Comments on "Grading" Schools with RD
by Abdulkadiroglu, Angrist, Narita, and Pathak

Comments by;
Ariel Pakes, Harvard University

January 3, 2019

Substantive Goal: Analyze how educational outcomes vary with
the school a student attends.

- Current paper: correct for selection in OLS regressions of
  educational outcome variables on school attended.
- Econometrics: generalization of prior techniques that
  - look at the lowest priority accepted (or the marginal group) at
    oversubscribed schools (e.g. charters, or exam schools) .
  - Among the marginal group, students are accepted by
    randomization device.
  - Compare outcomes for those who are accepted to those who
    are not.
- Generalization: develop techniques that allow one to control
  for selection among marginal priorities when a DA algorithm is
  used to allocate students among schools.

DA allocations

- Consider the simple case of schools with no "running variable", and large market asymptotics.
- Each student is assigned a priority and a randomization score at each school ($\pi_{i,s} = \rho_{i,s} + R_{i,s}$).
- Priority ordering is lexicographic. If the school fills its seats, there is a marginal priority within which randomization selects.
- DA allocation insures that if a student is allocated to oversubscribed school "s" then the student's priority must be at least as high as the cutoff priority at "s" and the priority at all schools "b" that the student ranks higher than "s" (i.e. $b \succ_i s$) must be lower than the cutoffs for those schools, or

$$\pi_{i,s} \geq \xi_s, \ \& \ \pi_{i,b} \leq \xi_b, \ \text{if} \ b \succ_i s.$$

NYC High Schools.

- NYC assigns grades to schools (A to F)
- Look at students who applied to A schools and were in the marginal priority group in those schools.
- Compare OLS and IV regressions for educational outcome variables on school attended for those students.
- Instruments for school attended in IV results are the random number the student obtains.
- I have greatly simplified; econometrically unfair to them as their techniques enable you to get more precise estimates then if you followed the above procedure literally; but I don't want to focus on the econometrics.

Results.

- Outcome variables: SAT Math score, SAT Reading score, and whether they graduated (they also have College and Career prepared, and College-ready; but they never explain what they are so I am going to ignore them).
- OLS vs IV: OLS substantially higher; factor of
  - 3 for SAT Math,
  - 7.5 for SAT reading, and
  - 1.25 for graduating.
- Substantive results; point estimates & statistical significance.
  - Math: marginally significant, raise scores by $\approx .4\%$.
  - Reading: insignificant, raise scores by $\approx .1\%$.
  - Graduation: sometimes significant, raises rates $\approx 4\%$.

What have we learned?

Accepting all assumptions we learned that conditional on the

- taste distribution of students applying to public schools, &
- the distribution of the facilities offered by those schools,

among students who both

- applied to type A schools, and
- had marginal priority at type A schools

those who were randomized into the type A school did about .4% better in Math SAT scores, and may have had 4% better graduation rates then those randomized out.

We do not know the characteristics of the students who did better.

These numbers should be interpreted carefully.

- If we either
  - changed the distribution of school characteristics (e.g.'s; increase a particular schools capacity, or change the facilities or the priority definitions at a school), or
  - changed the taste distribution of the applicants (e.g.'s the unemployment rate falls and a parents who used to be at home after school are no longer home and want to send their children to schools near relatives, or there were changes in the composition of a neighborhood.)

the estimates provided here would no longer be relevant, as the marginal student body at all schools would change.

- Of course small changes in the environment might not change the results much. However no metric is provided to tell us what might be considered "small".

Can we go further?

• We could analyze a model that determined what caused improvements in scores or graduation rates. This would allow us to analyze what would happen if either policy or exogenous factors changed the environment, but would require:

  1. A model for the preferences of the applicants as a function of their characteristics.
  2. A model for how students with different characteristics are allocated to different schools.
  3. A model for how those characteristics interact with the facilities of the school to produce educational outcomes.

• Though this may be viewed as a demanding research program, as I now explain, the institutional environment and the data available for analyzing educational attainment are as close to perfect as we get in analyzing different allocation mechanisms.

The data on, and the analysis of, preferences.

- The data on preferences is as good as preference data gets.
- The individual submits an ordered list of preferences and knows that the list submitted determines the allocation. Assuming they understand the incentive compatibility constraint in DA the list can be used to estimate a rich preference ordering that allows for important unobserved choice determining characteristics
    - Take the prior case with a change in employment status; not only would this induce a change in location of the first choice school, but also the second, the third.... Multiple ordered choices yields an ability to account for characteristics which are not observed (MicroBLP, 2004).
- Indeed this has already been done by two of this papers' coauthors and Agarwal in a paper in the *AER* (2017); It was used to generate a revealing & important analysis of the benefits of different allocation mechanisms.

The allocation mechanism.

• Most economic studies of the impact of policy or environmental change on a market allocation rely on an equilibrium assumption of some form. None of us believe that after any change we instantaneously go to a new equilibrium. We treat the equilibrium as a rest point and believe it generates a good approximation.

• Here we know the exact form of the allocation mechanism, and, as shown in the AER paper, can mimic it when either the allocative rules (e.g. the priorities) or the tastes of participants change; i.e. there is no approximation error. This enables both

  • a clean analysis of counterfactuals and

  • a determination of which changes are indeed "small" or large.

The transformation function from school and student characteristics to outcomes.

• An education economist would be needed to know the data on school characteristics (capacity, facilities, teachers, etc.) that are available to combine with student characteristics in order to estimate the transformation function needed to analyze the distribution of student outcomes generated by any allocation.

• Estimation problems might arise, but by now there are a large number of techniques available to mitigate them.

• Notice that this kind of analysis would not only tell us how many students would be helped, but also which students would be helped, were we to change school characteristics.