps even promoting success after school entry. How do we know if a preschool program works? If we simply observe the outcomes of a child before and after preschool

The long tradition of evaluating preschool programs stems from an interest in understanding whether a given program is having its intended effects in terms of preparing children for school and perhaps even promoting success after school entry.

participation, we may see gains (or declines) in their development, but we will not know whether those changes were the result of the program or caused by some other factors, such as characteristics of the child or family. In other words, the child's outcomes may have been equally favorable if he or she had not attended the program. To measure the cause-and-effect relationship of the program, we need to compare a child's outcomes when he or she participates in the program with what that child would have experienced if there been no program at all or if the child had participated in an alternative program (i.e., the status quo), holding everything else constant. Because we cannot observe outcomes for the same child both when he or she participates in the preschool program and when the child does not, the counterfactual must be generated by observing the outcomes of a control or comparison child who experiences the alternative but is otherwise the same as the participating child. Collectively, we seek to compare the outcomes for participating children with those of a control or comparison group of similar nonparticipants.

Experimental studies randomly assign individuals to participate in a program (the treatment group) or not to participate in a program (the control group). When implemented well, the random assignment ensures that the treatment and control groups are matched on observable characteristics, as well as

characteristics that cannot be observed by the researcher. Any differences in outcomes between the two groups, at any given point in time after random assignment, would then be attributable to the only systematic factor that differs between the them—that is, participation in the program. One caution is that experimental studies do not necessarily provide the most-rigorous evidence, because poor execution (e.g., assignment to groups that is not truly random) or other problems—such as high rates of nonrandom attrition when participants are followed over time—can compromise the results. Experiments with small sample sizes in the treatment and control groups also have low statistical power to detect small program effects compared with experiments that have large study samples. Moreover, random assignment studies, while often considered the gold standard in program evaluation, are not always feasible because of universal provision of the program, resource constraints, ethical concerns, or other issues.[8]

In the absence of random assignment, researchers face the challenge of controlling for all other factors that might explain differences between children who participate in the preschool program and those who do not. Quasi-experimental methods aim to generate a control group that is as close to the randomized gold standard as possible, often by using statistical methods to mimic random assignment.[9] One strategy is to use a control group that has similar characteristics as the treatment group but lives in an area where the program is not available or includes individuals on the waiting list for an oversubscribed program. In such cases, members of the comparison group are likely candidates for the program, but they do not participate because of factors outside their families' control (e.g., the neighborhood in which the services are located or a lottery that selects participants from a waiting list).

Another quasi-experimental method employed in a number of recent preschool evaluations is to use the "randomness" of birth as a type of natural experiment that determines which children enter a program (because they meet a strict age cutoff) versus those who do not (e.g., children who are age four by September 1 and eligible for preschool, and those who turn four on September 1 or after and are therefore not eligible). This approach uses what is a called a regression discontinuity research design because the birth date cutoff creates a break or "discontinuity" in the continuous age spectrum between one cohort of children that participates in the program in a given year versus those that must wait another year to enroll.[10] One drawback of the regression discontinuity design is that it can be employed to study only the effects of preschool on kindergarten readiness and then only for a one-year program. Other statistical methods are required to capture preschool program impacts in kindergarten and beyond or for a two-year program. To measure such impacts, researchers have turned to other methods to create valid comparison groups among preschool nonparticipants, such as propensity score matching.

The quasi-experimental methods, in addition to defining an appropriate control group, use statistical methods to control for possible (observable) confounding factors that may be associated with the child outcomes of interest (e.g., child characteristics, parent education, motivation, or other family background factors). Ultimately, the quality of any quasi-experimental evaluation rests on the ability of the control or comparison group to serve as a valid counterfactual for the preschool program participants. The regression discontinuity design is generally viewed as a valid alternative to an experiment evaluation for causal inference. Other methods are subject to the critique that the analyst has not controlled for potential confounding factors that could explain the observed outcomes.[11] Nevertheless, when an experimental evaluation is not possible, such methods are the only option for making inferences about the effects of a preschool program beyond school readiness.

The use of experimental and rigorous quasi-experimental evaluation methods provides a higher level of confidence that researchers have measured the true causal effect of a preschool program on children's outcomes. It lessens the possibility that results will be biased by failing to control for other confounding factors. For this reason, we focus the discussion that follows on the evaluations of 15 preschool programs that use either random assignment or a quasi-experimental design. Table A.2 in Appendix A records the evaluation method and the point of last follow-up.

Regardless of the evaluation design, a key issue in interpreting the results of a preschool program evaluation is the nature of the counterfactual (i.e., control group) condition. Especially for more-recent evaluations of large-scale preschool programs in the United States, the control group is not necessarily a "no preschool" comparison, as children experiencing the status quo may participate in some other preschool program or type of center-based care. Indeed, the counterfactual condition has been changing through time. For example, when the Perry Preschool program was evaluated in the early 1960s using an experimental evaluation, children assigned randomly to the control group likely did not have access to any other subsidized early learning programs. Thus, the impacts for Perry Preschool

The use of experimental and rigorous quasi-experimental evaluation methods provides a higher level of confidence that researchers have measured the true causal effect of a preschool program on children's outcomes.

capture the effect of the program compared with not participating in any kind of early learning program. In contrast, when the Head Start program was evaluated using a randomized design in the early 2000s, 48 percent of the children in the control group attended some other center-based child care or preschool program, many of them a different Head Start program.[12] Thus, rather than evaluating Head Start against an alternative of no preschool program, the national experimental study evaluated the effectiveness of Head Start compared with the array of other early learning programs in the community. Because the number of early learning programs has expanded over time, it is likely that when evaluating any given preschool program in more recent years, many children in the control group will experience some other form of preschool rather than no preschool at all. If the alternative programs experienced by the control group produce any favorable impacts relative to no program, we would expect the impacts from more-recent preschool evaluations to be smaller compared with those conducted decades in the past.[13] It is also important to note that parental investments in children, such as providing books in the home, have increased over time, which also potentially lessen the impact of high-quality preschool compared with the past when parents were less likely to make such contributions.[14]

Evidence of Preschool Program Effectiveness for Full-Scale Publicly Funded Programs

We now turn to the findings from the evaluations of the programs listed in Tables A.1 and A.2 in Appendix A. We begin with evidence of the effects of preschool participation on school readiness. We then turn to evidence of effects on outcomes during the school-age years and beyond.

Effects of Preschool on School Readiness

Following evaluations of programs like Perry Preschool and Head Start in the 1960s, much of the initial interest in preschool program evaluation concerned the ability of programs to boost children's cognitive development, measured by IQ or specific skills in such domains as early language, literacy, and mathematics. Fewer studies at that time and since then have considered the effects of programs on social and emotional skills.

Table 1 lists the 13 programs that have measured the causal effect of preschool participation on cognitive measures of school readiness, assessed either at the end of the preschool program or at the time of kindergarten entry. Results from the evaluations of Perry Preschool and Head Start are listed in panels (a) and (b), respectively. State and district program results are shown in panels (c) and (d), respectively. The Table 1 findings are based on an experimental design in the case of Perry Preschool and Head Start; statistical controls in the case of the Chicago Child-Parent Centers (CPC) program; propensity score matching in the case of the Tennessee program; and the regression discontinuity design for all others.[15]

The table shows results for a set of developmental assessments in early mathematics, reading, and language skills, although most studies used different combinations of the listed measures. In particular, most evaluations used one or more subtests of the Woodcock-Johnson Test of Achievement, which measures premathematics reasoning (the Applied Problems subtest), prereading or reading skills (the Letter-Word Identification subtest), and prewriting or spelling skills (the Spelling subtest). A majority of the studies also used the Peabody Picture Vocabulary Test (PPVT), a test of receptive vocabulary that is predictive of general cognitive abilities. Five evaluations used a measure of print awareness[16] (namely, the print awareness subtest of the Preschool Comprehensive Test of Phonological and Print Processing [Pre-CTOPP] or the Test of Preschool Early Literacy [TOPEL]). The Chicago CPC evaluation assessed chil-

Table 1. Effects of Preschool Programs on Cognitive Measures of School Readiness

Program (PreK Cohorts Studied[a])	Effect Size: Woodcock-Johnson Subtest: Applied Problems	Woodcock-Johnson Subtest: Letter-Word Identification	Woodcock-Johnson Subtest: Spelling	Vocabulary (PPVT)	Print Awareness (Pre-CTOPP)	Cognitive Readiness[b]
a. Demonstration Program						
Perry Preschool (1962–1964)	–	–	–	1.02*	–	–
b. National Program						
Head Start (2002)	0.15	0.32*	0.24*	0.08	–	–
c. State Programs						
Arkansas (2005)	0.24*	–	–	0.36*	–	–
Georgia (2011)	0.51*	1.05*	–	–	–	–
Michigan (2004)	0.47*	–	–	–0.16	0.96*	–
New Jersey (2004)	0.23*	–	–	0.36*	0.50*	–
New Mexico (2005)	0.38*	–	–	0.35*	1.17*	–
New Mexico (2006)	0.50*	–	–	0.25*	0.59*[b]	–
New Mexico (2007)	0.43*	–	–	0.17*	1.15*[b]	–
Oklahoma (2004)	0.35	–	–	0.29*	0.43	–
South Carolina (2004)	–	–	–	0.05	0.79*	–
Tennessee (2009, 2010)	0.17*	0.41*	0.29*	–	–	–
West Virginia (2004)	0.11	–	–	0.14	0.83*	–
d. District Programs						
Boston (2008)	0.59*	0.62*	–	0.44*	–	–
Chicago CPC (1983–1985)	–	–	–	–	–	0.46*
Tulsa (2002)	0.38*	0.79*	0.64*	–	–	–

SOURCES: See Table A.2 in Appendix A and evaluation studies list in Appendix B.
NOTES: For full program names, see Table A.1 in Appendix A. The effect sizes are for the treatment-on-treated program impacts, i.e., the program's effects on those who participate. The estimate for Perry Preschool is after the end of the first program year for children who entered at age three or four. The estimate for Chicago CPC is the lower bound from various estimation methods that control for potential selectivity bias. Estimates for Head Start are after the first program year for children who entered the program at age four. Estimates for Michigan, New Jersey, Oklahoma, South Carolina, and West Virginia are based on the pooled sample regression discontinuity model. * = statistically significant at the 5-percent level or better. – = not measured.
[a] Indicates year(s) for fall entry of preschool cohorts studied.
[b] Print awareness assessed using the print knowledge subtest of the TOPEL.

dren at kindergarten entry using the Iowa Test of Basic Skills, a cognitive school readiness subtest. In all cases, we report preschool program impacts in terms of effect sizes, which provide a standardized measure of impact across all outcome measures.[17] It is common in the literature to consider effect sizes of 0.2 to be small, 0.5 to be medium, and 0.8 to be large, although in comparison to other education interventions, effect sizes of 0.3 to 0.5 are considered to be reasonably large.[18]

A comparison across the rows of Table 1 indicates that each program had at least one statistically significant impact on a measure of school readiness. However, the pattern of significant coefficients and their magnitudes varies considerably across the programs. The Boston Public School PreK program—which is open to all children in the district regardless of income and currently reaches about half of all entering kindergartners in the district—shows some of the largest impacts, with effect sizes ranging from 0.44 to 0.62. Tulsa's universal preschool is in a similar range from 0.38 to 0.79. A number of other programs produced effect sizes in the range from 0.5 to 1.0, although some outcomes like the measure of print awareness are understood to be easier to realize developmental gains. The PPVT is the only measure of readiness at the start of kindergarten that

the Perry Preschool evaluation has in common with the other studies in Table 1. The effect size for Perry Preschool of 1.0 is about two to four times as large as what was achieved in the large-scale programs for the same measure. Such attenuation of impacts in moving from a small-scale demonstration program to programs implemented at full scale is what is typically expected.[19]

Table 1 illustrates the estimated impacts from specific preschool programs that are implemented at scale (with the exception of Perry Preschool) and evaluated using rigorous designs. Looking across a wider set of preschool evaluations using experimental or quasi-experimental designs and pooling across assessments for language, literacy, and mathematics measured soon after the preschool program ended, one recent meta-analysis concluded that the average impact across the 84 preschool programs studied was an effect size of 0.21 to 0.35, which is equivalent to about one-fifth to one-third of a year of learning.[20] Consistent with this finding, a different meta-analysis of 59 preschool program evaluations estimated an average effect size of 0.31 for measures of child development immediately after the preschool program ended.[21]

Fewer studies have assessed preschool program effects on social and emotional skills in the short or long term. The evaluation of Boston's PreK program is one exception, with evidence of significant favorable impacts on executive function and emotional regulation prior to kindergarten entry, even though the program did not directly target the development of these skills. However, the effect sizes are in the 0.2–0.3 range, smaller than those for the cognitive assessments (see Table 1).[22] The evaluation of Tulsa's universal preschool program has assessed the effects on socio-emotional outcomes, with findings of favorable effects on timidity and attentiveness, but again the gains are relatively modest (effect sizes of 0.15 to 0.19) compared with the cognitive benefits.[23] In evaluating Georgia's universal program, the researchers found no significant effect on teacher-assessed measures of social skills and problem behaviors.[24]

Evidence also comes from evaluations of specific preschool curricula that focus on social and emotional learning. For example, a large-scale evaluation of three approaches—Incredible Years, Preschool PATHS, and Tools of the Mind–Play—designed to promote children's social emotional development in Head Start centers generated mixed results. Two of the approaches—Incredible Years and Preschool PATHS—showed a positive effect on children's learning, social behaviors, and emotional knowledge. All three approaches had a positive effect on teacher practices, although which teacher practices were affected differed by approach.[25]

> The Boston Public School PreK program—which is open to all children in the district regardless of income and currently reaches about half of all entering kindergartners in the district—shows some of the largest impacts.

Effects of Preschool on Success in School and Beyond

With an expectation of continued benefits for preschool participants as they move through the elementary grades, several studies of large-scale preschool programs have estimated the effects of subsequent measures of school performance, including academic achievement scores, grade retention, and special education use. Keeping in mind that, with the exception of the experimental evaluations, estimating these downstream impacts requires researchers to use other quasi-experimental designs in comparing children who did and did not participate in the preschool program, approaches that are not as rigorous as those used to assess effects on school readiness (i.e., those reported in Table 1).

An exception is the Perry Preschool program, which has the advantage of long-term follow-up and an experimental evaluation, albeit for a small-scale demonstration program. Key findings from the Perry Preschool evaluation through the follow-up to age 40 are listed in panel (a) of Table 2. Notably, Perry participants had significantly higher reading and mathematics assessments as late as sixth grade. Although the test score advantage did not persist into later grades, participants had a lower rate of special education use and a higher high school graduation rate, along with other long-term impacts.[26]

Table 2. Preschool Programs with Sustained Effects in the Elementary Grades and Beyond

Program (PreK Cohorts Studied[a]) and Outcome	Grade or Age of Measurement	Unit of Measurement	Treatment–Control Difference	Effect Size
a. Demonstration Program				
Perry Preschool (1962–1964)				
Reading achievement	Grade 6	Normed score	3.9*	0.34*
Mathematics achievement	Grade 6	Normed score	4.4*	0.33*
Grade retention	By age 27	Years	–0.2	–0.15
Special education use	By age 19	Percentage of years	–12*	–0.29*
High school completion	By age 27	Percentage points	21*	0.43*
b. State Programs				
Michigan (1995)				
Satisfactory on state reading test	Grade 4	Percentage points	8*	–
Satisfactory on state mathematics test	Grade 4	Percentage points	8*	–
Grade retention (ever)	By grade 12	Percentage points	12*	–
On time high school graduation	By grade 12	Percentage points	14*	–
New Jersey (2005)				
Literacy / language arts achievement[b]	Grade 5	Normed score	–	0.18* / 0.22
Mathematics achievement[b]	Grade 5	Normed score	–	0.14 / 0.29*
Grade retention[b]	By grade 5	Percentage points	–8* / –7*	–
Special education placement[b]	Grades 3 to 5	Percentage points	–6* / –4*	–
North Carolina (2002, 2003)				
Special education use	Grade 3	Percentage reduction in placements per $1,000 spending	–32*	–
Washington (2003–2008)				
Reading achievement	Grade 5	Standardized score	–	0.23*
Mathematics achievement	Grade 5	Standardized score	–	0.16*
c. District Programs				
Chicago CPC (1983–1985)				
Reading achievement	Grade 8	Normed score	5.4*	0.24*
Mathematics achievement	Grade 8	Normed score	4.3*	0.23*
Grade retention (ever)	By age 15	Percentage points	–15.4*	–0.34*
Special education use	By age 18	Percentage points	–10.2*	–0.26*
High school completion	By age 28	Percentage points	6.4*	–
Any substance abuse (excluding alcohol)	By age 28	Percentage points	–5.2*	–
Any arrest	By age 27	Percentage points	–6.4*	–
Tulsa (2005)				
Mathematics achievement	Grade 3	Normed score	18*	0.18*

SOURCES: See Table A.2 in Appendix A and evaluation studies list in Appendix B.
NOTES: For full program names, see Table A.1 in Appendix A. * = statistically significant at the 5-percent level or better; – = not reported.
[a] Indicates year(s) for fall entry of preschool cohorts studied.
[b] Effect for one year of participation followed by effect for two years of participation.

Among the full-scale programs listed in Table A.1, eight have conducted follow-up studies beyond kindergarten entry (see Table A.2). To examine outcomes beyond school entry, sufficient time must pass to observe later outcomes. Given the expense of collecting data from students, most studies rely on student-level administrative data, which means a focus on outcomes from the third grade onward when most states begin using statewide standardized tests. Administrative records also may provide information on special education use and grade retention. The Chicago CPC evaluation, which extends to age 28, collected some information through surveys of individuals in the treatment and comparison groups.

Panels (b) and (c) of Table 2 list findings from six of the eight full-scale state- or district-level preschool programs that have used various methods to identify longer-term effects *and* found sustained impacts.[27] Sustained effects of programs were found for reading and mathematics standardized achievement tests as late as eighth grade (Chicago CPC), but at least through third, fourth, and fifth grades (Michigan, New Jersey, North Carolina, Washington, and Tulsa [for one cohort of children]).[28] It is important to note that the magnitudes of the effects on student achievement are in the range of 0.2 to 0.3; smaller in magnitude compared with those at kindergarten entry (see Table 1). Again, this is consistent with findings from meta-analyses, which show a drop in the magnitude of achievement differences between preschool program participants and nonparticipants at each successive grade.[29]

At the same time, significant reductions in grade retention and special education use were found as well over the entire schooling years (Michigan and Chicago CPC) or in the elementary grades when that was the last follow-up (New Jersey and North Carolina). For example, the percentage of preschool program participants repeating a grade by age 15 in Michigan was lower by 15 percentage points and the rate of being placed in special education was lower by 10 percentage points. In New Jersey, the effects were about half as large but measured only through fifth grade. Michigan and Chicago CPC found eventual favorable effects on the high school graduation rate, with increases of 6 percentage and 14 percentage points, respectively. And Chicago CPC demonstrated impacts well into adulthood on substance abuse and contact with the criminal justice system, among other long-term impacts.[30]

Is Fadeout a Concern?

Table 2 highlights large-scale preschool programs in four states and two school districts that have produced sustained effects into later grades and even into adulthood. At the same time, there is growing attention to the issue of the "fadeout" of preschool program impacts.[31] A narrowing of the achievement test score advantage for those who attended preschool versus those that did not is a phenomenon present in a number of preschool program evaluations that continue to follow participating children after they enter kindergarten.[32] Two studies with longitudinal follow-up, in particular, have garnered attention in this regard: the experimental evaluation of the Head Start program conducted in the early 2000s, and the ongoing quasi-experimental evaluation of Tennessee's statewide preschool program.

Head Start

The national Head Start Impact Study, after finding modest effects on measures of school readiness (see Table 1), showed no sizable advantage to Head Start participation by the end of kindergarten, first grade, and third grade.[33] However, several subsequent reanalyses of the national Impact Study data suggest a more-nuanced interpretation of the findings. First, as noted earlier, with nearly half of the children in the control group in another early learning program (including other Head Start centers), the estimated impacts of Head Start are not relative to a no-preschool alternative. When impacts are estimated for Head Start participants who otherwise would have been in home care, the positive effects of Head Start are strengthened.[34] Second, researchers have looked for differential effects of Head Start based on the child's background characteristics and found larger impacts of Head Start for some groups of children, such as Spanish-speaking children with larger skills deficits when they started Head Start and children whose mothers do not have a high school degree.[35]

Another consideration in interpreting the evaluation findings for Head Start is that the program is not uniform across the centers operated by grantees. Rather, Head Start is implemented in each community following federal standards that currently allow some variation in the education and training background of lead teachers, choice in the curriculum to employ, and differences in other key program features. At the time of the national Head Start Impact Study, fewer than half of Head Start teachers nationwide had a bachelor's degree, as contrasted with most of the preschool programs listed in Table 2 with longer-term impacts—namely, those in Boston, Chicago, Tulsa, Michigan, and New Jersey—all of which require a bachelor's degree and specialized early childhood training (see Table A.1). Other data show that at the time of the national evaluation, compared with the programs in Boston or Tulsa,

Table 3. Effects by Subgroups of Tulsa, Oklahoma, Preschool Program on School Readiness

Population Group	Woodcock-Johnson Subtest Effect Size		
	Applied Problems	Letter-Word Identification	Spelling
All children	0.38*	0.79*	0.64*
By race-ethnicity			
White	n.s.	0.76*	0.72*
Hispanic	0.99*	1.50*	0.98*
Black	0.38*	0.74*	0.52*
Native American	0.60	0.89*	0.72*
By free lunch eligibility status[a]			
Eligible for free lunch	0.45*	0.81*	0.65*
Eligible for reduced-price lunch	n.s.	1.04*	0.97*
Not eligible	0.29	0.63*	0.54*

SOURCE: William T. Gormley, Ted Gayer, Deborah Phillips, and Brittany Dawson, "The Effects of Universal Pre-K on Cognitive Development," *Developmental Psychology*, Vol. 41, No. 6, 2005, pp. 872–884.
NOTES: * = statistically significant at the 5-percent level or better. n.s. = not significant at the 10-percent level and effect size not reported.
[a] Children in families with income below 130 percent of the federal poverty level are eligible for the free lunch program; students in families with income from 130 percent to 185 percent of the federal poverty level are eligible for the reduced-price program.

Head Start programs on average scored lower on global quality measures that capture key dimensions of teacher-child interactions.[36] Such differences in quality, on average, may explain the lower initial effects of Head Start participation on measures of school readiness compared with the initial effects (see Table 1) and sustained effects (see Table 2) for programs like those in Boston, Chicago, Oklahoma, or New Jersey that are implemented with uniform, high-quality standards across all schools and centers.

It is also worth noting that another body of research on Head Start has examined evidence for longer-term effects beyond what can be learned from the national experimental evaluation. These studies use a variety of quasi-experimental methods to control for other factors that might confound our ability to measure the causal effects of Head Start at older ages.[37] Together, these studies have found persistent favorable effects of Head Start on academic achievement scores at older ages; on educational attainment, including high school graduation and college enrollment; on other lifecourse outcomes such as criminal behavior, teen pregnancy, and aspects of health, albeit with some differences in the magnitude of the impacts across subgroups defined by gender or race-ethnicity.[38]

Tennessee Voluntary Prekindergarten

More recently, results for a quasi-experimental substudy of the evaluation of the Tennessee state-funded preschool program found that the test score boost for preschool attendees versus nonattendees observed at the end of the preschool year (see Table 1) was no longer evident by the end of kindergarten, nor at the end of the first, second, and third grades.[39] Indeed, by the end of second grade, Tennessee children who attended the preschool program had lower outcomes on most measures compared with the nonparticipants. To date, the evaluation of the Tennessee program has published preliminary findings for other measures of school performance based on administrative data for the first cohort of the full experimental study sample. In particular, preschool participants were significantly less likely to repeat a grade by the end of kindergarten (4 percent for preschool attendees versus 8 percent for nonattendees) and had significantly higher attendance by the end of first grade (a differential of three and a half days).[40] Neither effect is particularly large, but future analyses of school administrative data will determine whether these impacts (and others such as special education use) are amplified as children progress through the early elementary grades, as has been found in other studies (see Table 2).

In the debate over the Tennessee findings, the same concerns mentioned for Head Start also are relevant. First, for the children who were randomly not offered a space in the state-funded program, the comparison condition was not a "no preschool program" group. Although about half (51 percent) were cared for by a parent or guardian, 38 percent attended a Head Start program or private child care center that was not part of the state-funded preschool program.[41] Second, the quality of the Tennessee program may not reach the level of quality achieved by the preschool programs that have demonstrated sustained impacts, such as those in Table 2. Notably, the average spending per child in Tennessee's program is lower than the spending levels in Boston, Oklahoma, or New Jersey

(see Table A.1). The rapid expansion of the program in the past decade to more than 900 classrooms in 95 counties—without a mechanism to ensure high-quality implementation, such as the quality monitoring and improvement features employed in the New Jersey Abbott districts—may contribute to the mixed level of quality. Indeed, measures of classroom quality collected as part of the Tennessee evaluation show considerable variation across programs in several quality measures, with about 7 percent of classrooms in the "minimal" to "inadequate" quality range on a commonly used measure of global quality and just 15 percent of programs in the "good" to "excellent" quality range.[42] Information collected on time use in the classrooms shows, on average, about 40 percent of the four hours in the program morning had no instructional content.

Hypotheses Regarding Fadeout

In light of the findings for Head Start and Tennessee, as well as the pattern of declining test-score effect sizes, even among programs with sustained impacts, researchers have debated the reasons for the fadeout phenomenon as it pertains to achievement scores in particular.[43] One explanation is that, in settings where high-quality preschool programs have been evaluated, teachers in kindergarten and beyond have not had sufficient time to adapt their curricula and teaching practices to the children arriving in their classrooms with more advanced skills. If they continue to emphasize material that has been mastered, such as basic counting or shapes, children with the preschool experience are not able to advance their learning. A second hypothesis is that once children arrive in elementary school, teachers devote compensatory time to those children without a preschool experience in order to help them catch up with their better-prepared peers.

According to these two explanations, the pattern is one of convergence rather than fadeout, as the children who arrive well-prepared for school are not able to capitalize on the gains they made in the preschool year and the children who arrived behind are provided with more resources in order to catch up. This suggests that, as participation in preschool programs rises, there will be fewer children that need the compensatory supports and teachers will be able to work with all children in the classroom to build on earlier gains and advance them at a similar pace. Indeed, the fact that there is some evidence of sustained effects on achievement scores in Oklahoma's universal program and New Jersey's program, which is universal in the Abbott districts, is consistent with this hypothesis. If the universal program in Boston shows evidence of sustained impacts beyond kindergarten entry, this also would be further support for the importance of reaching a high level of participation in preschool, particularly for disadvantaged children.

Another hypothesis for the fadeout pattern is that children in effective preschool programs may continue on to a lower-quality elementary setting, so the readiness gains are not sustained. There is evidence that this has been the case for children in Head Start programs.[44] One explanation for the longer-lasting effects of the Chicago CPC program is that the two-year program was aligned with a K–3 program in the same schools, which maintained the quality of the early learning experience. Notably, the evidence from the quasi-experimental evaluation of Chicago CPC indicates that children in the preschool-to-third grade (P–3) intervention, compared with a preschool-only intervention, showed greater increases in academic skills and lower dropout and remediation rates.[45] Thus, as we discuss later in this report, it is important to consider the alignment between investments in preschool with the opportunities for continued development once children enter kindergarten and beyond.

Whether labeled "convergence" or "fadeout," it is important to keep in mind that the pattern is typically seen for achievement scores, and, when measured, for social and emotional skills as well. As noted in Table 2, longer-term follow-up for programs like Chicago CPC and Michigan has demonstrated sustained favorable impacts on other areas of educational performance, such as special education use, grade retention, and high school completion or beyond (in the case of Chicago CPC). Evaluations of Tennessee's program, and oth-

Researchers have debated the reasons for the fadeout phenomenon as it pertains to achievement scores in particular.

This means that investing in preschool programs without investing in high-quality programs may not yield even partial benefits relative to the most-effective programs.

ers in Table 2 where fadeout is evident, have not yet analyzed longer-term effects on these other educational outcomes.

Features of Effective Preschool Programs

As noted earlier in this report, the preschool programs we considered in our discussion vary in important ways, such as the requirements for lead teachers' educational backgrounds, the curriculum to use in the classroom, who is eligible (targeted versus universal), and the intensity of the program (part-day versus full-day). To what extent do these program features affect the outcomes of participating children? The evaluation literature provides some answers to this question, although there are important gaps in our understanding of which program features are most important.

Process and Structural Quality

Evaluations of preschool programs have largely assessed the short- or longer-term impacts of specific bundles of preschool features that comprise a given program, such as those summarized in Table A.1 for our reviewed programs. The evaluation methods do not support making inferences about which features of a program are essential for obtaining a given impact or if some modified combination of program features would have an even larger effect. Thus, researchers have inferred which program features appear to be consistently present in successful programs or those that are strongly predictive of children's developmental gains based on naturally occurring variation in program features.

Drawing on this evidence base, recent syntheses conclude that features defined as "process quality"—the experiences children have in the preschool setting through interactions with teachers, other adults, and peers—are the strongest predictors of developmental gains in cognitive, social, and emotional domains.[46] Key aspects of process quality in preschool programs include a warm and supportive emotional climate characterized by responsive teacher-child relationships, and teacher-child interactions that provide a language-rich environment supporting learning in specific content areas (e.g., early literacy and math), while also promoting higher-order thinking skills.

Research also suggests a role for a set of "structural quality" features that provide the conditions that support the realization of high process quality. These structural features include those often measured or regulated: the group size, staff-child ratio, teacher qualification, and curriculum. Effective programs achieve and sustain process and structural quality through ongoing systematic measurement tied to quality improvement. For teachers, this is realized, in part, through coaching, mentoring, and other forms of professional development that foster effective use of a curriculum, provide feedback from experienced professionals, encourage reflective practice, and support improvement of knowledge, skills, and competencies.[47] As further research refines this list, there also is a need to develop valid and reliable measures to capture these important dimensions of program quality.

Another important consideration is that the relationship between program quality and children's development may not be linear—in other words, the expectation may not hold that increasing quality will lead to better outcomes for children no matter where a program begins on the quality spectrum. Indeed, there is evidence to suggest that there are thresholds in quality, below which children receive little benefit from preschool participation.[48] This means that investing in preschool programs without investing in high-quality programs may not yield even partial benefits relative to the most-effective programs.

Targeted Versus Universal

Beyond aspects associated with quality, another key feature of publicly funded preschool programs is whether enrollment is limited to qualifying children or available to all age-eligible children regardless of circumstances. Targeted programs, such as Head Start, have their origins in a compensatory model of preschool education that aims to provide early learning sup-

ports for children most at risk of developmental delays or not being ready to enter kindergarten.[49] Risk status is typically associated with low income, but targeting may be based on other risk factors, such as children who are homeless or in the child welfare system. Targeted programs are designed to first serve those children who can benefit most from a high-quality early learning experience. At that same time, because risk status may be a function of multiple factors and vary through time, the targeting mechanisms may be imperfect. Some children who would benefit are not eligible and many who are eligible do not participate, perhaps because of the hurdles of eligibility determination or other factors, such as stigma. Universal programs, because they are available to all children regardless of circumstances, do not require establishing or implementing eligibility rules. Consequently, stigma may be minimized and higher rates of participation realized, albeit with higher costs as more children are eligible. In addition, if children are more likely to be in mixed-income classrooms, universal programs may be more conducive to peer learning. A hybrid model is also possible, in which all children are eligible to enroll but the degree of public funding is a function of need (e.g., a sliding scale fee).[50]

Tables 1 and 2 demonstrated favorable effects on school readiness and other indicators of success in school and beyond for both targeted programs (the majority of those in Tables 1 and 2) and universal programs. Indeed, there is no indication that universal programs, such as those in Boston and Oklahoma, realize any smaller effects from serving children across the income spectrum. Table 3 shows additional results for the evaluation of the Tulsa, Oklahoma, universal preschool program. The effects of the program on school readiness for the full sample were provided in Table 1. Table 3 shows that the favorable effects on school readiness were shared across diverse groups defined by race-ethnicity and economic status as measured by eligibility for the free and reduced-price school lunch program.

For the three measures of school readiness, effect sizes were almost always statistically significant across the race-ethnic and economic status subgroups. The school readiness effects tended to be larger for Latino students compared with children from other race-ethnic backgrounds. Effects were consistently smaller for the higher income group (those not eligible for free or reduced-price lunch). Effects were not always the largest, however, for the children in the lowest income group. Similar findings of larger effects for more-disadvantaged children were also found in the evaluations of the preschool programs in Boston, New Jersey, and North Carolina, among others.[51]

Much of the evidence demonstrating the effectiveness of Oklahoma's universal preschool program comes from the evaluation of the program as implemented in Tulsa. Like Cincinnati, young children in Tulsa are more likely to live in families with income below the federal poverty line compared with the national pattern. As of 2014, about 33 percent of Tulsa's children younger than age six lived in families with income below poverty compared with the national average of 24 percent (see Table 4). In 2014, a family with one adult and two children was counted as poor if its cash income was below $19,073. For a family of two adults and two children, the poverty threshold stood at $24,008. About 59 percent of Tulsa's children lived in families with income below 185 percent of poverty, the income cutoff that determines eligibility for a free or reduced-price lunch. Nearly 80 percent live in families with income below 300 percent of poverty or about $57,000 for a family of three. Thus, in Tulsa, a universal preschool program could be classified as a targeted

Table 4. Distribution of Children Ages 0 to 5 in Tulsa, Oklahoma, and Cincinnati, Ohio, by Family Income Relative to Poverty: 2014

Population Group	Tulsa, Oklahoma		Cincinnati, Ohio	
	Percentage	Cumulative Percentage	Percentage	Cumulative Percentage
Children ages 0 to 5 by poverty status[a]				
Income below 100% of poverty	33.3	33.3	46.8	46.8
Income from 100% to 185% of poverty	25.2	58.5	18.2	65.0
Income from 185% to 300% of poverty	19.8	78.3	12.7	77.6
Income greater than 300% of poverty	21.7	100.0	22.4	100.0

SOURCE: U.S. Census Bureau, *American Fact Finder*, 2016.
[a] Poverty status is based on income thresholds set by the U.S. Census Bureau that vary by family size and the number of dependent children. In 2014, for example, the poverty threshold was $19,073 for a family of three with two children, and $24,008 for a family of four with two children.

program that is primarily reaching children in low- to lower-middle-income families.[52] Boston's public preschool program likewise is open to all children regardless of income, but it is effectively a targeted program as two-thirds of the participating children live in families with income below poverty. The same would hold for Cincinnati, where 47 percent of children live in families with income below poverty and about 78 percent live in families with income below 300 percent of poverty.

Intensity of Preschool Program

The intensity of the preschool program—whether the programs offers one year or two, and whether the program is delivered in a part-day session or a full-day session (typically six or six and a half hours) has potential implications for program impact, as well as program cost. Together, the total annual hours, or combined annual hours across multiple program years, determine the dosage the participating child will experience.

The preschool programs included in Table 1 are primarily one-year programs, or, for those programs offering two years, program effects have been measured only for the program serving four-year-olds (see Table A.1). Exceptions include Perry Preschool, Head Start, Chicago CPC, and the Abbott program in New Jersey. In each case, there is some evidence about the effects of two years versus one year. However, none has used a randomized design to test for differences in effects based on program duration. The general finding from those studies is that there are additional gains from a two-year program, but they do not always produce twice the impact as a one-year program.[53] A recent meta-analysis concluded that, compared with one-year programs, two-year programs produced an average effect size gain in cognitive skills at school entry of 0.10, but the difference was not statistically significant, in part because there were relatively few studies from which to make this comparison.[54] It may be possible to generate larger gains from an additional year of preschool by strengthening the curriculum and teacher practices to more effectively capitalize on the additional preschool exposure afforded by a two-year program.

Likewise, the preschool programs reviewed in this study include those offering only part- or full-day programs, although many offer both options. The effective programs in Tables 1 and 2 include examples of each, although there have been no experimental evaluations to rigorously test the differential effectiveness of part- versus full-day models. For programs that offer both options, without random assignment or another method to control for which children participate in which model, it is not possible to conclude whether there is differential effectiveness. For example, an initial analysis of the Tulsa program showed larger gains for children in the part-day program compared with the full-day program, but that may reflect the differential participation of children in these options based on race, family income, or other factors.[55] A subsequent evaluation of the Tulsa universal preschool program found that, within a given income group, effects on average test scores at kindergarten entry were higher for students in the full-day program compared with those in the part-day program.[56] However, for all but the reduced-price lunch group, effect sizes were less than twice as large for the full-day program compared with the part-day program. In other words, the results suggest that there may be added benefit from a full-day program compared with a part-day one, but the impact may not increase in step with the additional annual hours (i.e., doubling the annual hours may not double the impact). Again, this may reflect the limits of current practice in full-day programs, if programs are not currently implemented to use the available time efficiently in support of developmental activities. Even if full-day programs are not as efficient, there may be other reasons to make such programs available, especially for working families who need access to full-time care.

Alignment of Preschool with Early Elementary Grades

The earlier discussion regarding explanations for potential fadeout or convergence of cognitive skills across preschool program participants versus nonparticipants points to the importance of the structure and quality of elementary school

There is no indication that universal programs, such as those in Boston and Oklahoma, realize any smaller effects from serving children across the income spectrum.

programs. Research is just beginning to explore this issue. For example, using data from the Tennessee preschool evaluation, researchers find that there is small positive effect on some academic outcomes for children who attended preschool and subsequently received high-quality teaching in first grade.[57] Other research points to the importance of moving beyond basic content as children begin formal schooling. For example, using a large, nationally representative dataset, researchers examined whether children who received exposure to advanced academic content compared with basic content in kindergarten performed better on a skills assessment at the end of the school year.[58] Results indicate that children benefited—regardless of preschool experience—from being taught more-advanced content. Results such as these suggest that high-quality teaching and exposure to new academic content in early elementary school may help sustain the gains made in preschool.

This research is in line with a growing emphasis of the alignment of preschool programs with the early elementary grades through P–3 systems. As noted earlier, the evaluation of the Chicago CPC program demonstrated additional benefits from the extended supports incorporated from kindergarten to third grade. More generally, P–3 initiatives that have been implemented to date typically include increased access to preschool for three- and four-year-olds; support for the transition from preschool to kindergarten; alignment between learning standards, the curriculum, and assessment across the P–3 continuum; aligned professional development for teachers; instructional practices that focus on the whole child; parent engagement practices across the P–3 continuum; and use of data for accountability and quality improvement.[59] A recent quasi-experimental evaluation of a P–3 initiative with these features in Hawai'i showed a modest impact (effect size of 0.10) at the school level on third grade reading scores after five years of participation in the P–3 initiative.[60] Effective strategies for linking preschool investments with K–12 systems continues to be an active area of research.

The Key Lessons

The expanding investment in publicly funded preschool programs at the national, state, and local level has been accompanied by a growing contingent of rigorous evaluations that seek to understand the short- and longer-term impacts of the programs on participating children. A consistent finding across these studies is that high-quality preschool programs, operating at full scale, can produce meaningful gains in school readiness, especially for more-disadvantaged children. In addition,

> It may be possible to generate larger gains from an additional year of preschool by strengthening the curriculum and teacher practices to more effectively capitalize on the additional preschool exposure afforded by a two-year program.

where longer-term follow-up has been possible, several states and localities show evidence of sustained effects on academic achievement but also on other important aspects of educational performance such as grade retention, special education use, and high school graduation.

At the same time, this research evidence cautions against assuming that all preschool programs will, by their nature, generate favorable effects in the short or long term. Although ongoing research seeks to understand which program features are the most effective, the available evidence indicates that quality matters. Programs that have the largest initial impacts and demonstrate sustained effects into the early elementary grades on key aspects of education performance share several features in common: the quality of teacher-child interactions; the quality of instructional support for children; quality improvement supports through coaching and professional development; a systematic approach to monitoring and improving quality; and employing a proven curriculum and providing teachers with the training to implement it well.[61] Provided these features are in place, the research indicates that preschool programs can be effective under a range of alternative designs, including universal or targeted, part-day or full-day, and serving children for one or two years before they begin kindergarten. To further capitalize on preschool investments, there is also a growing recognition of the importance in strengthening the alignment between high-quality preschool programs and the quality of

the experience from kindergarten through third grade, although further research is needed to guide decisionmakers in how best to make these linkages.

EVIDENCE THAT PRESCHOOL PAYS OFF

With a growing demand for results-based accountability, policymakers in the public and private sectors interested in early childhood investments—including targeted or universal preschool programs—have increasingly turned to benefit-cost analyses to compare the economic cost of a given early childhood investment with the economic value of the outcomes generated by the investment. The resulting estimates of net present value benefits, the benefit-cost ratio, or the internal rate of return provide metrics of the return on investment to the program. Before presenting the findings from benefit-cost analyses of effective preschool programs, we first briefly consider the challenges of using this economic evaluation method in the context of early childhood programs, including one- or two-year preschool programs.

Benefit-Cost Analysis as a Tool for Measuring Economic Returns

The essential ingredients for conducting a benefit-cost analysis of a given program (or intervention) include (1) a well-specified program model with a clearly defined counterfactual condition (i.e., the alternative condition that the program group is compared with); (2) an evaluation of the program that provides causal evidence of its effects relative to the counterfactual; (3) estimates of the full economic cost required to implement the program relative to the counterfactual; and (4) estimates of the economic value of the outcomes affected by the program.[62] With these ingredients in place, a benefit-cost analysis proceeds by considering the year-by-year stream of program costs and the year-by-year economic value of the resulting program outcomes. After discounting the future dollar values of both costs and outcomes to account for the time value of money (i.e., that a dollar in the future is worth less than a dollar today), the analyst compares the cumulative present value costs with the cumulative present value outcomes as a measure of the net present value benefits of the program. When net benefits are positive, the program produces a positive economic return, as the value of the year-by-year impacts of the program outweigh the dollar value of the resources required to implement it, after discounting. Net present value benefits will be negative when the value of the program impacts is not sufficient to cover the program cost. Positive returns also mean that the benefit-cost ratio (ratio of net present value benefits to net present value costs) will exceed one. Negative returns mean the ratio is less than one. Another summary metric, the internal rate of return, reflects the discount rate where present value benefits equals present value costs.

There are several challenges with the application of benefit-cost analysis to preschool programs. Although the first two ingredients are often satisfied—as summarized in the prior section, there are numerous well-defined preschool models and many with rigorous evidence of program impact—many programs have not collected comprehensive measures of program cost. More importantly, many of the earliest impacts generated by high-quality preschool programs—gains in prereading skills, language and literacy skills, premathematics skills, and social and emotional development—are hard to express in monetary terms. When economic values or prices for program outcomes are not readily available, economists look to "shadow prices" as measures of economic value that do not derive from observed market prices but other sources of information about dollar values.

One solution to the challenge of valuing early impacts from preschool program participation is to extend the follow-up period in the evaluation so that potential outcomes of the program during the school-age years and beyond can be measured, outcomes that are more readily valued. However, such long-term follow-up takes time—requiring 13 years to measure the effects on high school graduation and waiting even longer to capture the effects on earnings and other adult outcomes. The Perry Preschool program offers one example of long-term

There are several challenges with the application of benefit-cost analysis to preschool programs.

follow-up (to age 40). The Chicago CPC program is another example (with follow-up to age 28). The benefit-cost analyses in those two examples are able to value a wider range of program effects in tallying up the benefits from the preschool program. However, programs with long-term follow-up were necessarily implemented in the past, when conditions were different compared with the present, such as the availability of early learning and care programs. Evaluations of current programs are potentially more indicative of the impacts that would be realized under today's conditions, but then only short-term impacts can be used in the benefit-cost analysis.

An alternative to long-term follow-up is to make a projection of the likely future impacts of a preschool program based on the observed outcomes. For example, if there is a known relationship between a test score in kindergarten and high school graduation, the early outcome can be linked to a later outcome that is more readily valued (e.g., estimating the difference in lifetime earnings between a high school graduate and a high school dropout). The benefit-cost analysis tool created by the Washington State Institute of Public Policy (WSIPP) includes many such linkages between early and later outcomes based on evidence from longitudinal research studies.[63] These linkages are used to generate estimates of the economic returns to preschool programs that do not have long-term follow-up.

When favorable outcomes from preschool participation are not readily valued in dollar terms, the benefit-cost analysis will provide an underestimate of the economic returns. Returns also will be underestimated if beneficial impacts are not measured as part of the evaluation, but they could be overstated if unmeasured unfavorable effects are not considered as well. For example, most evaluations of preschool focus on the outcomes of participating children, but they rarely capture favorable (or unfavorable) impacts on parents or other parties (e.g., siblings, peers). If a preschool program in providing a regular and stable source of care allows parents to work more and increase their productivity (e.g., by a reduction in lost work time), family income would increase and employers would gain from a reduction in absenteeism and perhaps reduced costs of turnover.[64] Most economic evaluations of preschool programs do include the value to parents of the subsidized early care and education services, but often omit some potential sources of economic value.

A final issue is that the application of benefit-cost analysis to such social programs as preschool has yet to converge on a common analytic approach. Thus, benefit-cost analyses may vary in the choice of discount rates, the specific shadow prices to use, and so on. Differences across studies in which outcomes are measured and in the length of the follow-up period introduce additional sources of variation that may affect the magnitude of the estimated return. This variation in methods often makes it difficult to compare results across studies. For this reason, studies may be useful for demonstrating that a particular program generates positive economic returns, but the research does not always support identifying programs with the highest rates of return because the benefit-cost studies used for comparison do not use a consistent methodology.[65]

Estimates of Economic Returns for Full-Scale Publicly Funded Preschool Programs

An estimated return to society for the Perry Preschool program of $17 for every dollar invested is perhaps the most widely cited estimate of the return on investment for preschool programs.[66] However, given the unique circumstances of the Perry Preschool program (e.g., small scale, setting with a "no preschool" comparison group), it is not realistic to expect social returns of that magnitude in real-world programs. Table 5 instead focuses on benefit-cost analysis results for full-scale programs based on the findings from their rigorous evaluations. These programs include a subset of those featured in Table 1, namely two targeted programs: Head Start and Chicago CPC; and one universal program, the Oklahoma universal program as implemented in Tulsa. A remaining row entry in the table is based on a meta-analysis of the findings of preschool program impacts. In each case, we summarize the program intensity, the population served, and the age at last follow-up. In particular, the Tulsa program estimates are specific to the one-year full-day program, whereas the other programs are a weighted combination of one- or two-year programs with part- or full-day options. The resulting estimates of present value costs, benefits, net benefits and the benefit-cost ratio are all from the societal perspective (i.e., full social return) and converted to 2015 dollars.

Excluding for the moment the result for Chicago CPC, the other row entries show benefit-cost ratios that range from about $2 to $4 for every dollar invested. Moreover, the estimates for Tulsa's universal program demonstrate that the favorable economic returns accrue for lowest-income children (those eligible for a free lunch, meaning income is below 130 percent of the federal poverty level), as well as children somewhat higher up on the income scale, those eligible for a reduced price lunch (income from 130 percent to 185 percent of the federal poverty level) and

Table 5. Benefit-Cost Analysis Results for Full-Scale Publicly Funded Preschool Programs

Program	Program Intensity / Population Served / Follow-up Period	Per Child (in 2015 $)			
		PDV Costs to Society	PDV Benefits to Society	NPV Benefits to Society	Benefit-Cost Ratio for Society
Head Start (meta-analysis)	One or two years, part or full day Targeted (children in families with income below federal poverty level) Varied follow-up ages	8,841	23,178	14,337	2.63
State and district preschool programs for low-income 3- and 4-year-olds (meta-analysis)	Varied intensity Targeted (low-income children) Varied follow-up ages	7,199	30,154	22,955	4.20
Chicago Child-Parent Centers (CPC)	One or two years, part day Targeted (children in high-poverty neighborhoods) Followed to age 26	9,730	105,419	95,688	10.83
Tulsa, Oklahoma	One year, full day Universal: free lunch eligible subgroup Followed to K entry	10,687	33,023	22,336	3.09
Tulsa, Oklahoma	One year, full day Universal: reduced-price lunch eligible subgroup Followed to kindergarten entry	10,687	36,870	26,183	3.45
Tulsa, Oklahoma	One year, full day Universal: full-price lunch subgroup Followed to kindergarten entry	10,687	30,137	19,450	2.82

SOURCES: For Head Start and state and district programs: Washington State Institute for Public Policy, *Early Childhood Education for Low-Income Students: A Review of the Evidence and Benefit-Cost Analysis*, Olympia, Wash., 2014; for Chicago CPC: Arthur J. Reynolds, Judy A. Temple, Barry A. White, Suh-Ruu Ou, and Dylan L. Robertson, "Age-26 Cost-Benefit Analysis of the Child-Parent Center Early Education Program," *Child Development*, Vol. 82, No. 1, 2011, pp. 379-404; for Tulsa: Timothy J. Bartik, William Gormley, and Shirley Adelstein, "Earnings Benefits of Tulsa's Pre-K Program for Different Income Groups," *Economics of Education Review*, Vol. 31, 2012, pp. 1143–1161.
NOTES: All dollar values were converted to 2015 dollars using the Consumer Price Index for All Urban Consumers. The benefit-cost ratios are the ratio of the present discounted value of total benefits to society as a whole (participants and the rest of society) divided by present discounted value of program costs. The discount rate is 3 percent and discounting is to ages three or four.

those in the full-price lunch group who are not eligible for any subsidy.[67]

The benefit-cost analysis of the Chicago CPC program is the most comprehensive because the follow-up to age 26 allowed estimates of the impacts of participating in the program on a range of outcomes that are economically important. Notably, the estimated return to society of almost $11 for every dollar invested in this targeted preschool program includes economic benefits from the following:

- net savings to K–12 education from lower grade retention, lower rates of special education use, and higher high school graduation rates
- increased lifetime earnings and tax payments to federal, state, and local governments
- reduced cost to the criminal justice system from reductions in crime, as well as reduced victim costs from crime reduction
- reduced child abuse and neglect
- improved health and health behaviors (e.g., reduced depression, smoking, and substance use).[68]

The estimated returns for the other programs in Table 5 with only short-term follow-up could be as high as Chicago CPC, if there are long-term impacts of a similar magnitude for these same outcomes.

Forecasts of Economic Returns for Proposed Programs

In addition to estimating the economic returns to existing programs, analysts have forecasted the expected returns to new or expanded preschool programs under consideration. These include a RAND study, which estimated the economic returns to a universal part-day one-year preschool program for four-year-olds in California.[69] The benefit-cost analysis projected a return to society of approximately $2 to $4 for every dollar spent on the high-quality program, depending on assumptions about the gains to children already in preschool but in a lower-quality program and the gains to new preschool participants. As another example, a series of studies provides estimates of the returns to universal one- or two-year part-day preschool programs in Arkansas, Massachusetts, Ohio, and Wisconsin.[70] These analyses focused on returns to the public sector (e.g, state and local governments). Because of similar assumptions, the resulting estimates showed returns ranging from about $1.20 to about $1.90 for every dollar invested. These estimates exclude the private benefits to the preschool participants themselves (e.g., higher earnings over their lifetime), so the returns to society as a whole (both private gains and benefits to the public sector) would be more in line with the estimated range of returns for a universal preschool program in California.

The Bottom Line

Although numerous full-scale preschool programs have been evaluated using rigorous methods (see Table 1), only a handful of them have been the focus of a formal benefit-cost analysis (see Table 5). This stems from the absence of cost data in some cases and the challenges of valuing the outcomes from preschool programs when evaluation evidence is limited to short-term effects on outcomes that are not readily valued in dollars. The few preschool programs with benefit-cost findings provide proof of the principle that preschool programs can be an investment that pays off, both in terms of the direct effects on meaningful outcomes for children but also in terms of dollars and cents. Such favorable returns are demonstrated for part- and full-day programs, for one- and two-year programs, and for both targeted and universal programs.

It is important to keep in mind that these findings apply to high-quality program models implemented with attention to high standards. In the same way that the miles per gallon of gasoline for a given automobile will differ from the Environmental Protection Agency rating depending on conditions of the road, the maintenance of the vehicle, and the driving speed, among other factors, so too will the exact return on investment for a given preschool program vary from the results shown in Table 5 depending on the preschool program model, how well it is implemented, the characteristics of the population served, and the local context within which it is delivered. Despite the expectation that "your mileage may vary," benefit-cost analyses indicate that high-quality preschool programs, when implemented well, can produce impacts of sufficient magnitude as to generate economic benefits that readily outweigh the upfront preschool program investment.

INFORMING INVESTMENTS

As stakeholders in Cincinnati, Ohio, consider options for expanding preschool access and quality, there is a rich research base that can provide critical guidance for designing a citywide preschool initiative and anticipating the potential return on the planned investment. In this final section, we highlight the key findings from that evidence base reviewed in this report.

There are numerous examples of full-scale preschool programs with rigorous evaluations that show improvements in school readiness for participating children. These examples come from high-quality programs implemented at full scale in various states and communities. The size of the effects from these real-world programs are typically smaller than those found in

The estimates for Tulsa's universal program demonstrate that the favorable economic returns accrue for lowest-income children, as well as children somewhat higher up on the income scale.

small-scale demonstration programs (such as Perry Preschool), but they represent meaningful gains in children's readiness for school.

Favorable impacts have been demonstrated for part- and full-day preschool programs, as well as one- and two-year programs, but the research is not definitive about the comparative effectiveness of these options. Preschool programs proven to be effective include those operating with either a part- or full-day schedule. Likewise, favorable impacts have been demonstrated for programs that begin with either three- or four-year-olds. The available evaluation evidence does not support definitive conclusions about the additional gains from programs with more hours per day or from programs offering two years of preschool instead of one. The research suggests that children experience additional benefit from a program with more hours or from a second year of attendance, but the additional gains may not be proportional to the increase in preschool dosage. This may be because existing programs are not structured to fully capitalize on the added time.

High quality is a common element among the preschool programs with the largest effects on school readiness and with sustained effects at older ages. These effective programs include such features as well-trained classroom teachers who are provided with ongoing professional development supports through coaching and other mechanisms, a learning environment that supports teachers and children, a well-defined curriculum that is implemented with fidelity in the classroom and aligned with the early elementary grades, and ongoing monitoring of program quality and other metrics that support continuous quality improvement.

Children across the income spectrum may benefit from high-quality preschool but the impacts tend to be larger for more disadvantaged children. Because of funding constraints, most large-scale publicly funded preschool programs serve children in low-income families or who face other risks to healthy development. Where programs have been made universally available such as Oklahoma's universal preschool program, beneficial effects have been found for children across the income spectrum, albeit the effects are largest for the most-disadvantaged children. But like the universal program evaluated in Tulsa, Oklahoma, a universal program in Cincinnati would effectively be a targeted program: Half of the city's children younger than age six live in families with income below the federal poverty level (well above the state average of 28 percent), and almost 80 percent of children younger than age six have income less than 300 percent of poverty.

Although differences in achievement scores between preschool program participants and nonparticipants tend to narrow as they advance through the elementary grades, high-quality preschool programs show sustained effects on other aspects of school performance. Rigorous evaluations of a number of high-quality full-scale preschool programs such as those in Chicago, Michigan, New Jersey, North Carolina, Oklahoma, and Washington have shown sustained favorable effects through at least third grade on student achievement in reading and/or mathematics. Even when evaluations find that those who did not attend preschool eventually catch up to their preschool counterparts in terms of academic achievement measures, the evaluations often find that preschool participants have experienced favorable effects for other aspects of educational performance, such as special education use, grade retention, and high school completion.

Improving the alignment between preschool and the early elementary grades may help sustain the initial boost in cognitive and noncognitive skills from preschool participation. Although research is ongoing to identify the factors that may contribute to the fadeout or catch-up phenomenon, a well-aligned preschool to elementary school system offers a promising strategy for ensuring that children who experience a high-quality preschool program can continue to build upon their early success.

High-quality preschool programs represent a significant investment of resources, but that investment may be paid back through improved outcomes during the school-age years and beyond. Estimates of the economic returns to full-scale high-quality preschool programs range from about $2.50 per dollar invested to $4.20 per dollar invested. The results for Chicago CPC—the one full-scale program with long-term follow-up—suggest that the economic returns may be even higher once longer-term impacts can be observed and valued. It is important to keep in mind that the actual return on investment experienced for any given publicly funded preschool program will depend on the population of children served, the quality of the preschool program implemented, and the impacts of the program relative to the status quo.

APPENDIX A Table A.1. Preschool Programs with Rigorous Evaluations

Table A.1 provides additional information about the preschool programs with rigorous evaluations reviewed in this report. Table A.2. provides information on the evaluation methods and citations to the evaluation studies.

Program (Years in Operation)	Scale of Operation	Targeted or Universal	Ages Served (Evaluated)	Intensity	Teacher–to–Child Ratio	Maximum Class Size	Teacher Education	Per Child Spending per Year[a] (2015 US$)
a. Demonstration Program								
Perry Preschool (1962–67)	Single-site program	Targeted, very disadvantaged minority children	3, 4 (3, 4)	Part day, school year	1:6	13	Required bachelor's degree with training in ECE	$9,020
b. National Program								
Head Start (1965–present)	National	Targeted, low income (< 100% FPL)	3, 4 (3, 4)	Part day and full day, school year	1:10	20	Require associate's degree or higher[b]	$9,270
c. State Programs								
Arkansas Better Chance for School Success (2004–present)	Statewide	Targeted, low-income (< 200% FPL) or other risk factors	3, 4 (4)	Full day (7 hours), school year	1:10	20	Require associate's degree or higher with training in ECE	$5,550[d]
Georgia Universal PreK Program (1995–present)	Statewide	Universal	4 (4)	Full day (6.5 hours), school year	1:11	22	Require bachelor's degree with training in ECE	$3,750[d]
Michigan Great Start Readiness Program (1985–present)	Statewide	Targeted, low-income (< 300% FPL) or other risk factors	4 (4)	Part day (3 hours) and full day (6.5 hours), school year	1:8	18	Require bachelor's degree with training in ECE	$5,710[d]
New Jersey Abbott Preschool Program (1998–present)	Statewide	Targeted, all students in high-poverty school districts	3, 4 (3, 4)	Full day (6 hours), school year	2:15	15	Require bachelor's degree with training in ECE	$13,350[c]
New Mexico PreK Program (2005–present)	Statewide	Targeted, students in attendance zones of Title I elementary schools	4 (4)	Part day (2.5 to 3.5 hours), school year	1:10	20	Require bachelor's degree with training in ECE for public school programs only	$3,560[d]
North Carolina Pre-K Program North (2001–present)	Statewide	Targeted, low income (< 75% SMI) or other risk factors	4 (4)	Full day (6.5 hours), school year	1:9	18	Require bachelor's degree with training in ECE	$7,360[d]

Table A.1—Continued

Oklahoma Early Childhood Four-Year-Old Program (1990–present)	Statewide	Universal	4 (4)	Part day (2.5 hours) and full day (6.5 hours), school year	1:10	20	Require bachelor's degree with training in ECE	$7,690[d]
South Carolina Half-Day Child Development Program (4K) (1984–present)	Statewide	Targeted, low income (< 185% FPL) or other risk factors	4 (4)	Part day (2.5 to 5 hours), school year	1:10	20	Require bachelor's degree with training in ECE for public school programs only	$4,300[d]
Tennessee Voluntary Prekindergarten (TVP) Program (2004–present)	Statewide	Targeted, low income (< 185% FPL) or other risk factors	4 (4)	Full day (5.5 hours), school year	1:10	20	Require bachelor's degree with training in ECE	$5,900[d]
Washington Early Childhood Education and Assistance Program (ECEAP) (1985–present)	Statewide	Targeted, low income (< 110% FPL) or other risk factors	3, 4 (4)	Local decision (minimum 320 hours per year), school year	1:10	20	Require associate's degree or higher with training in ECE	$6,670[d]
West Virginia Universal Pre-K System (1983–present)	Statewide	Universal	4 (4)	Local decision (minimum 14 hours per week), school year	1:10	20	Require bachelor's degree with training in ECE	$8,810[d]
d. District Programs								
Boston Public School PreK (1998–present)	Boston public schools	Universal	3, 4 (4)	Full day (6 hours), year-round	1:11	22	Require bachelor's degree for public school teachers; master's within five years	$12,390
Chicago Child-Parent Centers (CPC) Program (1967–present)	Chicago public schools	Targeted, children in disadvantaged center-city neighborhoods	3, 4 (3, 4)	Part day (3 hours), school year	2:17	17	Require bachelor's degree with training in ECE	$6,400
Tulsa Oklahoma Early Childhood Four-Year-Old Program (1990–present)	Tulsa public schools (part of statewide program)	Universal	4 (4)	Part day (2.5 hours) and school day (6.5 hours), school year	1:10	20	Require bachelor's degree with training in ECE	$7,690[d]

SOURCE: W. Steven Barnett, Megan E. Carolan, James H. Squires, Kristy Clarke Brown, and Michelle Horowitz, *The State of Preschool 2014: State Preschool Yearbook*, New Brunswick, N.J.: National Institute for Early Education Research, 2015 and evaluation studies listed in Appendix B.

NOTES: Program descriptions are for 2013–14 program year. Head Start requirements are for center-based programs serving primarily four-year-olds. ECE = early childhood education; FPL = federal poverty level.

[a] Inflated to 2015 dollars using the Consumer Price Index, All Urban Consumers. Costs for these two-year programs are shown on a per-year basis.

[b] Nationwide, Head Start programs are required to have at least half of their teachers with a bachelor's degree or higher.

[c] Spending for the six-hour-per-day, 185-day school year. Additional costs apply for wrap-around services for full-day, year-round care.

[d] Spending per child may exclude local contributions.

Table A.2. Citations and Evaluation Features for Studies Cited in Tables 1 and 2

Program	Evaluation Citations	Evaluation Method	Preschool Cohort(s) Studied	Last Follow-up
a. Demonstration Program				
Perry Preschool	Schweinhart et al. (2005)	RCT	1962, 1963, 1964	Age 40
b. National Program				
Head Start	Ludwig and Phillips (2007); U.S. DHHS (2005, 2012)	RCT	2002	Grade 3
c. State Programs				
Arkansas Better Chance for School Success	Hustedt et al. (2007)	QE RD	2005	K entry
Georgia Universal PreK Program	Peisner-Feinberg et al. (2014)	QE RD	2011	K entry
Michigan Great Start Readiness Program	Wong et al. (2008); Malofeeva, Daniel-Echol, and Xiang (2007); Schweinhart et al. (2012); Xiang and Schweinhart (2002)	QE RD QE	2004 1995	K entry, Grade 12
New Jersey Abbott Preschool Program	Wong et al. (2008); Frede et al. (2007); Barnett et al. (2013)	QE RD QE	2004 2005	K entry, Grade 5
New Mexico PreK Program	Hustedt et al. (2007, 2008, 2009)	QE RD	2005, 2006, 2007	K entry
North Carolina Pre-K Program	Ladd, Muschkin, and Dodge (2014); Peisner-Feinberg et al. (2010)	QE	2002, 2003	Grade 3
Oklahoma Early Childhood Four-Year-Old Program	Wong et al. (2008)	QE RD	2004	K entry
South Carolina Half-Day Child Development Program (4K)	Wong et al. (2008)	QE RD	2004	K entry
Tennessee Voluntary Prekindergarten (TVP) Program	Lipsey et al. (2011, 2013); Lipsey, Farran, and Hofer (2015)	RCT and QE	2009, 2010	Grade 3
Washington Early Childhood Education and Assistance Program	Bania et al. (2015)	QE	2003–2008	Grade 5
West Virginia Universal Pre-K System	Wong et al. (2008)	QE RD	2004	K entry
d. District Programs				
Boston Public School PreK	Weiland and Yoshikawa (2013)	QE RD	2008	K entry
Chicago Child-Parent Centers (CPC) Program	Reynolds and Temple (1995); Reynolds (1998); Reynolds et al. (2011)	QE QE	1983–1985 1983–1985	K entry Age 28
Tulsa Oklahoma Early Childhood Four-Year-Old Program	Gormley et al. (2005); Gormley et al. (2011); Hill, Gormley, and Adelstein (2015)	QE RD QE QE	2002 2005 2000, 2005	K entry K entry Grade 3

SOURCES: Cited in table's second column. See Appendix B for complete references.
NOTES: QE = quasi-experimental; RCT = randomized control trial; RD = regression discontinuity.

APPENDIX B. EVALUATION STUDIES

This appendix provides the references to the evaluation studies cited in Table A.2. The studies are listed by program following the order of Table A.2.

Perry Preschool

Schweinhart, Lawrence J., Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores, *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40 (Monographs of the High/Scope Educational Research Foundation, No. 14)*, Ypsilanti, Mich.: High/Scope Press, 2005.

Head Start

Ludwig, Jens and Deborah Phillips, "The Benefits and Costs of Head Start," *Social Policy Report*, Vol. XXI, No. 3, 2007, pp. 3–18.

U.S. Department of Health and Human Services, *Head Start Impact Study: First Year Findings*, Washington, D.C.: Administration for Children and Families, U.S. Department of Health and Human Services, June 2005.

U.S. Department of Health and Human Services, *Head Start Impact Study: Final Report*, Washington, D.C.: Administration for Children and Families, U.S. Department of Health and Human Services, January 2010.

U.S. Department of Health and Human Services, *Third Grade Follow-Up to the Head Start Impact Study*, Washington, D.C.: Administration for Children and Families, U.S. Department of Health and Human Services, October 2012.

Arkansas Better Chance for School Success

Hustedt, Jason T., W. Steven Barnett, Kwanghee Jung, and Jessica Thomas, *The Effects of the Arkansas Better Chance Program on Young Children's School Readiness*, New Brunswick, N.J.: National Institute for Early Education Research, January 2007.

Georgia Universal PreK Program

Peisner-Feinberg, Ellen S., Jennifer M. Schaaf, Doré R. LaForett, Lisa M. Hildebrandt, and John Sideris, *Effects of Georgia's Pre-K Program on Children's School Readiness Skills: Findings from the 2012–2013 Evaluation Study*, Chapel Hill: The University of North Carolina, FPG Child Development Institute, 2014.

Michigan Great Start Readiness Program

Malofeeva, Elena, Marijata Daniel-Echol, and Zongping Xiang, *Findings from the Michigan School Readiness Program 6 to 8 Follow-Up Study*, Ypsilanti, Mich.: High Scope Educational Research Foundation, 2007.

Schweinhart, Lawrence J., Zongping Xiang, Marijata Daniel-Echols, Kimberly Browning, and Tomoko Wakabayashi, *Michigan Great Start Readiness Program Evaluation 2012: High School Graduation and Grade Retention Findings*, Ypsilanti, Mich.: High Scope Educational Research Foundation, 2012.

Wong, Vivian C., Thomas D. Cook, W. Steven Barnett, and Kwanghee Jung, "An Effectiveness-Based Evaluation of Five State Pre-Kindergarten Programs," *Journal of Policy Analysis and Management*, Vol. 27, No. 1, 2008, pp. 122–154.

Xiang, Zongping, and Lawrence J. Schweinhart, *Effects Five Years Later: The Michigan School Readiness Program Evaluation Through Age 10*, Ypsilanti, Mich.: High Scope Educational Research Foundation, 2002.

New Jersey Abbott Preschool Program

Barnett, W. Steven, Kwanghee Jung, Min-Jong Young, and Ellen C. Frede, *The Abbott Preschool Program Longitudinal Effects Study: Fifth Grade Follow-Up*, New Brunswick, N.J.: National Institute for Early Education Research, March 2013.

Frede, Ellen, Kwanghee Jung, W. Steven Barnett, Cynthia Esposito Lamy, and Alexandra Figueras, *The Abbott Preschool Program Longitudinal Effects Study (APPLES), Interim Report*, New Brunswick, N.J.: National Institute for Early Education Research, June 2007.

Wong, Vivian C., Thomas D. Cook, W. Steven Barnett, and Kwanghee Jung, "An Effectiveness-Based Evaluation of Five State Pre-Kindergarten Programs," *Journal of Policy Analysis and Management*, Vol. 27, No. 1, 2008, pp. 122–154.

New Mexico PreK Program

Hustedt, Jason T., W. Steven Barnett, and Kwanghee Jung, *The Effects of the New Mexico PreK Initiative on Young Children's School Readiness*, New Brunswick, N.J.: National Institute for Early Education Research, August 2007.

Hustedt, Jason T., W. Steven Barnett, Kwanghee Jung, and Alexandra Figueras, *Impacts of New Mexico PreK on Children's School Readiness at Kindergarten Entry: Results from the Second Year of a Growing Initiative*, New Brunswick, N.J.: National Institute for Early Education Research, June 2008.

Hustedt, Jason T., W. Steven Barnett, Kwanghee Jung, and Alexandra Figueras-Daniel, *Impacts of New Mexico PreK on Children's School Readiness at Kindergarten Entry: Results from the Third Year of Implementation*, New Brunswick, N.J.: National Institute for Early Education Research, September 2009.

North Carolina Pre-K Program

Ladd, Helen F., Clara G. Muschkin, and Kenneth A. Dodge, "From Birth to School: Early Childhood Initiatives and Third-Grade Outcomes in North Carolina," *Journal of Policy Analysis and Management*, Vol. 33, No. 1, 2014, pp. 162–187.

Peisner-Feinberg, Ellen S., and Jennifer M. Schaaf, *Long-term Effects of the North Carolina More a Four Pre-Kindergarten Program: Children's Reading and Math Skills at Third Grade*, Chapel Hill: Frank Porter Graham Child Development Institute, University of North Carolina, 2010.

Oklahoma Early Childhood Four-Year-Old Program

Wong, Vivian C., Thomas D. Cook, W. Steven Barnett, and Kwanghee Jung, "An Effectiveness-Based Evaluation of Five State Pre-Kindergarten Programs," *Journal of Policy Analysis and Management*, Vol. 27, No. 1, 2008, pp. 122–154.

South Carolina Half-Day Child Development Program

Wong, Vivian C., Thomas D. Cook, W. Steven Barnett, and Kwanghee Jung, "An Effectiveness-Based Evaluation of Five State Pre-Kindergarten Programs," *Journal of Policy Analysis and Management*, Vol. 27, No. 1, 2008, pp. 122–154.

Tennessee Voluntary Prekindergarten Program

Lipsey, Mark W., Dale C. Farran, Carol Bilbrey, Kerry G. Hofer, and Nianbo Dong, *Initial Results of the Evaluation of the Tennessee Voluntary Pre-K Program*, Nashville, Tenn.: Peabody Research Institute, Vanderbilt University, April 2011.

Lipsey, Mark W., Dale C. Farran, and Kerry G. Hofer, *A Randomized Control Trial of a Statewide Voluntary Prekindergarten Program on Children's Skills and Behaviors Through Third Grade*, Nashville, Tenn.: Peabody Research Institute, Vanderbilt University, September 2015.

Lipsey, Mark W., Kerry G. Hofer, Nianbo Dong, Dale C. Farran, and Carol Bilbrey, *Evaluation of the Tennessee Voluntary Prekindergarten Program: Kindergarten and First Grade Follow-Up Results from the Randomized Control Design*, Nashville, Tenn.: Peabody Research Institute, Vanderbilt University, August 2013.

Washington Early Childhood Education and Assistance Program

Bania, N., N. Kay, S. Aos, and A. Pennucci, *Outcome Evaluation of Washington State's Early Childhood Education and Assistance Program*, Document No. 14-12-2201, Olympia: Washington State Institute for Public Policy, 2015.

West Virginia Universal PreK System

Wong, Vivian C., Thomas D. Cook, W. Steven Barnett, and Kwanghee Jung, "An Effectiveness-Based Evaluation of Five State Pre-Kindergarten Programs," *Journal of Policy Analysis and Management*, Vol. 27, No. 1, 2008, pp. 122–154.

Boston Public School PreK

Weiland, Christina, and Hirokazu Yoshikawa, "Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills," *Child Development*, Vol. 84, No. 6, 2013, pp. 2112–2130.

Chicago Child-Parent Centers Program

Reynolds, Arthur J., "The Chicago Child-Parent Center and Expansion Program: A Study of Extended Early Childhood Intervention," in Jonathan Crane, ed., *Social Programs that Work*, New York: Russell Sage Foundation, 1998, pp. 110–147.

Reynolds, Arthur J., and Judy A. Temple, "Quasi-Experimental Estimates of the Effects of a Preschool Intervention," *Evaluation Review,* Vol. 19, No. 4, 1995, pp. 347–373.

Reynolds, Arthur J., Judy A. Temple, Suh-Ruu Ou, Irma A. Arteaga, and Barry A. B. White, "School-Based Early Childhood Education and Age-28 Well-Being: Effects by Timing, Dosage, and Subgroups," *Science*, Vol. 333, July 15, 2011, pp. 360–364.

Tulsa Oklahoma Early Childhood Four-Year-Old Program

Gormley, William T., Ted Gayer, Deborah Phillips, and Brittany Dawson, "The Effects of Universal Pre-K on Cognitive Development," *Developmental Psychology*, Vol. 41, No. 6, 2005, pp. 872–884.

Gormley, William T., Deborah A. Phillips, Katie Newmark, Kate Welti, and Shirley Adelstein, "Social-Emotional Effects of Early Childhood Education Programs in Tulsa," *Child Development*, Vol. 82, No. 6, 2011, pp. 2095–2109.

Hill, Carolyn J., William T. Gormley Jr., and Shirley Adelstein, "Do the Short-Term Effects of a High-Quality Preschool Program Persist?" *Early Childhood Research Quarterly*, Vol. 32, 2015, pp. 60–79.

NOTES

[1] According to the 2014 American Community Survey, 46.8 percent of children in Cincinnati younger than age six live in families with income below the federal poverty level. Another 30.9 percent have income from 100 percent to 300 percent of the federal poverty level. U.S. Bureau of Census, *American Fact Finder*, web page, January 2016.

[2] Sean F. Reardon, "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations," in Greg J. Duncan and Richard Murnane, eds. *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, New York: Russell Sage Foundation, pp. 91–116; and Emma Garcia, *Inequalities at the Starting Gate: Cognitive and Noncognitive Skills Gaps Between 2010–2011 Kindergarten Classmates*, Washington, D.C.: Economic Policy Institute, 2015.

[3] Flavio Cunha and James J. Heckman, "The Technology of Skill Formation," *American Economic Review*, Vol. 97, No. 2, 2007, pp. 31–47.

[4] For example, although it is often counted as a preschool program, we did not include the Abecedarian program in our review because it begins serving children soon after birth and continues until kindergarten entry. For a review of the Abecedarian program and other early childhood interventions, see Lynn A. Karoly, M. Rebecca Kilburn, and Jill Cannon, *Early Childhood Interventions: Proven Results, Future Promise*, Santa Monica, Calif.: RAND Corporation, MG-341, 2005.

[5] For other recent syntheses that overlap with (but are not identical to) our focus, see Hirokazu Yoshikawa, Christina Weiland, Jeanne Brooks-Gunn, Margaret R. Burchinal, Linda M. Espinosa, William T. Gormley, Jens Ludwig, Katherine A. Magnuson, Deborah Phillips, and Martha J. Zaslow, *Investing in Our Future: The Evidence Base on Preschool Education*, New York: Foundation for Child Development and Ann Arbor, Mich.: Society for Research in Child Development, 2013; and Sneha Elango, James J. Heckman, Jorge Luis Garcia, and Andréd Hojman, "Early Childhood Education," Human Capital and Economic Opportunity Global Working Group Working Paper No. 2015-017, Chicago: University of Chicago, 2015.

[6] Lawrence J. Schweinhart, Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores, *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40 (Monographs of the High/Scope Educational Research Foundation, No. 14)*, Ypsilanti, Mich.: High/Scope Press, 2005.

[7] The Tulsa district program is part of the Oklahoma statewide universal preschool program, but the Tulsa program has been the subject of more-intensive evaluation in the past decade.

[8] Gary Burtless, "The Case for Randomized Field Trials in Economic and Policy Research," *Journal of Economic Perspectives*, Vol. 9, 1995, pp. 63–84.

[9] Thomas D. Cook and Donald T. Campbell, *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Boston: Houghton Mifflin, 1979.

[10] In the regression discontinuity studies we cite, the child assessments are conducted at the same time for both the treatment group (children who attended preschool in the prior year because they made the age cutoff and are now entering kindergarten) and the control group (children who are just entering preschool because they did not meet the age cutoff the prior year). The models estimate the relationship between child assessments and age (controlling for other factors), separately for the treatment and control groups. The treatment effect is measured as the difference between the expected assessment score of a treated child who just made the age cutoff and the expected assessment score of a control child who just missed the age cutoff. For more detail on this approach and some of the limitations, see Mark W. Lipsey, Christina Weiland, Hirokazu Yoshikawa, Sandra Jo Wilson, and Kerry G. Hofer, "The Prekindergarten Age-Cutoff Regression-Discontinuity Design: Methodological Issues and Implications for Application," *Educational Evaluation and Policy Analysis*, Vol. 37, No. 3, 2015, pp. 296–313.

[11] At issue is whether any given quasi-experimental method accounts for the potential bias from failing to control for factors that are correlated with preschool program participation and that have a causal link to the observed outcomes. When all potential confounding factors are not controlled for, participation in preschool will be credited with causing the outcomes instead of the unmeasured factors. The nature of the bias in estimating the impact of preschool participation depends upon the omitted factors, but typically the expectation is that failing to control for the unobserved factors would tend to make estimates of the preschool program's impacts too large.

[12] U.S. Department of Health and Human Services, *Head Start Impact Study: First Year Findings*, Washington, D.C.: Administration for Children and Families, U.S. Department of Health and Human Services, June 2005.

[13] Duncan and Magnuson make this point in their review article of preschool program impacts. They also argue that home environments in low-income families also have likely improved over time with the increase in maternal education. See Greg J. Duncan and Katherine Magnuson, "Investing in Preschool Programs," *Journal of Economic Perspectives*, Vol. 27, No. 2, 2013, pp. 109–132.

[14] Daphna Bassok, Jenna E. Finch, RaeHyuck Lee, Sean F. Reardon, and Jane Waldfogel "Socioeconomic Gaps in Early Childhood Experiences, 1998 to 2010," working paper, University of Virginia, 2015.

[15] We note that the evaluation of the Tennessee's state preschool program is frequently described as using an experimental design. While the study has a randomized component, the main findings pertaining to children's cognitive development at the end of preschool and subsequent early elementary grades are based on data collected for a subset of the randomized treatment and control groups where the parents gave consent for direct developmental assessments. This substudy approach negates the experimental design and requires the use of quasi-experimental methods to account for the selectivity of which parents consented to the developmental assessments. Hence, the evaluation has relied on propensity score matching as the primary analytic method for comparing preschool participants and nonparticipants. For additional detail, see Mark W. Lipsey, Dale C. Farran, and Kerry G. Hofer, *A Randomized Control Trial of a Statewide Voluntary Prekindergarten Program on Children's Skills and Behaviors Through Third Grade*, Nashville, Tenn.: Peabody Research Institute, Vanderbilt University, 2015.

[16] The print awareness measure assesses whether children are able to recognize letters, can match sounds to letters, and can distinguish between symbols and letters.

[17] The effect size is usually calculated as the ratio of the program effect (the difference in means between the treatment and control groups in the case of an experimental evaluation) divided by the standard deviation of that effect estimate (i.e., the pooled standard deviation in a difference of means). In some cases, effect sizes are calculated using the standard deviation of the control group mean in the denominator.

[18] Jacob Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Hillsdale, N.J.: Lawrence Erlbaum Associates, 1988; and Jonathan Crane, *Social Programs that Work*, New York: Russell Sage Foundation, 1998.

[19] For example, to allow for attenuation, the Washington State Institute for Public Policy benefit-cost analysis model applies a discount up to 50 percent to any impact parameter from a non-real-world study. Washington State Institute for Public Policy, *Benefit-Cost Technical Documentation*, Olympia, Wash., 2015.

[20] Duncan and Magnuson, 2013. The authors report a meta-analysis average effect size of 0.35. A more conservative estimate is an effect size of 0.21, which is found when program effect sizes are weighted to account for the precision (i.e., standard error) of each estimate.

[21] Washington State Institute for Public Policy, *Early Childhood Education for Low-Income Students: A Review of the Evidence and Benefit-Cost Analysis*, Olympia, Wash., 2014. See also Gregory, Sadako Vargas, Sharon Ryan, and W. Steven Barnett, "Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development," *The Teachers College Record*, Vol. 112, 2010, pp. 579–620.

[22] Christina Weiland and Hirokazu Yoshikawa, "Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills," *Child Development*, Vol. 84, No. 6, 2013, pp. 2112–2130.

[23] William T. Gormley, Deborah A. Phillips, Katie Newmark, Kate Welti, and Shirley Adelstein, "Social-Emotional Effects of Early Childhood Education Programs in Tulsa," *Child Development*, Vol. 82, No. 6, 2011, pp. 2095–2109.

[24] Ellen S. Peisner-Feinberg, Jennifer M. Schaaf, Doré R. LaForett, Lisa M. Hildebrandt, and John Sideris, *Effects of Georgia's Pre-K Program on Children's School Readiness Skills: Findings from the 2012–2013 Evaluation Study*, Chapel Hill: University of North Carolina, FPG Child Development Institute, 2014.

[25] Pamela Morris, Shira Mattera, Nina Castells, Michael Bangser, Karen Bierman, and Cybele Raver, *Impact Findings from the Head Start CARES Demonstration: National Evaluation of Three Approaches to Improving Preschoolers' Social and Emotional Competence (OPRE Report 2014-44)*, Washington, D.C.: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2014.

[26] Schweinhart et al., 2005.

[27] Table 2 features the statistically significant sustained effects of the listed programs. It is important to note that these programs do not find significant effects on all outcomes measured. For example, relative to the comparison group as of the follow-up to age 28, Chicago CPC participants did not show a significant difference in food stamp participation or in the percentage completing a two- or four-year college degree. Additionally, in the Tulsa preschool program, which used a matching technique to evaluate third-grade outcomes, no significant differences were found for any academic outcomes measured at third grade for one of the two cohorts of children studied.

[28] Research for North Carolina has exploited the rollout of the program across counties and other time and found significant effects on student achievement test scores for every $1,000 increase in spending, along with the effect shown in Table 2 for special education use. Helen F. Ladd, Clara G. Muschkin, and Kenneth A. Dodge, "From Birth to School: Early Childhood Initiatives and Third-Grade Outcomes in North Carolina," *Journal of Policy Analysis and Management*, Vol. 33, No. 1, 2014, pp. 162–187.

[29] Duncan and Magnuson, 2013; Gregory Camilli, Sadako Vargas, Sharon Ryan, and W. Steven Barnett, "Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development," *Teachers College Record*, Vol. 112, pp. 579–620; Washington State Institute for Public Policy, 2014; and Drew Bailey, Greg J. Duncan, Candice Odgers, and Winnie Yu, *Persistence and Fadeout in the Impacts of Child and Adolescent Interventions*, Brisbane, Australia: University of Queensland Life Course Centre, 2015.

[30] Arthur J. Reynolds, Judy A. Temple, Barry A. White, Suh-Ruu Ou, and Dylan L. Robertson, "Age-26 Cost-Benefit Analysis of the Child-Parent Center Early Education Program," *Child Development*, Vol. 82, No. 1, 2011, pp. 379–404.

[31] For further discussion of this topic, see Duncan and Magnuson, 2013; and Lipsey, Farran, and Hofer, 2015.

[32] Duncan and Magnuson, 2013; Camilli et al., 2010; and Bailey et al., 2015.

[33] U.S. Department of Health and Human Services, *Head Start Impact Study: Final Report*, Washington, D.C.: Administration for Children and Families, U.S. Department of Health and Human Services, January 2010; and U.S. Department of Health and Human Services, *Third Grade Follow-Up to the Head Start Impact Study*, Washington, D.C.: Administration for Children and Families, U.S. Department of Health and Human Services, October 2012.

[34] Avi Feller, Todd Grindal, Luke W. Miratrix, and Lindsay C. Page, "Compared to What? Variation in the Impact of Early Childhood Education by Alternative Care-Type Settings," working paper, 2016; Patrick Kline and Christopher Walters, "Evaluating Public Programs with Close Substitutes: The Case of Head Start," working paper No. w21658, Cambridge, Mass.: National Bureau of Economic Research, 2015; and Christopher Walters, "Inputs in the Production of Early Childhood Human Capital: Evidence from Head Start," National Bureau of Economic Research Working Paper No. 20639, Cambridge, Mass.: National Bureau of Economic Research, 2014.

[35] Marianne Bitler, Hilary W. Hoynes, and Thurston Domina, "Experimental Evidence on Distributional Effects of Head Start," working paper No. w20434, Cambridge, Mass.: National Bureau of Economic Research, 2013; and Walters, 2014.

[36] Marianne P. Bitler and Lynn A. Karoly, "Intended and Unintended Effects of the War on Poverty: What Research Tells Us and Implications for Policy," *Journal of Policy Analysis and Management*, Vol. 34, No. 3, 2015, pp. 639–696; and Howard S. Bloom and Christina Weiland, *Quantifying Variation in Head Start Effects on Young Children's Cognitive and Socio-Emotional Skills Using Data from the National Head Start Impact Study*, New York: MDRC, 2015.

[37] Reviews of these studies have been conducted by Janet Currie and Maya Rossin-Slater, "Early-Life Origins of Life-Cycle Well-Being: Research and Policy Implications," *Journal of Policy Analysis and Management*, Vol. 34, 2015, pp. 208–242; and Chloe Gibbs, Jens Ludwig, and Douglas L. Miller, "Does Head Start Do Any Lasting Good?" in Martha J. Bailey and Sheldon Danziger, eds., *Legacies of the War on Poverty*, New York: Russell Sage Foundation, 2013.

[38] Janet Currie and Duncan Thomas, "Does Head Start Make a Difference?" *American Economic Review*, Vol. 85, No. 3, June 1995, pp. 235–262; Eliana Garces, Duncan Thomas, and Janet Currie, "Longer-Term Effects of Head Start," *American Economic Review*, Vol. 92, No. 4, September 2002, pp. 999–1012; David Deming, "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start," *American Economic Journal: Applied Economics*, Vol. 1, 2009, pp. 111–134; and Jens Ludwig and David Miller, "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," *Quarterly Journal of Economics*, Vol. 122, No. 1, 2007, pp. 159–208.

[39] Lipsey, Farran, and Hofer, 2015.

[40] Mark W. Lipsey, Kerry G. Hofer, Nianbo Dong, Dale C. Farran, and Carol Bilbrey, *Evaluation of the Tennessee Voluntary Prekindergarten Program: Kindergarten and First Grade Follow-Up Results from the Randomized Control Design*, Nashville, Tenn.: Peabody Research Institute, Vanderbilt University, August 2013.

[41] Based on kindergarten and first grade administrative data for the two cohorts in the full experimental sample, the Tennessee evaluation also reported descriptive results for special education use (i.e., they did not provide experimental impact estimates, in part because of challenges with coding special education status). In kindergarten (both cohorts) and first grade (first cohort only), the percentage of students using any special education was higher for the preschool participants compared with nonparticipants, but new entry into special education use in either kindergarten or first grade was higher for students who did not attend preschool compared with those who did. These descriptive findings suggest that preschool attendance may lead to earlier identification of the need for special education services. Further follow-up is needed with the experimental design to determine whether cumulative special education use across grades is reduced as a result of preschool participation See Mark W. Lipsey, Dale C. Farran, Carol Bilbrey, Kerry G. Hofer, and Nianbo Dong, *Initial Results of the Evaluation of the Tennessee Voluntary Pre-K Program*, Nashville, Tenn.: Peabody Research Institute, Vanderbilt University, 2011.

[42] Theses figures are based on the Early Childhood Environment Rating Scale–Revised (ECERS–R), which is scored on a scale of 1 to 7, with 1 being "minimal" quality and 7 being "excellent" quality. Other quality measures collected for the Tennessee classrooms included the Early Language and Literacy Classroom Observation (ELLCO) and the Narrative Record (to measure the allocation of time across various activities). See Dale C. Farran, Kerry Hofer, Mark Lipsey, and Carol Bilbrey, *Variations in the Quality of TN–VPK Classrooms*, presentation delivered at the Society for Research on Educational Effectiveness, March 8, 2014.

[43] Duncan and Magnuson, 2013; Lipsey, Farran, and Hofer, 2015; and Carolyn J. Hill, William T. Gormley Jr., and Shirley Adelstein, "Do the Short-Term Effects of a High-Quality Preschool Program Persist?" *Early Childhood Research Quarterly*, Vol. 32, 2015, pp. 60–79.

[44] Janet Currie and Matthew Neidell, "Getting Inside the 'Black Box' of Head Start Quality: What Matters and What Doesn't," *Economics of Education Review*, Vol. 26, No. 1, pp. 83–99; Fuhua Zhai, C. Cybele Raver, and Stephanie M. Jones, "Academic Performance of Subsequent Schools and Impacts of Early Interventions: Evidence from a Randomized Controlled Trial in Head Start Settings," *Children and Youth Services Review*, Vol. 34, No. 5, 2012, pp. 946–954.

[45] Kimber Bogard and Ruby Takanishi, "PK–3: An Aligned and Coordinated Approach to Education for Children 3 to 8 Years Old," *Social Policy Report*, 2005, pp. 3–24.

[46] Yoshikawa et al., 2013.

[47] Jim Minervino, *Lessons from Research and the Classroom: Implementing High-Quality Pre-K that Makes a Difference for Young Children*, Seattle, Wash.: Bill and Melinda Gates Foundation, 2014; and Robert C. Pianta, W. Steven Barnett, Margaret Burchinal, and Kathy R. Thornburg, "The Effects of Preschool Education: What We Know—How Public Policy Is or Is Not Aligned with the Evidence Base, and What We Need to Know," *Psychological Science in the Public Interest*, Vol. 10, 2009, pp. 49–88.

[48] Martha Zaslow, Rachel Anderson, Zakia Redd, Julia Wessel, Louisa Tarullo, and Margaret Burchinal, *Quality Dosage, Thresholds, and Features in Early Childhood Settings: A Review of the Literature (OPRE 2011-5)*, Washington, D.C.: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2010.

[49] Edward F. Zigler and Susan Muenchow, *Head Start: The Inside Story of America's Most Successful Educational Experiment*, New York: Basic Books, 1992.

[50] For further discussion of the trade-offs between targeted versus universal programs, see Lynn A. Karoly, *Preschool Adequacy and Efficiency in California: Issues, Policy Options, and Recommendations*, Santa Monica, Calif.: RAND Corporation, MG-889, 2009. For evidence of peer learning, see Christina Weiland and Hirokazu Yoshikawa, "Does Higher Peer Socio-Economic Status Predict Children's Language and Executive Function Skills Gains in Prekindergarten?" *Journal of Applied Developmental Psychology*, Vol. 35, 2014, pp. 422–432.

[51] Minervino, 2014.

[52] The evaluation of the Tulsa preschool program showed about 65 percent of participating children had income below 185 percent of the federal poverty level (i.e., free or reduced-price lunch eligible), while the remaining 35 percent had income above that threshold. See William T. Gormley, Ted Gayer, Deborah Phillips, and Brittany Dawson, "The Effects of Universal Pre-K on Cognitive Development," *Developmental Psychology*, Vol. 41, No. 6, 2005, pp. 872–884.

[53] Yoshikawa et al., 2013.

[54] James Leak, Greg Duncan, Weilin Li, Katherine Magnuson, Holly Schindler, and Hirokazu Yoshikawa, "Is Timing Everything? How Early Childhood Education Program Cognitive and Achievement Impacts Vary by Starting Age, Program Duration and Time Since the End of the Program," working paper, 2010.

[55] Gormley et al., 2005.

[56] Timothy J. Bartik, William Gormley, and Shirley Adelstein, "Earnings Benefits of Tulsa's Pre-K Program for Different Income Groups," *Economics of Education Review*, Vol. 31, 2012, pp. 1143–1161.

[57] Walker A. Swain, Matthew G. Springer, and Kerry G. Hofer, "Early Grade Teacher Effectiveness and Pre-K Effect Persistence," *AERA Open*, Vol. 1, 2015, doi: 10.1177/2332858415612751.

[58] Amy Claessens, Mimi Engel, and F. Chris Curran, "Academic Content, Student Learning, and the Persistence of Preschool Effects," *American Educational Research Journal*, Vol. 51, 2014, pp. 403–434.

[59] Kristie Kauerz, *P–3: What It Is, Why It's Important, and State Approaches*, National Association of Latino Elected and Appointed Officials (NALEO) Legislative Hearing, September 23, 2008.

[60] Gail L. Zellman and M. Rebecca Kilburn, *Final Report on the Hawai'i P–3 Evaluation*, Santa Monica, Calif.: RAND Corporation, RR-1100-PPE, 2015.

[61] Minervino, 2014; and Pianta et al., 2009.

[62] Eugene Steurle and Leigh Miles Jackson, eds., *Advancing the Power of Economic Evidence for Making Informed Investments in Children, Youth, and Families*, Washington, D.C.: National Academies Press, forthcoming.

[63] Washington State Institute for Public Policy, 2015.

[64] See for example Timothy J. Bartik, *From Preschool to Prosperity: The Economic Payoff to Early Childhood Education*, Kalamazoo, Mich.: W.E. Upjohn Institute, 2014.

[65] Lynn A. Karoly, "Toward Standardization of Benefit-Cost Analyses of Early Childhood Interventions," *Journal of Benefit-Cost Analysis*, Vol. 3, No. 1, Article 4, 2012.

[66] Recent examples citing this high benefit-cost ratio include Institute for a Competitive Workforce, *Why Business Should Support Early Childhood Education*, Washington, D.C.: U.S. Chamber of Commerce, 2010; and Committee for Economic Development, *Unfinished Business: Continued Investment in Child Care and Early Education Is Critical to Business and America's Future*, Washington, D.C., 2012. The estimated return is based on W. Steven Barnett, Clive R. Belfield, and Milagros Nores, "Lifetime Cost-Benefit Analysis," in Lawrence J. Schweinhart, Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores, *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40 (Monographs of the High/Scope Educational Research Foundation, No. 14)*, Ypsilanti, Mich.: High/Scope Press, 2005, pp. 130–157.

[67] The benefit-cost analysis for the Tulsa program also provides estimates of the return for the part-day version of the preschool program. The estimates show that costs are half as much relative to the full-day program ($5,344 per child in 2015 dollars), but because the program impacts are somewhat more than half as large for the lowest and highest income groups, the estimated benefit-cost ratio is slightly higher for those groups (4.08 and 3.44, respectively), but somewhat lower for the income group in between (2.99). See Bartik, Gormley, and Adelstein, 2012.

[68] Reynolds et al., 2011.

[69] Lynn A. Karoly and James H. Bigelow, *The Economics of Investing in Universal Preschool Education in California*, Santa Monica, Calif.: RAND Corporation, MG-349, 2005.

[70] Clive R. Belfield, *An Economic Analysis of Pre-K in Arkansas*, Washington, D.C.: Pre-K Now, 2006; and Clive R. Belfield, "The Fiscal Impacts of Universal Pre-K: Case Study Analysis for Three States," Invest in Kids Working Paper No. 6, July 2006.

About This Report

Cincinnati, Ohio, is one of a growing number of U.S. cities seeking to expand access to and raise the quality of preschool programs, especially for the city's most vulnerable children. In a companion report, we document the current preschool landscape in Cincinnati and consider design options for allowing more low-income children to enroll in high-quality programs. In this report, we aim to compile the most-reliable research evidence concerning whether high-quality preschool programs benefit participating children and whether the economic value of outcomes experienced by participating children exceeds the economic costs to implement the program. Our review draws on evidence from full-scale preschool programs implemented at the national, state, and local levels.

The authors would like to thank Stephanie Byrd of the United Way of Greater Cincinnati, Greg Landsman of the Cincinnati Preschool Promise, Emily Lewis of the Strive Partnership, and Gary Lindgren and Kristi Phillips of the Cincinnati Business Committee and Cincinnati Regional Business Committee for their guidance on the project, along with feedback provided by their colleagues on the Cincinnati Preschool Review Committee. The authors also benefited from the thorough and constructive reviews provided by Becky Kilburn of RAND and Christina Weiland of the University of Michigan.

This research was supported through a contract from the Cincinnati Business Committee and the United Way of Greater Cincinnati. The report is part of a larger study examining options for expanding preschool quality and access in Cincinnati.

This research was conducted jointly in RAND's Education, and Labor and Population units. Additional information about RAND is available at www.rand.org.

Limited Print and Electronic Distribution Rights

For more information on this publication, visit www.rand.org/t/rr1461.

www.rand.org

RR-1461-CBC/UWGC