Title:

Prognostic Score-Based Difference-in-Differences Strategy for Education Policy Evaluation

Justification

Under the No Child Left Behind Act, many states and school districts have required standardized testing in selected grade levels. Such repeated cross-sectional cohort data may provide a unique opportunity for education policy evaluation. Yet a major concern with regard to the causal validity is that a policy of interest may be introduced in the midst of other contextual changes. Hence the policy effect will likely be confounded by the impact of other concurrent events. Difference-in-differences (DID) methods resort to a non-equivalent comparison group unaffected by the policy that is often located in the same jurisdiction or adjacent to the experimental group affected by the policy. The average outcome change in the comparison group is solely attributable to the confounding effect of the concurrent events. Subtracting the average outcome change in the comparison group from that in the experimental group produces an estimate of the policy effect of interest. Yet the conventional DID strategy relies on the strong assumption that the average confounding effect is the same for the comparison group and the experimental group. This assumption will be violated and therefore the DID results will be biased, for example, if the confounding effect varies by individual characteristics and if the experimental group and the comparison group differ in such characteristics.

This symposium includes three papers that systematically introduce, evaluate, and illustrate a new DID strategy that offers an alternative to the existing DID methods. The new strategy calibrates the adjustment for confounding on the basis of a pair of prognostic scores representing the predicted pre-policy outcome and the predicted post-policy outcome of a unit if unaffected by the policy. The difference between the two prognostic scores estimates the predicted amount of confounding attributable to concurrent events for the individual unit. Pooling the results from DID analyses within subclasses defined by this pair of prognostic scores, one obtains an unbiased estimate of the policy effect under identification assumptions that are distinct from and often weaker than those invoked by the existing DID methods.

The first paper derives the theoretical results for the prognostic score-based DID strategy, clarifies its identification assumptions, and develops an analytic procedure. This new strategy is then extended to multilevel, multi-cohort education accountability data. The second paper reviews the existing alternative DID strategies, compares their identification assumptions with those of the prognostic score-based DID strategy, and evaluates the relative strengths and limitations of the new strategy through a series of simulations. The third paper applies the prognostic score-based DID strategy to an evaluation of a policy adopted by the Chicago Public Schools requiring all ninth graders to take algebra. Through illustrating the implementation of this new strategy to multilevel multi-cohort data typically seen in education accountability systems, the application study reveals issues that arise in practice.