Quantifying and Predicting Variation in the Medium-Term Effects of Oversubscribed Prekindergarten Programs

Rebecca Unterman & Christina Weiland

Background/Context: Research using cutting-edge methods (Bloom, Raudenbush, Weiss, & Porter, 2017; Weiss et al., 2017) has demonstrated that intervention effects can vary substantially across program sites. At the same time, there has been considerable interest in the conditions that can sustain the preschool boost, most notably via the "sustaining environments" hypothesis, which posits that the quality (broadly defined) of children's educational settings after preschool is critical in sustaining the preschool boost (Bailey, Duncan, Odgers, and Yu, 2016).

Purpose/Objective/Research Question: In the present paper, we unify these two new directions in the literature. Specifically, we use new methods for estimating variation in whether the boost from preschool is sustained through third grade on key child academic outcomes to answer: Does the impact of the Boston pre-K program on students' grade retention, special education identification, and third-grade state standardized mathematics and English language arts (ELA) test scores differ across program sites? We then predict this variation as a function of a parsimonious set of site characteristics related to the quality of a student's school experience to answer: Do BPS prekindergarten programs located within higher-quality elementary schools produce different impacts than those located within lower-quality elementary schools? Does the answer depend on how elementary school quality is measured? If so, are there measurable features of the students' school experiences that account for the differences in impacts?

Setting and Population: Our child-level sample comes from the population of students who applied to the Boston prekindergarten program for four-year-olds from which we identified naturally occurring lotteries for students' first-choice school involving 3,182 students, or 25 percent of all appliers and 32 percent of those who applied through the standard process. Our key outcomes are children's K-2 grade retention, K-3 special education placement, and third-grade state test scores. Based on prior literature (i.e., Zhai, Raver, & Jones, 2012), we also use site characteristics as proxies for the quality of a student's school experience – the demand for the school during the assignment process, school-level third-grade ELA and Math Proficiency rates, the school's Median Student Growth Percentile state ranking, the proportion of low-income students, and measures of the school's climate from the district's annual student and teacher school climate survey.

Intervention/Program/Practice: Study data come from the Boston Public Schools (BPS) Prekindergarten program, a relatively large-scale program that is based entirely in the public schools, pays teachers on the same scale as K-12 teachers, subjects teachers to the same educational requirements as K-12 teachers, and is open to any child in the city, regardless of income. The Boston program has shown strong effects on children's language, literacy, mathematics, and executive function skills at kindergarten entry in a

large-scale regression discontinuity study that used the program's long-standing September 1 cutoff as its source of exogeneity (Weiland & Yoshikawa, 2013).

Research Design: As a first step in our analysis of impact variation (and to answer our first research question), we quantified and illustrated the distribution of intent-to-treat (ITT) effects across sites using the framework set forth by Bloom, Raudenbush, Weiss, and Porter (2017) and applied by Weiss and colleagues (2017). In this method, as they suggest, we assumed that our study sites are a sample drawn from a "super population" of prekindergarten sites, and our goal is to generalize to the larger population from which we have drawn. We estimated key statistics for these distributions using a two-level hierarchical linear model and illustrated the distributions using site-level constrained empirical Bayes impact estimates, which, as shown in Bloom and colleagues (2017), constrain the cross-site variance to match that estimated by the model. This is preferable to an empirical Bayes model, which may slightly underestimate cross-site variation.

To estimate whether key site characteristics predict variation in impacts (and answer our second research question), we selected a parsimonious set of prekindergarten program and elementary school characteristics measured at baseline (before school assignment). For each of these site characteristics, we estimated whether the main effect of treatment is moderated — that is, whether it is affected in direction or strength by its values.

Data Collection and Analysis: We addressed our research questions using data made available to the project team by BPS and the Massachusetts Department of Education. We began with data on students' choices and baseline demographics during the BPS assignment process from the spring of the 2006-07 through 2009-10 school years (for enrollment in prekindergarten in 2007-08 through 2010-11). We merged these data, by each student's unique identifier, with district and state administrative records covering the school years students were age-eligible for prekindergarten through third grade. We drew school context data from the state of Massachusetts and school climate data from BPS. We merged these school-level data, by follow-up year, onto each student's file by the school identifier for the school he or she was enrolled in for the longest period of time during that school year.

Findings: A complementary paper that presents the average effects of the BPS program across oversubscribed sites has been accepted by Child Development and is awaiting publication. While we cannot share this paper's findings before that paper is released, at SREE, we will present results of our work quantifying the distribution of intent-to-treat (ITT) effects across sites using the framework set forth by Bloom, Raudenbush, Weiss, and Porter (2017) and applied by Weiss and colleagues (2017) and whether site characteristics predict this variation.

Conclusions and Policy Implications: Solving the convergence puzzle is one of the chief challenges facing the field of early childhood education. Stakeholders' ability to create

conditions in which preschool benefits can last — particularly in large-scale programs — is currently limited by the lack of empirical evidence. In the present paper, we help to advance the science of early childhood education by exploring variation in the medium-term effects of prekindergarten within a unique sample — children who participated in oversubscribed lotteries for the BPS prekindergarten program — in the hopes that the findings may provide much-needed insight and hypothesis generation.