Leveraging Lotteries for School Value-Added: Testing and Estimation

Joshua Angrist, Peter Hull, Parag Pathak, and Christopher Walters

Public school districts increasingly use value-added models (VAMs) to assess teacher and school effectiveness. Conventional VAMs compare test scores across classrooms or schools after regression-adjusting for students’ demographic characteristics and earlier scores, with remaining achievement differences attributed to differences in teacher or school quality. Some districts use teacher value-added to guide personnel decisions, while others use VAMs to generate “report cards” that allow parents to compare schools. Value-added estimation is a high-stakes statistical exercise: low VAM estimates can lead to school closures and teacher dismissals, while a growing body of evidence suggests the near-term achievement gains produced by effective teachers and schools persist throughout adulthood (see, e.g., Chetty et al. [2011] and Chetty, Friedman and Rockoff [2014b] for teachers and Angrist et al. [2016a] and Dobbie and Fryer [2015] for schools).

Because the stakes are high, the use of VAMs for assessment remains controversial. Critics note that VAM estimates may be misleading if the available control variables are inadequate to ensure ceteris paribus comparisons. VAM estimates may also reflect considerable sampling error. The accuracy of teacher value-added models is the focus of a large and expanding body of research. This work demonstrates that teacher VAM estimates have predictive value, but has yet to generate a consensus on the substantive importance of bias or guidelines for “best practice” VAM estimation (Kane and Staiger 2008; Rothstein 2010; Kane et al. 2013; Chetty, Friedman and Rockoff 2014a; Chetty, Friedman and Rockoff 2014b; Rothstein 2016). While the social significance of school-level VAMs is similar to that of teacher VAMs, their validation has received less attention.

The proliferation of partially-randomized urban school assignment systems provides a new tool for measuring school value-added. Centralized assignment mechanisms based on the theory of market design, including those used in Boston, Chicago, Denver, New Orleans, and New York, use information on student preferences and school priorities to allocate scarce admission offers. These matching algorithms typically employ random sequence numbers to distinguish between students with the same priorities, creating stratified assignment lotteries. Similarly, independently-run charter schools often use admissions lotteries when oversubscribed. Scholars increasingly use these lotteries to identify causal effects of enrollment in various school sectors, including charter schools, pilot schools, small high schools, and magnet schools (Cullen, Jacob and Levitt 2006; Hastings and Weinstein 2008; Abdulkadiroğlu et al. 2011; Angrist, Pathak and Walters 2013; Bloom and Unterman 2014; Deming et al. 2014). Lottery-based estimation

---

of individual school value-added is less common, however, as lottery samples for many schools are small, while others are under-subscribed.

This paper develops econometric methods to leverage school admissions lotteries for VAM testing and estimation, accounting for the partial coverage of lottery data. Our first contribution is the formulation of a new lottery-based test of conventional VAMs. This test builds on recent quasi-experimental validation strategies, including the work of Kane and Staiger (2008), Deutsch (2013), Kane et al. (2013), Chetty, Friedman and Rockoff (2014a) and Deming (2014). In contrast with earlier studies, which implicitly look at average-across-schools validity in a test with one degree of freedom, ours is an over-identification test that looks at each of the orthogonality restrictions generated by a set of lottery instruments. Intuitively, our test asks whether conventional VAMs correctly predict the effect of randomized admission at every school that has a lottery.

Application of this test to data from Boston reveals moderate but statistically significant bias in conventional VAM estimates. This finding notwithstanding, conventional VAM estimates may nevertheless provide a useful guide to school quality if the degree of bias is modest. To assess the practical value of VAM, we develop and estimate a hierarchical random coefficients model describing the joint distribution of value-added, VAM bias, and lottery compliance across schools. The model is estimated via a simulated minimum distance procedure that matches moments of the distribution of conventional VAM estimates, lottery reduced forms, and first stages to those predicted by the random coefficients structure. Estimates of the model indicate substantial variation in both causal value-added and selection bias across schools. The estimated joint distribution of these parameters implies that conventional VAM estimates are highly correlated with school effectiveness.

A second contribution of our study is to use the random coefficients framework and lottery variation to improve conventional VAM estimates. Our approach builds on previous estimation strategies that trade reduced bias for increased variance (Morris 1983; Judge and Mittelhammer 2004). Specifically, we compute empirical Bayes hybrid posterior predictions that optimally combine relatively imprecise but unbiased lottery-based estimates with biased but relatively precise conventional VAM estimates. Importantly, our approach makes efficient use of available lottery information without requiring a lottery for every school. Hybrid estimates for under-subscribed schools are improved by information on the distribution of bias. The hybrid estimation procedure generates estimates that, while still biased, have lower mean squared error than conventional VAM estimates. Our framework provides a general recipe for combining non-experimental and quasi-experimental estimators and may therefore be useful in other settings.

Finally, we quantify the consequences of bias in conventional VAM estimates and the pay-off to hybrid estimation using a Monte Carlo simulation calibrated to our Boston estimates. Simulation results show that policy decisions based on conventional estimates are likely to boost achievement. For example, replacing the lowest-ranked Boston school with an average school is predicted to generate a gain of 0.24 test score standard deviations (σ) for affected students, roughly two-thirds of the benefit obtained when true value-added is used to rank schools (0.37σ). Hybrid estimates are highly correlated with conventional estimates (the rank correlation is 0.74), and hybrid estimation generates modest additional gains, reducing mean squared error by 30 percent and increasing the benefits of school closure policies by 0.08σ (33 percent). Conventional school VAMs therefore appear to provide a useful guide for policy-makers, while hybrid estimators generate worthwhile improvements in policy targeting.
References


