Angrist et al. on KIPP Lynn (MA)

KIPP schools are aimed at low-income students; KIPPs have strict discipline, encourage student work ethic.
The KIPP students perform at better-than-average levels, but why?
   Does KIPP simply attract the better students in the first place?
   Or does KIPP actually make students better?

Research strategy: exploit the fact that, during 2005-2008, KIPP applications were “oversubscribed” – so many kids wanted to go to a KIPP that the state had to hold a lottery to determine who would actually be allowed to attend the KIPP.
Being a “lottery winner” is a *random variable*, so it can’t be correlated with unobservable characteristics (= the error term e) such as motivation, IQ, or anything else that we can’t explicitly measure –

*but*, being a winner is certainly correlated with attending a KIPP (X)!

In other words, the lottery gives us the equivalent of a random experiment:

   winners = experimentals, losers = controls.

So, being a lottery winner is correlated with attending a KIPP, but (since it’s random) can’t possibly be correlated with anything else (such as being a good or motivated student).

Thus, *we break the connection between X and e*!
Three equations:

the “equation of interest”:

(1) \( Y = \alpha + \rho S + \text{other variables} + e \)

where \( Y \) = achievement test score, \( S \) = years in KIPP

Can’t estimate this directly since \( S \) and \( e \) are likely to be correlated (kids who are more motivated/smarter/harder working etc. may end up in KIPP, spending more years there) -- but we can do so *indirectly*, using two-stage least squares:

the “first-stage equation”: gives determinants of \( S \)
(a lottery win, \( Z \), and other variables determine \( S \))

(2) \( S = \mu + \pi Z + \text{other variables} + u \)

Now substitute (2) into (1), to get the *reduced-form equation*:

\[ Y = \alpha + \rho [\mu + \pi Z + \text{other variables} + u] + e, \]  or

(3) \[ Y = [\alpha + \rho \mu] + \rho \pi Z + \text{other variables} + [\rho u + e] \]
So now we have

(1) \( Y = \alpha + \rho S + \text{other variables} + e \)
(2) \( S = \mu + \pi Z + \text{other variables} + u \)
(3) \( Y = [\alpha + \rho \mu] + \rho \pi Z + \text{other variables} + [\rho u + e] \)

We can estimate (2) since \( Z \) and \( u \) are uncorrelated. We can estimate (3) since \( Z \) and the composite error term \([\rho u + e]\) are uncorrelated. With these results, divide the coefficient on \( Z \) in (3), = \( \rho \pi \), by the coefficient on \( Z \) in (2), = \( \pi \), to get the coefficient on \( S \), namely \( \rho \), in the “equation of interest,” (1).
See results in Table 2, p. 242:

results for $Y = \text{math score}$:

<table>
<thead>
<tr>
<th>variables included in the regression:</th>
<th>$\pi$</th>
<th>$\rho_\pi$</th>
<th>$\rho_\pi/\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>basic variables</td>
<td>1.222</td>
<td>0.431</td>
<td>0.353</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.116)</td>
<td>(0.095)</td>
</tr>
<tr>
<td>basic variables + baseline test scores, demographic variables</td>
<td>1.228</td>
<td>0.425</td>
<td>0.346</td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td>(0.066)</td>
<td>(0.052)</td>
</tr>
</tbody>
</table>

(Can you interpret the results for ELA (English proficiency) in the same way?)
One other result (p. 243): Testing for interaction of “lottery winner” × prior test score T implies that “KIPP Lynn raises achievement more for weaker students” (!) – the coefficient on Z × prior test score T is negative.

In effect, the equation “with interactions” is as follows:

\[ Y = a + \rho S + \beta [S \times T] + \lambda T + \nu \]

This can also be estimated using 2SLS. Here, a negative \( \beta \) means that the lower the prior test score T is, the higher the test score after KIPP (Y).
OLS is NOT terribly different from 2SLS (Table 2, p. 242, and unpublished OLS results):

<table>
<thead>
<tr>
<th>variables included in the regression:</th>
<th>$\pi$</th>
<th>$\rho\pi$</th>
<th>$\rho\pi/\pi$</th>
<th>2SLS</th>
<th>OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“first stage” (equation (2))</td>
<td>reduced form (equation (3))</td>
<td>(equation (1))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>basic variables</td>
<td>1.222</td>
<td>0.431</td>
<td>0.353</td>
<td>0.304</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.116)</td>
<td>(0.095)</td>
<td>(0.048)</td>
<td></td>
</tr>
<tr>
<td>basic variables + baseline test scores, demographic variables</td>
<td>1.228</td>
<td>0.425</td>
<td>0.346</td>
<td>0.317</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td>(0.066)</td>
<td>(0.052)</td>
<td>(0.032)</td>
<td></td>
</tr>
</tbody>
</table>