Abstract Title Page

Title:

The Persistence of Pre-K Effects and Early Grade Teacher Quality: Evidence from the Tennessee-Voluntary Pre-K Experiment

Authors and Affiliations:

Matthew G. Springer, Vanderbilt University’s Peabody College; Tennessee Consortium on Research, Evaluation, and Development

Walker A. Swain, Vanderbilt University’s Peabody College; Tennessee Consortium on Research, Evaluation, and Development

Kerry G. Hofer, Vanderbilt University’s Peabody College; Peabody Research Institute
Abstract Body

Background / Context:

In recent years, state-financed pre-school programs have expanded dramatically. Enrollment in the past decade in state programs has more than doubled, with several states going as far as offering universal programs (Hustedt & Barnett 2011). President Obama has made a concerted effort to push legislation that would make universal pre-K access federal law. However, much of the research cited by politicians supporting these types of broad expansions comes from a few resource-intensive targeted experimental programs that have demonstrated remarkable benefits (Nores et al. 2005; Heckman et al. 2010; Campbell et al. 2008; Campbell et al. 2012). Unfortunately, these model programs may have little in common with the type of programs states have implemented at scale and propose to grow in the future.

Purpose / Objective / Research Question / Focus of Study:

Advocates of early childhood education generally view pre-schooling intervention as a vital and underutilized tool to narrow racial and socioeconomic outcome gaps in school and beyond (i.e. Dogget & Wat 2010). Opponents tend to argue that the benefits of the programs are too short-lived to justify the costs (i.e. Dalmia & Snell 2008). One important way research can inform this debate is by developing a stronger understanding of the factors that contribute to or inhibit the persistence of pre-school benefits. In this paper we utilize data from Vanderbilt’s Peabody Research Institute’s (PRI) public pre-K evaluation in Tennessee, matched with school administrative records and data from a new teacher evaluation program, to examine the interaction between pre-K participation and a factor that is as elusive to measure as it is universally accepted as vital to student outcomes—teacher quality.

Setting:

The setting for this study is public elementary schools in Tennessee.

Population / Participants / Subjects:

Prior to the start of the 2009-10 and 2010-11 school years, two cohorts of over 3,000 total children were offered admission to the Tennessee Voluntary pre-K program (TN-VPK) off randomly-ordered applicant lists at oversubscribed sites across the state that were willing to participate in the experimental evaluation. PRI has followed (and is continuing to follow) both groups of students into later grades, those who were randomly admitted to the TN-VPK program and those who were randomly left on the wait list, collecting state data on the whole sample and directly assessing a smaller subsample (referred to as the Intensive Subsample, or ISS). While the counterfactual is a “business as usual” control that makes no adjustments to account for parents who sought out alternative preschool programs, a parent survey revealed that more than half of the ISS analytic sample stayed at home with a parent or other caregiver. Just over 11 percent enrolled in Head Start, and another 15 percent had formal private childcare.
In order for a child to be included in the ISS, the child had to (1) meet all eligibility criteria for participation in the experiment; (2) have parental consent to participate in the study; and (3) be assessed by PRI staff at least once during the pre-K or kindergarten school year. Additionally, the consented, eligible children on a school’s applicant list were only included in the ISS if there were at least one consented, assessed child who participated in TN-VPK and at least one consented child that did not participate from that randomized applicant list. Furthermore, to construct the analytic sample, PRI restricted the sample to the 1076 students who were assessed at the end of pre-K. Participating students were given a battery of individually-administered achievement tests at the beginning and end of the pre-K year and then again at the end of each subsequent school year. All results reported in this paper are based on assessment data for the 1076 students in the ISS analytic sample.

**Intervention / Program / Practice:**

See above.

**Research Design:**

For this study, we draw on information and data from PRI’s TN-VPK experiment matched with school administrative records and data from Tennessee’s teacher evaluation program.

**Data Collection and Analysis:**

The data in this study come from three primary sources. First, student-level data was collected by researchers at PRI as part of a large-scale experimental evaluation of the TN-VPK program. Second, teacher evaluation records were collected by the Tennessee Department of Education (TNDOE) and cleaned by the Tennessee Consortium on Research, Evaluation, and Development (the Consortium). With regard to teacher evaluation records, we focus on the overall level of effectiveness score which is calculated using individual and school-level student growth scores and achievement data as well as classroom observations for teachers in tested and untested subjects and grades.

Our primary research question concerns the interaction of pre-K participation and 1st grade teacher quality as it correlates with student achievement, which we interpret as the effect of early grade teacher quality on the persistence of pre-K effects. We model this interaction as a second difference using binary indicators for three potential measures of teacher quality for the student’s 1st grade teacher. We estimate interactions separately for (1) teachers whose overall evaluation rating makes them subject to sanctions or increased scrutiny (a level-3 rating or below); (2) teachers whose ratings indicate exemplary performance (level-5 rating); and (3) a teacher’s

---

1 To be eligible for inclusion in the randomized sample, children had to meet certain criteria, which included: (1) assigned to either TN-VPK or the control condition on the basis of their position on the randomized applicant list (i.e., not automatically let in because of sibling preference, etc.); (2) age-eligible (age and subsequent grade progression indicated that the child was old enough to attend kindergarten the next school year); (3) income-eligible (based on the exclusion of those children who did not meet standard for the FRPL program; and (4) placed in a regular TN-VPK classroom (not a blended or special education classroom).
specific final evaluation rating on the 1 to 5 scale. We use OLS regression techniques. Select specifications include robust standard errors clustered at the school list level.

Findings / Results:

All findings and results are preliminary and should not be cited at this time.

For each indicator of 1st grade teacher quality, students who participated in TN-VPK perform significantly higher than their control group counterparts on the Woodcock Johnson composite score in 1st grade when taught by better 1st grade teachers. With 1st grade teachers who are rated poorly on their educator evaluations, students in the control group outperform the pre-K participants. However, with each step up the teacher evaluation rating scale, students who participated in TN-VPK performed better. When taught by a highly-rated 1st grade teacher, TN-VPK participants outperformed similarly-situated control group students by roughly a tenth of a standard deviation on the composite cognitive score. The positive relationship between TN-VPK participation and 1st grade teacher quality is relatively consistent across the different Woodcock Johnson subscales, with treatment students sometimes scoring as much as .35 standard deviations higher than control group peers when taught by highly-rated 1st grade teachers. Preliminary analyses of separate student subgroups indicates that the interaction between teacher quality and TN-VPK participation is strongest among Hispanic students, non-native English speakers, and students whose mothers have lower levels of educational attainment.

To illustrate our main preliminary findings we model the relationship between 1st grade teacher ratings and the persistence of pre-K effects graphically by plotting the estimated pre-K effect at three points in time (baseline, end of pre-K and after first grade) for students in classrooms taught by teachers of different quality. Figure 1 shows how the pattern of convergence or fade-out of the initial test score gap for treatment and control groups can vary by teacher quality. As described in Lipsey and colleagues’ 2013 reports, TN-VPK participants and non-participants performed identically at baseline on the composite cognitive measure after controlling for measurable covariates, but by the end of pre-K, treatment students outperformed controls by roughly a third of a standard deviation. On average the difference between TN-VPK participants and non-participants on the Woodcock Johnson composite was statistically negligible by the end of 1st grade. However, this study suggests that when students are taught by a poorly-rated 1st grade teacher, TN-VPK students actually performed significantly worse than controls. Conversely, when taught by highly-rated 1st grade teachers, TN-VPK students still outperformed control group peers by roughly a tenth of a standard deviation.

[Insert Figure 1 Here]

Conclusions:

The results offer some potentially important patterns that merit further exploration. Our findings with respect to the persistence of cognitive benefits seem to contradict those presented in Ruhm and colleagues 2007 study, where more desirable academic settings (smaller class size or high instructional time classrooms) helped non-preschool students catch up. One possible explanation is that the high-rated teachers in our study have adjusted their teaching emphasis to account for higher levels of preparedness due to the expansions of pre-K in recent years. To the extent that
the teacher teaches to the full group, this elevated content emphasis could have negative effects on the control students but benefit students who participated in pre-K.

The differences could also be related to the inadequacy of the prior study’s measure of teacher quality. Spending more time on instruction is not necessarily good teaching if the material one covers is something most the students already know. Recent work by Engel and colleagues (2013) demonstrated a negative association between overemphasis of low level math concepts and student achievement. Similarly, if poorly-rated teachers struggle with classroom management, the disruptive environment could result in a convergence of cognitive measures, as pre-K students fail to build on earlier gains. If the higher-rated 1st grade teachers in our study had stronger classroom management skills or spent more time on challenging concepts and tasks, both of which are factors in the Tennessee observation rubric, it is plausible that such an environment would facilitate greater learning among the pre-K participants than their control group peers.

The positive interaction between TN-VPK participation and teacher quality could also be interpreted as TN-VPK preparing students (particularly those with language deficits) to benefit more from high-quality teachers. Our analysis of the main effects of having a high-rated teacher indicates that the primary benefits were in the categories of letter word recognition, oral comprehension, and quantitative concepts. If TN-VPK helped prepare students who might have otherwise struggled with these types of skills, it is not surprising that we see large positive interactions between teacher quality and pre-K participation in these areas.

The question at the core of this study is the intersection between two of the most prominent aims of contemporary education policy—expanding access to pre-school learning opportunities and access to high-quality teachers for traditionally disadvantaged students. Our findings, though preliminary, are consistent with the theory that either intervention without the other is inefficient if not inadequate. When we invest public dollars in providing pre-school to low income students, the effects fade rapidly in classrooms with inexperienced or poorly rated teachers. Alternatively, students who have experienced a strong pre-school program, tend to benefit more from high quality teachers than they otherwise would. These findings suggest that policymakers interested in maximizing the impacts of preschool initiatives, should work to insure access to high quality early grade teachers. This could be facilitated through policies that discourage school administrators from shifting experienced or effective teachers to later tested grades or the institution of recruitment and retention bonuses for talented early grade teachers in hard-to-staff schools.

As the students from the Tennessee Pre-K experiment progress into tested grade areas, our analysis will expand to estimate cognitive effects for the full sample (approximately 3 times the size and not subject to bias from differential consent rates) and to include teacher test score value-added measures in our construct of teacher quality. Furthermore, the inclusion of additional years of teacher evaluation data will enable more rigorous techniques to account for the potential bias contributed by non-random sorting of students into classrooms, which (along with differential consent rates) poses the most substantial barrier to causal interpretations of the present results. If the findings presented here hold, we can begin to establish a working theory around the ingredients necessary for measures of preschool cognitive gains to persist.
Appendix A. References


Appendix B. Tables and Figures

Figure 1: Pre-K Effect Size Over Time By First Grade Teacher Quality (TEAM Rating)

![Figure 1: Pre-K Effect Size Over Time By First Grade Teacher Quality (TEAM Rating)](image-url)

- Pre-Test
- End of Pre-K
- End of 1st Grade

Effect Size (Standard Deviations)

- Level 5
- ≤ Level 4
- ≤ Level 3
- All