Intensive Professional Development at Scale: Measuring the Casual Effect of Massachusetts' RETELL Initiative

Jesse Bruhn
Assistant Professor
Brown University
jbruhn@princeton.edu

Nathan Jones
Associate Professor
Boston University.
ndjones@bu.edu

Yasuko Kanno
Associate Professor
Boston University
yakanno@bu.edu

Marcus A. Winters*
Associate Professor
Boston University
marcusw@bu.edu

*Corresponding/Presenting Author

Marcus A. Winters, Ph.D.
2 Silber Way
Boston, MA 02215
617-353-3307
marcusw@bu.edu

Background/Context:

Research demonstrating that teacher quality varies substantially and is an important determinant of student outcomes has sparked a variety of reforms and interventions within U.S. public schools. Policies that focus on retaining more effective teachers have shown some promise. But improving teacher quality at scale in a timely manner likely requires affecting teachers currently in classrooms. Professional development (PD) is perhaps the most common and longstanding strategy for improving the quality of in-service teachers. American school systems annually spend about $18 billion on PD despite little evidence that such training improves classroom effectiveness, and no convincing evidence that it can do so at a meaningful scale.

Purpose/Objective/Research Question

We provide the first causal estimates for the impact of an intensive professional development delivered statewide to all teachers working with a critical subpopulation of students, English Learners (ELs). Specifically, we measure the causal effect of a core academic teacher obtaining an endorsement in Sheltered English Immersion (SEI) under Massachusetts' Rethinking Equity in the Teaching of English Language Learners (RETELL) initiative on the performance of ELs and other students they instruct. RETELL is a unique example of a rigorous and time-intensive professional development delivered at a nearly unprecedented scale. Over a five-year period about 35,000 in-service public school teachers throughout Massachusetts completed a training equivalent in scope to a college-level semester-long course.

Setting

Massachusetts provides an especially salient policy setting in which to estimate the effects of this growing type of teacher training requirement. State policymakers initiated RETELL in response to pressure from the U.S. Department of Justice, which argued that so few of the state's teachers had received adequate training to instruct ELs that it was in violation of federal law.

Population/Participants/Subjects

We use longitudinal administrative data for the universe of Massachusetts public school students and their teachers for school years 2010-11 through 2017-18. We combine the administrative data with records of RETELL course completion.

Intervention/Program/Practice

The overall purpose of the SEI endorsement is to help general education teachers develop proficiency in instructional strategies in making academic content accessible to ELs as well as scaffolding their English language development in the context of general education classrooms. The 45-hour SEI Teacher Endorsement course consisted of 12 face-to-face sessions, 9-hour online coursework, and a 2-hour small group capstone presentation.

Research Design

We expect that a naive comparison of outcomes between students with trained and untrained teachers will be biased by two sources of selection. First, the roll-out of the program across the state effectively prioritized teachers in urban and under-performing school districts, where ELs were frequently clustered. This source of selection would lead us to conflate student socio-economic characteristics with the impact of RETELL training. Second, we worry that after teachers have received training, administrators within schools may endogenously sort students to trained teachers on the basis of ability.

We leverage cross-teacher variation in the timing of training via a generalized difference-in-difference strategy. Intuitively, we would like to compare the classroom level trend in test score gains among teachers who have received the training to those that have not. Because the comparison across teachers is via trends, our identification strategy addresses the first source of selection by effectively differencing out variation in the socio-economic and demographic composition of students across schools and districts. To address the second source of selection, we hold constant test scores from prior years and hence focus our attention on trends in test-score gains. We ask, ``Do average test score gains increase suddenly in the classrooms of teachers that receive the training relative to the classrooms of teachers that do not?" Thus, compositional changes emerging from the endogenous sorting of students to teachers are effectively controlled for by differencing out average student ability at the classroom level.

Causal interpretation of our estimates hinges on the assumption that there are no time-variant factors that are associated with both the timing of the teacher completing the training and the outcomes of students in her class. We test the plausibility that this assumption holds by conducting an event-study analysis that measures changes in teacher effects in the years leading up to and following the training.

Findings/Results

For ELs in a teacher's classroom we find that completing the full training leads to modest improvements in math and ELA scores as well as their progress toward proficiency in English. The training also has positive spillovers on the average math and ELA test scores of students with disabilities and the larger group of students who have never been classified as an EL. However, estimates for the average effect of the training mask important heterogeneity at both the student and teacher level. The positive effects for ELs are confined to students in later grades, with ELs in elementary grades actually experiencing less growth in math if their teacher has received the training. We find some evidence that the training had a more positive effect for early-career teachers. Finally, we find that the training was most positive, and indeed of a meaningful magnitude in some cases, for students who the prior year had relatively high math or ELA test scores, while it tended to have a null or even negative effect on students with lower prior proficiency in the subject.

Conclusions

On the one hand, the rapid and substantial expansion of teachers in the state who had received training in SEI instructional techniques led to only modest improvements in the outcomes for ELs, on average. On the other hand, the existence of spillovers for other students and the

heterogeneous effects from teachers completing the training on the outcomes of both ELs and non-ELs suggests that the training did elicit improvements in teacher effectiveness.