49. Pre-K Education

Congress should

- recognize that the promises of large returns on investment for universal preschool programs are grossly overstated;
- recognize that the high-quality research on large-scale preschool programs fails to find lasting positive effects on participating students;
- understand that a universal preschool program is likely to cost tens of billions of dollars without measurably improving student outcomes;
- end direct federal subsidies of preschool programs; and
- refrain from enacting a universal preschool program.

Perhaps the most popular new entitlement that politicians are proposing is universal preschool education. Because such a program would carry a hefty price tag—President Barack Obama called for spending $75 billion over 10 years—proponents often make grand claims about its return on investment. These claims are not supported by the research.

The studies on which proponents rely are of small-scale, high-intensity programs operational several decades ago that bear little resemblance to the large-scale, lower-intensity programs being proposed today. No high-quality research has found such big and lasting positive impacts from larger-scale programs. Indeed, the highest quality studies of the federal Head Start program and other larger-scale preschool efforts tend to find little to no lasting impact. In one example, by the time students reached second grade, Tennessee’s preschool program had a negative impact on participating students.

The federal government has no constitutional authority to enact a universal preschool program. Even if it did, the research does not support
making such a huge financial investment in a program that is likely to produce little or no return. Moreover, because federal grants usually flow through state education agencies, they tend to fund programs that are housed in existing school facilities that are intended for older students and are unsuitable for three- and four-year-olds. By crowding out better-suited alternatives, government subsidies for preschool may actually make things worse.

**Probing the Promises about Preschool Programs**

Touting universal preschool is all the rage these days. In his 2013 State of the Union address, President Obama proposed “working with states to make high-quality preschool available to every single child in America” through a “Preschool for All” program that would cost $75 billion over 10 years. The president justified the price tag by pointing to research that supposedly indicates there would be a large return on the investment:

> Every dollar we invest in high-quality early childhood education can save more than seven dollars later on—by boosting graduation rates, reducing teen pregnancy, even reducing violent crime. In states that make it a priority to educate our youngest children, like Georgia or Oklahoma, studies show students grow up more likely to read and do math at grade level, graduate high school, hold a job, form more stable families of their own. We know this works.

Likewise, Hillary Clinton has claimed, “we already know that for every one dollar we spend on early childhood education, we reap seven dollars as a society in returns.” In recent years, numerous gubernatorial candidates and state legislators have echoed calls for universal pre-K.

Those promises are not supported by the research. Universal preschool proponents generally rest their claims on research on the High/Scope Perry Preschool Project (the source of the “seven-to-one” rate of return claim) and the Carolina Abecedarian Project. Although research on both programs showed lasting results, both were very small, expensive, high-intensity programs in the 1960s and 1970s that differ significantly from the large-scale, low-intensity programs being proposed today. Moreover, as detailed below, high-quality research on comparable large-scale early education programs has failed to replicate the strong positive results from Perry and Abecedarian. As David Armor, a leading scholar on education policy and quantitative methods, explained in a 2014 Cato Institute report:
There is a body of research that finds educational benefits from preschool programs, but these studies suffer from one of two problems: Either the preschool programs are not comparable with the programs being proposed today, or they use non-experimental designs that suffer from serious limitations, including an inability to track preschool effects into the early grades.

At present, there simply is insufficient evidence to conclude that large-scale preschool programs produce significantly positive results.

**Examining the Evidence on Early Education**

**Georgia and Oklahoma**

In his 2013 State of the Union, President Obama made specific reference to the state-funded preschool programs in Georgia and Oklahoma. Although research has suggested some benefits for disadvantaged students, there is little evidence that these programs have significantly improved educational outcomes for participating students overall. Moreover, the research used very limiting methods.

Georgia initially enacted a means-tested preschool program in 1992 and expanded it to include all children in 1995. Research finds some evidence that the program benefits some disadvantaged students, at least at first. A 2008 study by a researcher at Stanford University’s Institute for Economic Policy Research found that “disadvantaged children residing in small towns and rural areas” who attended preschool in Georgia were more likely to have higher reading and math scores in fourth grade. However, the study found no consistent and statistically significant benefits to middle-income students. The researcher concluded that universal preschool failed a cost-benefit analysis:

The costs of the program today ($302 million in 2007–2008) greatly outweigh the benefits in terms of potential increased taxable revenue. This is a very simple cost-benefit analysis and should therefore be interpreted with caution. However, it is at least suggestive that the government’s scarce resources would be better spent on more targeted early childhood interventions that have been shown to be more cost efficient, particularly if the goal is to increase wages through test scores.

Oklahoma enacted a universal preschool program in 1998. A study of participating children in Tulsa found much larger positive impacts than in Georgia, the equivalent of about eight months of learning for verbal
skills. Although the results appear quite impressive, they may have been an artifact of the research design. A later study that examined preschool programs in five states, including Oklahoma, failed to detect similarly large results.

However, both the Georgia and Tulsa studies employed regression discontinuity design (RDD) instead of random assignment (the gold standard). In these studies, the test scores of children who just made the age cutoff for preschool are compared with those who just missed it. As Armor explained in 2014, there are several potential problems with this approach. First, attrition from the treatment group can bias results upward. This was a particular problem in the Tulsa study. Second, the variable used in RDD analyses to separate the treatment and control groups “must not be confounded with any other characteristic that might cause different behaviors in the treatment and control group.” However, it is “likely that the age variable affects many other variables, such as the friends a child plays with, the amount of time spent in out-of-home care, and activities with parents,” all of which can distort the findings. For example, parents of a child who is not going to attend preschool for another year are likely to treat her differently than if she were enrolled in preschool. Third, RDD analyses can only estimate the impact of a policy in a particular year. RDD cannot compare treatment and control groups over time because the control group enters preschool in the following year, and therefore RDD cannot tell whether any perceived effects are lasting or fade out. This is crucial because high-quality research on large-scale preschool programs has often found that positive effects fade out after just a few years.

To get a clearer picture of the effect of universal preschool programs, we should look to random-assignment studies of large-scale preschool programs that track students over time.

The High/Scope Perry Preschool and Carolina Abecedarian Projects

Proponents of universal preschool often point to two random-assignment studies that found positive outcomes for disadvantaged students. However, the programs that they studied differed significantly from the types of efforts under discussion today.

Beginning in 1962, the High/Scope Perry Preschool Project studied 123 children from low-income households in Ypsilanti, Michigan. The study randomly assigned 58 children to a “treatment group” and enrolled those students in the Perry Preschool; the remaining children formed a “control group” of students who were not enrolled. The study tracked
the outcomes of both groups through age 40, finding that participants in the treatment group were less likely to be arrested and more likely to graduate from high school, obtain employment, and earn higher incomes than the control group. Accordingly, the researchers estimated a societal return on investment of $7.16 for every $1.00 expended, factoring in increased tax revenues, decreased welfare payments, lower crime rates, and so on.

Like Perry, the Abecedarian Project studied a small-scale, high-intensity program for mostly black students from low-income households. The project studied 111 students beginning in 1972 in Chapel Hill, North Carolina, with a treatment group of 57 students. Decades later, researchers found that the program produced positive outcomes, including lower rates of teenage pregnancy and higher rates of college matriculation and employment in skilled jobs.

However, these findings should be interpreted with great caution. First, the sample sizes—fewer than 60 students in the treatment group in each study—are tiny. Second, both studies had flaws in their randomization process that may have biased the results. Moreover, even if there had been no methodological issues, it would be unwise to assume that large-scale programs would produce similar results because the two earlier programs differed significantly from the sorts of universal preschool programs proposed today.

Program management. Both programs were run by people who were trying to prove that their model worked, rather than by the types of people who would be staffing preschool centers in a large-scale program.

Services. Both Perry and Abecedarian were high intensity. Perry offered a student-to-teacher ratio of about five or six to one, held regular group meetings with parents and teachers, and even had weekly home visits. Abecedarian students received full-time, year-round care for five years beginning in their first year of life; individualized education activities that changed as the child grew; transportation; a three-to-one student-to-teacher ratio for younger students that grew to six-to-one for older students; nutritional supplements; social services; and more. Those services are not comparable to standard preschool programs, which have significantly more students per classroom and offer few of the above services.

Cost. In 2016 dollars, Perry cost more than $21,000 per student and Abecedarian cost more than $22,000, compared with less than $7,000 per student on average in most state programs. No one is proposing spending anything remotely close to that amount per student today.


**Students.** Whereas the Perry Preschool and Abecedarian projects were targeted to at-risk students from low-income households, universal preschool programs would also include students from middle- and upper-income families who are not nearly as likely to reap such large benefits.

The Perry Preschool and Abecedarian projects simply bear no resemblance to the sorts of programs being proposed today. Grover J. “Russ” Whitehurst of the Brookings Institution colorfully cautioned against extrapolating from Perry and Abecedarian, which he said “demonstrate the likely return on investment of widely deployed state pre-K programs . . . to about the same degree that the svelte TV spokesperson providing a testimonial for Weight Watchers demonstrates the expected impact of joining a diet plan.”

Proponents of universal preschool also point to a few others studies, including those on the Abbott program in New Jersey, the Boston Preschool Program, and Chicago Child-Parent Centers. But none of those studies were gold-standard studies of large-scale programs that tracked students over time.

**Results from Rigorous Research on Large-Scale Programs**

As noted above, the best test of universal preschool is research that employs the most rigorous method—random assignment—and studies the impact of large-scale programs over time. Although studies meeting those criteria have found some short-term gains, the gains fade after just a few years. No rigorous research has found lasting gains from large-scale programs.

**Head Start**

Perhaps the most relevant research pertains to the federal Head Start program: that research is national in scope and tracked students through third grade. Enacted in 1965, Head Start provides educational and social services to low-income families nationwide. In 2015, it served nearly 1 million students (approximately 10 percent of the nation’s 3- and 4-year-olds) at an annual cost of about $8.5 billion.

According to Brookings’ Whitehurst, the federal government’s Head Start Impact Study (HSIS) was “one of the most ambitious, methodologically rigorous, and expensive federal program evaluations carried out in the last quarter century.” The HSIS followed a cohort of over 4,500 students who were 3 or 4 years old in the fall of 2002. It employed
random assignment, drawing a sample of about 2,600 students who had applied to more than 300 Head Start centers in 23 states. That sample was contrasted with a control group of about 1,800 children. The study employed statistical techniques to account for some treatment group students who got into but did not attend Head Start, as well as some control group students who found other preschools. The HSIS found modest gains for participating students on both reading and math scores during their preschool years and some positive effects for social behaviors. However, those effects did not last beyond kindergarten. On tests at the end of first and third grade, the treatment and control groups were statistically indistinguishable.

The results of the study were so disappointing for Head Start boosters that the Obama administration released the results in 2012 on the Friday before Christmas—a day on which they were least likely to get any attention.

In a reanalysis of the data, Dr. Peter Bernardy compared only treatment group students who remained in Head Start for at least two years against control group students who did not attend another preschool. Bernardy’s reanalysis likewise found that any early gains faded by the end of kindergarten. Moreover, as shown in Figure 49.1, the control group students scored two points higher than the Head Start students by the end of first grade, although this difference was not statistically significant.

Some Head Start proponents have theorized that perhaps there are “sleeper effects” from the program that only turn up much later. Whitehurst has criticized that theory, noting that “research on the impacts of early intervention consistently shows that programs with longer-term impacts also evidence shorter-term impacts in elementary school.” It is highly unlikely that Head Start is producing significant and lasting positive effects that are undetectable in the interim.

Nevertheless, one 2016 study by the Hamilton Project at the Brookings Institution purports to find such “sleeper effects,” including increased rates of high school graduation and post-secondary attainment. However, unlike the Impact Study, the Hamilton study did not employ random assignment. Instead, using data from the National Longitudinal Survey of Youth, the Hamilton study attempts to generate treatment and control groups by comparing the outcomes of people who had attended Head Start against the outcomes of siblings who did not. In order for this method to truly isolate the impact of Head Start, the sibling pairs must have the same average characteristics, except for Head Start attendance. The authors
concede that this assumption does not necessarily hold. Parents’ decisions to enroll one child in Head Start and not another may suggest significant yet unobserved differences.

In another 2016 study, researchers from Georgetown University found lasting gains for students who had participated in Tulsa’s Community Action Project Head Start, including positive effects on middle-school students’ math test scores (but not reading scores), grade retention, and chronic absenteeism overall, and for some subgroups, such as girls (but not boys) and white and Hispanic students (but not black students). Yet again, the study did not employ random assignment. Instead, the researchers matched students for comparison based on several demographic variables, meaning that the study may suffer from selection bias. Unlike the random-assignment Impact Study, the parents of the treatment group students in the Georgetown study actively chose to enroll their children in Head Start, whereas the parents of the matched comparison group
did not, meaning that there could be significant, unobserved differences between these two populations that cannot be eliminated merely by controlling for demographic characteristics. Moreover, as the authors concede, Tulsa’s CAP Head Start program “is not representative of Head Start programs across the country.” Tulsa’s program has had “higher scores on observational assessments of instructional quality and on time spent on academic instruction than was true of an 11-state sample of Head Start programs assessed at the same time.”

Although the Georgetown study provides some suggestive evidence that certain types of early education programs may have some lasting positive impacts for some types of students, the highest quality research on Head Start programs nationwide shows that any initial gains are, on average, short lived.

**Tennessee Voluntary Pre-K Initiative**

In 2005, Tennessee enacted the Voluntary Pre-K Initiative, which gave priority to high-risk students (e.g., those in poverty, children with disabilities, and students with limited English proficiency). By 2007, more than 18,000 students were participating. The program was deemed “high quality” because, like similar programs in Tulsa and Boston, teachers had to be licensed and instruction was offered for a minimum of 5.5 hours for five days a week (usually 6 to 8 hours). Tennessee was also one of the 18 states to receive some of the $226 million in federal Preschool Development Grants in 2015.

Researchers at Vanderbilt University conducted a random-assignment study with a sample size of 3,000 students. Because of a problem with the randomization—parental permission to collect data was obtained only after the students had been randomly assigned to control and treatment groups—the researchers had to rely on a statistical technique known as propensity analysis to compensate. As with Head Start, the study initially found significant positive effects on academic performance during preschool, but those effects faded. By second and third grade, the participating students actually scored worse than the control group.

The Tennessee example also demonstrates how federal subsidies can distort the market. Because the grants flow primarily through state education agencies, the funds generally support programs operated in existing district schools. When preschool programs are housed in facilities built for older students, bathrooms and cafeterias are often relatively far from their classrooms. Vanderbilt’s observational research found that transition
time between activities consumed the largest chunk of time during the day (about 25 percent). By contrast, only about 3 to 4 percent of the time was spent on outdoor play or gym time, and many students had no opportunity to run or play. Once again, housing a preschool in a district school was a contributing factor because their playgrounds are generally designed for older students.

**Conclusion and Recommendations**

Proponents of universal preschool rest their case on a thin empirical reed. The programs that produced large and lasting positive effects were small, high-intensity, prohibitively expensive, and not comparable to the sorts of programs being proposed today. In contrast, the most rigorous research on large-scale programs has consistently found that positive effects tend to fade within a few years. Even if the Constitution granted the federal government the authority to do so, the research literature does not support enactment of a universal preschool program. Instead, Congress should end subsidies for preschool programs, such as Head Start and the Preschool Development Grants.

**Suggested Readings**


Lipsey, Mark W., Dale C. Farran, and Kerry G. Hofer. “A Randomized Control Trial of the Effects of a Statewide Voluntary Prekindergarten Program on Children’s Skills and Behaviors through Third Grade.” Nashville, TN: Vanderbilt University, Peabody Research Institute, September 2015.


—Prepared by Jason Bedrick