Prospective Matching Methods in Education Research:
Recruiting an Active Comparison Sample for Causal Inference

Richard Correnti
Ally Thomas
Baeksan Yu
Jennifer Russell
University of Pittsburgh

Laura Booker
Nate Schwartz
Tennessee Department of Education

Mary Kay Stein
University of Pittsburgh

Background
In recent years there has been an expansion of research methodologies for generating causal inferences in the social and behavioral sciences (Rubin, 2006; Stuart, 2010). Much of which has utilized propensity score techniques to match participants on a host of observable characteristics in a post-hoc design (Guo & Fraser, 2010). A more under-developed method in education research, although used in other fields (e.g., medicine), is using prospective matching techniques to identify and actively recruit matched control participants (Song & Chung, 2010).

In this paper, we present our prospective matching technique currently being used to actively recruit participants for an IES supported continuous improvement coaching study in the state of Tennessee. Furthermore, we address the unique barriers for carrying out such large-scale studies in education, as well as the benefits for addressing causal research questions.

Purpose
After 2 years of designing, developing, and studying our mathematics coaching model with the Tennessee Department of Education (TDOE), we will investigate how our coaching model differs from naturally-occurring mathematics coaching in the state of TN. Specifically, we will compare and contrast a population of partner teachers coached through our coaching model \(n = 167\) with matched active comparison teachers receiving garden-variety coaching \(n = 100\) on the coaching process, as well as outcomes on teaching and learning.

Data
Prospective matching would have been more difficult without our partners at the TDOE providing access to several data files including: (1) state-wide 4th through 8th grade student achievement and demographics; (2) teacher value-added scores in mathematics, demographics, and responses to the TN educator survey; (3) school demographics; and (4) district demographics.

Potential active comparison teachers had to meet several criteria: (1) a 4th-8th grade math teacher; (2) linked in the data files to at least 10 students; (3) in a district that reported at least 1 coach serving the district or in a district served by a coach participating in our coaching model. This resulted in a sample of 7,887 unique teachers of mathematics students to match on.
Methodology

The process we are using to identify active comparisons includes: 1) identifying essential covariates at the classroom, teacher, school and district level; 2) conducting multiple imputation; 3) generating propensity scores; 4) checking for initial balance; 5) recruiting an active comparison sample; and 6) checking balance of our partner teachers and realized active comparisons.

 Identify covariates and impute missing values. We identified and developed covariates for 7,887 math teachers in districts across TN where there existed a potential for math coaching including, for example, teacher value-added scores and teacher responses about their Math instruction on a state-wide survey. We then performed a multi-level imputation model generating a total of five imputed datasets. After performing the multiple imputations, we engaged in data reduction activities (e.g., we generated factor scores of self-reports of mathematics teaching from items on the TN Educator survey) to form the final set of covariates used in the propensity score model.

Generate propensity score. Using the finalized set of 55 covariates, we generated propensity score estimates for each of the five imputed samples. To generate a propensity score for matching, we pooled estimates from the 5 data sets by calculating the average propensity score for each teacher (Hill, 2004; Mitra & Reiter, 2016). Because not all of our coaches identified their partner teachers by the beginning of September, we began by generating our propensity score and matches for wave 1 partner teachers (n = 123)\(^1\). We repeated this process for wave 2 (n = 20) late-identified teachers who appear in the master file of 7,887 teachers and wave 3 (n = 31) late-identified teachers who do not appear in that file.

In order to generate a large recruitment sample we utilized many-to-one matching with replacement and set a relatively conservative caliper. For each of our 123 treated teachers we first identified 3 matched comparisons (n = 372 matched controls) then later identified an additional 2 matched comparisons. When identifying matches, some control teachers were matched to more than one treated teacher, thus our first attempt yielded a total of 350 matched control teachers to recruit, and 314 of them had a recent email we could use to contact them.

Initial balance check. We examined whether there were any significant differences between our group of treated teachers and the 314 active comparison teachers we reached out to invite.

Recruiting control participants. Our initial email invite informs teachers about the study along with a link to a pre-survey to make sure they qualify for the study because they, 1) still teach 4\(^{th}\) through 8\(^{th}\) grade math, and 2) think they might receive coaching this upcoming year. Engaging in this process we have confronted barriers such as; invalid email addresses and/or firewalls, the need for our partners to send the initial invite so the offer has more legitimacy, many teachers switching to ineligible grades, many interested teachers not being coached, and some uninvited teachers attempting to enter the sample. We describe these and other issues we have confronted in the recruitment process.

Results

Within the first week, we had 20 responses of which 11 met study qualifications. Currently, we have over 100 teachers who have responded to the pre-survey and qualify for the study – 59 of them have officially enrolled in the study. We will report on our monitoring of the distribution and balance of our attained sample, as well as the balance obtained with our achieved sample.

\(^1\) Over the course of this month 8 of these teachers have left the treated sample for various reasons.
Conclusion

Prospective matching is an infrequent practice in large-scale education research. Our experience suggests it is fraught with barriers even when a trusting partnership with state leaders has been established providing unfettered access to data. Despite these troubles, we think it can help us answer important research questions allowing us to identify causal effects of our coaching model as well as causal explanations about how our coaching model differs from garden-variety coaching across the state. Combined, these analyses could provide much-needed guidance for mathematics professional learning.

References


