Early Childhood Research Quarterly xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade

Mark W. Lipsey*, Dale C. Farran, Kelley Durkin

Vanderbilt University, United States

ARTICLE INFO

Article history:
Received 9 June 2017
Received in revised form 20 February 2018
Accepted 9 March 2018
Available online xxx

Keywords:
Public pre-k
Randomized control trial
Longitudinal
Early childhood education
Achievement
Policy

ABSTRACT

This report presents results of a randomized trial of a state prekindergarten program. Low-income children ($N = 2990$) applying to oversubscribed programs were randomly assigned to receive offers of admission or remain on a waiting list. Data from pre-k through 3rd grade were obtained from state education records; additional data were collected for a subset of children with parental consent ($N = 1076$). At the end of pre-k, pre-k participants in the consented subsample performed better than control children on a battery of achievement tests, with non-native English speakers and children scoring lowest at baseline showing the greatest gains. During the kindergarten year and thereafter, the control children caught up with the pre-k participants on those tests and generally surpassed them. Similar results appeared on the 3rd grade state achievement tests for the full randomized sample – pre-k participants did not perform as well as the control children. Teacher ratings of classroom behavior did not favor either group overall, though some negative treatment effects were seen in 1st and 2nd grade. There were differential positive pre-k effects for male and Black children on a few ratings and on attendance. Pre-k participants had lower retention rates in kindergarten that did not persist, and higher rates of school rule violations in later grades. Many pre-k participants received special education designations that remained through later years, creating higher rates than for control children. Issues raised by these findings and implications for pre-k policy are discussed.

In 2015, 67% of U.S. children 4 years old and not in kindergarten were enrolled in preschool programs (McFarland et al., 2017). As in years past, higher income families were more likely to enroll their children in some form of center-based care than low-income families, and low-income children were more likely to be enrolled in public programs such as Head Start and state-funded prekindergarten programs. Many states have been prompted to increase funding for pre-k programs in order to serve a greater number of high-risk children (Parker, Workman, & Atchison, 2016) and most states currently offer some form of voluntary pre-k that is available to children from low-income families (Barnett et al., 2017).

State funding targeted to children from low-income families implies goals beyond merely providing daycare. For example, Mississippi began state funding of pre-k in 2014 after lobbying by Mississippi First about the role pre-k can play "... in closing the achievement gap while raising achievement for all learners" (http://www.mississippifirst.org/education-policy/pre-kindergarten/). In 2014–2015, the U.S. Department of Education allocated millions of dollars to states to expand pre-k, citing a white paper asserting that high quality early education narrows achievement gaps, boosts adult earnings, and results in savings of $8.60 for every $1 spent (Executive Office of the President of the United States, 2014). With such high expectations, it is especially important for policy to be informed by research on the effects of state-funded pre-k.

1. Pre-k effects at kindergarten entry

One relevant body of evidence demonstrates that state pre-k programs generally improve such aspects of children's readiness for kindergarten as letter recognition and print awareness (Gormley, Gayer, Phillips, & Dawson, 2005; Wong, Cook, Barnett, & Jung, 2008). Most of what is known about these immediate pre-k effects comes from age-cutoff regression-discontinuity designs (RDD). Though not without potential biases (Lipsey, Weiland, Yoshikawa, Wilson, & Hofer, 2015), this design has the twofold advantage of being recognized as a relatively strong design while also being easily applied to any program with an age cutoff for admission. Chil-

* Corresponding author at: Peabody Research Institute, Vanderbilt University, 230 Appleton Place, PMB 181, Nashville, TN 37203-5721, United States.
E-mail addresses: mark.lipsey@vanderbilt.edu (M.W. Lipsey), dale.farran@vanderbilt.edu (D.C. Farran), kelley.durkin@vanderbilt.edu (K. Durkin).

https://doi.org/10.1016/j.ecresq.2018.03.005

dren with birthdays on one side of the cutoff are admitted; those on the other side must wait until the following year. The outcomes of interest can be measured after the admitted group completes pre-k, and the group in waiting is just beginning, and then compared with statistical adjustments for the age difference.

The age-cutoff RDD was first used in an evaluation of the Tulsa pre-k program that found positive effects (Gormley et al., 2005). A number of similar age-cutoff RDD studies have been conducted since and have almost universally found positive effects (Wong et al., 2008). A recent example is the age-cutoff RDD study of the program in Boston carried out by Weiland and Yoshikawa (2013) that has received attention for its very positive findings. A disadvantage of the RDD, however, is that it does not allow investigation of pre-k effects after entry into kindergarten because, by then, the control group has also completed pre-k.

These studies demonstrate that state-funded pre-k programs can produce positive effects on various target outcomes prior to kindergarten entry. However, questions about the nature of the effects have been raised. This research has focused on basic pre-reading skills, but the influence of pre-k programs on other outcomes pertinent to children's cognitive and behavioral development, such as complex language skills, mathematics, self-regulation, and social skills, is less clear (Gormley et al., 2005; Jackson et al., 2007). Skills of this latter sort may be more critical for children's long-term performance in school and beyond and thus make better targets for pre-k intervention (Bailey, Duncan, Odgers, & Yu, 2017).

2. Long-term pre-k effects

State investments in pre-k are most often justified by the expectation of long-term effects (e.g., Executive Office of the President of the United States, 2014; Heckman, 2006). This expectation derives mainly from longitudinal research that reported positive outcomes on school completion, employment, marriage stability, criminal behavior, and the like for two model programs – Perry Preschool, mounted in the 1960s, and Abecedarian, begun in the 1970s. Both programs served a small number of children in a single location, and neither has been fully replicated in contemporary publicly funded programs. Indeed, the political feasibility of implementing them at scale is doubtful. These programs would cost more than any state currently allocates – $20,000 per child per year in today's dollars to implement Perry and $16,000–$40,000 for Abecedarian (Minervino & Pianta, 2014).

Attempts to evaluate longer-term effects of state-funded pre-k programs implemented in more recent times have been problematic. Random assignment of children to conditions in which some attend pre-k and others do not, or matching on relevant cognitive, family, and demographic baseline variables, requires that the research sample be identified prior to the beginning of the pre-k year. However, because state pre-k is voluntary, there are few situations in which families intending to enroll their children can be identified in advance and persuaded to participate in random assignment or provide adequate baseline data for matching those who follow through with enrollment and those who do not.

As a result, the largest group of studies of longer-term state pre-k effects compares outcomes for children identified sometime after the pre-k year who did and did not attend pre-k (e.g., Andrews, Jargowsky, & Kuhne, 2012; Gormley, Phillips, & Anderson, 2018; Huang, Invernizzi, & Drake, 2012; Peisner-Feinberg, Mokrova, & Anderson, 2017). These post hoc studies lacking both random assignment and true baseline measures collected prior to the pre-k year are quite vulnerable to selection bias from initial differences on unobserved variables. In short, why did some parents take advantage of a voluntary pre-k program while others did not, and how is that related to family and child characteristics that might influence later outcomes? The demographic variables collected in later years with which these samples are typically matched are unlikely to be sufficient to account for all the relevant differences between children whose parents made and sustained the effort to have them attend pre-k and those who did not.

Another distinct group of studies of longer-term effects of state-funded pre-k programs uses difference-in-difference (DD) methods that examine student outcomes before and after states or counties increased pre-k implementation compared to differences over the same period for areas with no analogous pre-k expansion. The challenge for these studies is to isolate the difference made by variation in pre-k implementation from other influential factors occurring in the same locations over the same period. To do so, they rely on complex statistical models, but those do not always yield robust results. Fitzpatrick (2008), for example, used a DD design to investigate the effects of the Georgia universal pre-k program that grew from participation rates of 14% in 1995 to 55% in 2008. Some analysis models showed positive effects on 4th grade NAEP reading and math scores while others did not. Similar effects that were generally positive, but sensitive to the selection of comparison states, were reported by Cascio and Schanzenbach (2013) for the Georgia and Oklahoma programs. In contrast, DD analyses with extensive data on the More at Four pre-k program in North Carolina showed effects on 3rd grade state achievement scores that were robust to a range of model variations, though the authors acknowledged that the resulting estimates were too large to plausibly represent direct pre-k effects and hypothesized that there must be spillover to nonparticipating children (Ladd, Muschkin, & Dodge, 2014).

The overarching theme in research on long-term effects of state pre-k programs is one of methodological challenge. When dealing with a voluntary program with children's participation always a matter of self-selection by parents, it is difficult for researchers to ensure that they are comparing outcomes for pre-k participants and nonparticipants who are similar in all ways that matter prior to their differential pre-k experience. The result is an uneven and inconclusive research literature. As the experts assembled by the Brookings Institute who recently reviewed virtually all the research on the effects of state pre-k programs reported, "Convincing evidence on the longer-term impacts of scaled-up pre-k programs on academic outcomes and school progress is sparse, precluding broad conclusions" (Phillips et al., 2017, p. 9).

In this context, the Head Start Impact study (Puma, Bell, Cook, & Heid, 2010; Puma et al., 2012) warrants attention. While not a study of state pre-k, it is the only previous randomized study of a public pre-k program. This study began in 2002 with a national sample of 5000 children who applied to 84 programs expected to have more applicants than spaces. Children were randomly selected for offers of admission with those not selected providing the control group. The 4-year-old children admitted to Head Start made greater gains across the pre-k year than nonparticipating children on measures of language and literacy, although not on math. However, by the end of kindergarten the control children had caught up on most achievement outcomes; subsequent positive effects for Head Start participants were found on only one achievement measure at the end of 1st grade and another at the end of 3rd grade. There were no statistically significant effects on social–emotional measures at the end of the pre-k or kindergarten years. A few positive effects appeared in parent reports at the end of the 1st and 3rd grade years, but teacher and child reports in those years showed either null or negative effects.

The positive short-term effects found in the Head Start study are consistent with those found for state pre-k programs. The mixed and null effects found thereafter in this methodologically strong study, however, raise questions about the expectation of substantial long-term benefits that has largely motivated investments in

Please cite this article in press as: Lipsey, M. W., et al. Effects of the Tennessee Prekindergarten Program on children's achievement and behavior through third grade. *Early Childhood Research Quarterly* (2018), https://doi.org/10.1016/j.ecresq.2018.03.005