

Assignment 2

Numerical Simulation

Yidan Sun

March 9, 2016

part 1

1

The authors address several research questions, also because they are discussing two separate experiments. I have identified the following questions, listed from most general to most specific:

1. Can providing inputs that address specific unmet needs in schools improve schooling quality?
2. What are the effects of a remedial education program, in which a 'balsakhi' works on basic skills with the children, on schooling quality?
3. What are the effects of a computer-assisted learning program, where children in grade four are offered two hours of shared computer time per week, on schooling quality?
4. How cost-effective are both of the programs?

The authors conduct two experiments that provide additional 'inputs' that specifically target the weaker students. Clearly, it is important to know how cost-effective these programs are, especially considering how much cheaper they are as compared to hiring additional regular teachers. Since the hypothesis is that targeting weaker students might be beneficial, it seems natural to test whether weaker students indeed gain more.

2

Now, for each individual student (analysis at the level of the school is also possible) there are two potential outcomes:

- y_i^1 is the outcome when treated
- y_i^0 is the outcome when non treated

Where y_i is a measure of learning outcomes, i.e. test score or improvement in test score over the year. Note the counterfactual nature of potential outcomes. Ideally, we would want to compare the learning outcome of a student if he received the treatment with his learning outcome had he not received the treatment. However, since we cannot turn back time this is impossible.

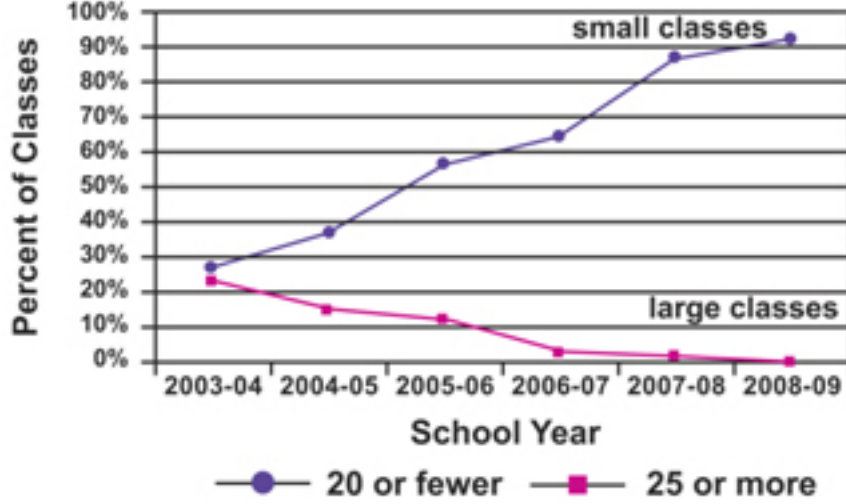
More practically speaking, they use what they call a value added specification:

$$y_{igj}POST - y_{igj}PRE = \lambda + \delta D_{jg} + \theta y_{igj}PRE + \epsilon_{igj}POST \quad (1)$$

That is, change in the test score of child i in grade g and school j is regressed on treatment assignment D_{jg} and the respective pretest score. This specification asks whether children that received the treatment increased their test score more, relative to their pretest score, than children that did not receive the treatment.

Secondly, they use a difference-in-differences specification (**only reported in working paper version**):

$$y_{igjk} = \lambda + \delta D_{jg} + \theta POST_k + \gamma(D_{ig} * POST_k) + \epsilon_{igjk} \quad (2)$$



Where k indexes the individual tests (pre and post potential treatment). Here the interaction term with parameter γ is the treatment effect; it is how much more test scores increased for individuals that received the treatment compared to individuals that did not receive the treatment.

4

Estimating equations (1) and (2) generates estimates of the average impact of the program on all children whose school-grade received a balsakhi. However, there might also be an indirect effect on children that did not get help from the balsakhi, as they will share the classroom with fewer students. This indirect channel could potentially work in two ways:

- A class size effect
- A peer quality effect

Since teachers were not prepared to assign the children to potential assignment a priori, without knowing whether or not they were going to get a balsakhi, there might be some selection effect issues. To address this the authors do the following:

The authors start by predicting assignment probability in the treatment schools as a flexible function of the pretest score.

$$P_{ijg} = (\pi_0 + \pi_1 y_{ijgPRE} + \pi_2 y_{ijgPRE}^2 + \pi_3 y_{ijgPRE}^3 + \pi_4 y_{ijgPRE}^4) * D_{jg} + \omega_{ijg} \quad (3)$$

Let M_{ij} be the vector $[1, y_{ijgPRE}, y_{ijgPRE}^2, y_{ijgPRE}^3, y_{ijgPRE}^4]$. Where P_{ijg} is a dummy indicating whether a child worked with the balsakhi.

Consider the following structural equation:

$$y_{ijgPOST} - y_{ijgPRE} = \gamma D_{jg} + \tau P_{ijg} + M_{ij}\alpha + \epsilon_{ijg} \quad (4)$$

Here γ is the indirect effect, while τ is the direct effect of having attended the remedial classes on top of being in a school that was assigned to the treatment group.

The authors use the predicted assignment probability as an instrument for actual assignment probability and then run an IV regression to estimate the parameters of equation (4).

The authors conclude that the effect of the program is concentrated on children that were indeed assigned to the balsakhi; they cannot reject the hypothesis that being in a balsakhi school had no effect for children who were themselves not sent to a balsakhi.

part 2

a

The unconfoundedness assumption is also called the *conditional independence assumption*. Basically, it means that selection into treatment (treatment assignment) is based on observable variables only. Formally:

$$y_i^0, y_i^1 \perp d_i | X_i \quad (5)$$

That is, potential outcomes should be independent of the treatment status, conditional on observed covariates X_i . Put differently, it means that the ones that get assigned treatment are a random sample of the population conditional on X . This implies the following:

$$E(y_i^0 | d_i = 1, X_i) = E(y_i^0 | d_i = 0, X_i) = E(y_i^0 | X_i) \quad (6)$$

$$E(y_i^1 | d_i = 1, X_i) = E(y_i^1 | d_i = 0, X_i) = E(y_i^1 | X_i) \quad (7)$$

To use a real world example, consider the STAR experiment discussed in class. There, students were randomly assigned to different class sizes, however the randomization was done within each school. Hence, treatment assignment was random conditional in each subset of the total population, i.e. within each school, but not necessarily across schools.

b

$$y_i^0 = \beta + X_i' \gamma + u_i \quad (8)$$

$$E(u_i) = E(y_i^0) - \beta - E(X_i' \gamma) \quad (9)$$

$$E(u_i | d_i, X_i) = E(y_i^0 | d_i, X_i) - \beta - X_i' \gamma \quad (10)$$

$$= E(y_i^0 | X_i) - \beta - X_i' \gamma \quad (11)$$

$$= \beta + X_i' \gamma + E(u_i | X_i) - \beta - X_i' \gamma \quad (12)$$

$$= E(u_i | X_i) = E(u_i) \quad (13)$$

Where (10) uses the fact that the expectation of X_i' conditional on X_i, d_i is simply X_i' , (11) uses the fact that $y_i^0, y_i^1 \perp d_i | X_i$, (12) uses equation (8) again and (13) uses the fact that $u_i \perp X_i$.

c

$$\alpha_i = y_i^1 - y_i^0 \quad (14)$$

$$E(\alpha_i | d_i, X_i) = E(y_i^1 - y_i^0 | d_i, X_i) \quad (15)$$

$$= E(y_i^1 - y_i^0 | X_i) \quad (16)$$

$$E(\alpha_i | d_i, X_i) = E(\alpha_i | X_i) \quad (17)$$

Here (14) simply uses the definition of α_i , (16) uses the unconfoundedness assumption and (17) combines (15) and (16), yielding the result. Hence, if the unconfoundedness assumption holds, no matter if the treatment effect is random or constant, it is uncorrelated to treatment assignment, conditional on the observed covariates X_i .

α_i might for example be the effect of a certain medicine. Perhaps we expect that this medicine might have different effects depending on age. Hence, we could construct several age groups, and then within each age group randomly assign treatment and placebos. α_i might be random because different people might react differently to a medicine, although on average it might have a certain effect. In such a setting, treatment assignment is random conditional on the age group one is in. One could then for example estimate the ATE by running a linear regression.