14.11 Lecture 6: The quality of education in developing countries

Prof. Esther Duflo

March 21, 2006

Until relatively recently, the main question policy makers were trying to solve about education in developing countries was: how to get children in school.

Access is not so much of a problem anymore, although many children remain out of school. The next important problem is that of the quality of education: the quality of education in developing countries is poor.

Example:

-In India, a nationwide survey found that 65% of the children enrolled in grades 2 to 5 in government primary schools could not read a simple paragraph, and 50% could not do simple subtraction or division (Pratham, 2006).

-Teacher absence rates in schools are high, with higher absence in poorer countries and regions (see table 1 and graph, taken from nationally representative surveys in several countries (Chaudhury, Kremer, and others)).

-Children's absence rates are high too, even when children are actually enrolled (suggesting that a drive to increase enrollment may not be sufficient!).

One could think of various levers to improve school quality. What enters into the education "production function," and what could be done about it?

-

-

-

The literature on field experiments in education in developing countries is very rich (this is probably the richest field), and many things have been tried.

Today, in order to draw on what we've learned from previous lectures, we will focus on the question of how to provide incentives to the agents in the system (children, teachers, and parents), and their impact. Now that we are reasonably familiar with reading the results of experiments, we will go through the details of one, and more quickly through the results of several others, so we obtain an overall coherent picture.

1   Incentives for parents and children

1.1   PROGRESA

Probably the best known program of providing incentives to parents to send their children to school is the PROGRESA program in Mexico. In part, it is so well known *because* it had a randomized evaluation attached to it.

The program is a *conditional cash transfer* program. This means that it is above all a cash transfer program for the poor (it is means tested), but to get it, you need to satisfy a number of conditions: in particular, kids need to be enrolled in school and attending for at least 85% of the time, and the family also needs to obtain preventative health care.

The history of the evaluation is interesting: the previous PRI government introduced it, and was worried about losing power (and that the new government would ditch the program). They therefore wanted to provide very rigorous evidence that it worked, and to this end, they set up a randomized pilot program in 495 villages (half the villages got the full program, and half did not). They invited several well-known scholars to study the program, they made the data public, they encouraged wide scrutiny.

Research showed that the program was indeed effective in improving health and education outcomes. The PRI lost power. The program continued, was expanded to rural areas, and spread like wild fire all over Latin America!

See tables and graphs for the PROGRESA fee structures and impact. Why is the effect larger for secondary school?

Can we interpret the PROGRESA results in a simple way? Why or why not?

A problem with PROGRESA as an incentive program is that the incentive is based on the teachers' reports of absence. This is easy to manipulate (and evidence from India suggests that it is). If all the children are already enrolled in primary school, and you get no gain in

primary school enrollment because the enrollment figures are doctored, it is not clear that this is a terribly cost-effective way to get children in school more.

Two alternatives: cooked school meals (which you get ONLY if you come!) and incentives based on results. Both have been evaluated by Michael Kremer (from Harvard) and associates in Kenya. Angrist and Lavy did another evaluation in Israel.

1.2 Incentives to learn: Kenya and Israel

Study effort (presence in school, paying attention, and studying at home) is arguably an important component of the quality of education. How to encourage effort at home?

Merit-based scholarships used to be an important component of educational policy. They went away in the 60s and 70s, but they are coming back. They are essentially absent in developing countries.

From our last lectures, or your intuition, what are the pros and cons of financial incentives for studying?

PROS:

-

-

-

CONS:

-

-

-

Kremer, Miguel, and Thornton: Incentives to Learn

Experimental design:

-127 primary schools (K to 8), half in treatment, half in controls

-Scholarship for girls enrolled in class 6 at beginning of school year; winners decided at the end of the school year; would provide for school fees for class 7 and 8, as well as $12 a year for school supplies and public recognition in awards ceremony.

-Winners were awarded for the top 15% of the girls in the treatment schools in the girls' districts (about 100 winners each year).

The questions we are particularly interested in studying regarding this program:

-

-

-

Results:

-Test scores: average, externality, inequality.

-Attendance

What do we conclude?

Methodological aside: Computing the right standard errors (1): clustering

Almost all of the interventions we look at in this lecture use child-level interventions, but we are looking at interventions that are common to all children in a school (for example, all girls were eligible for the scholarship, and all boys were potentially affected by the spillovers).

Do we have reason to think that the test scores of students would be correlated across schools?

-

-

This affects our standard errors. Little reminder from econometrics:

$$y = X\beta + \epsilon$$

$$E[\epsilon|y] = 0$$

$$\hat{\beta} = (X'X)^{-1}X'y$$

Let's show that:

$$VAR\hat{\beta} = E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'] = \sigma^2(X'X)^{-1}$$

if ϵ is iid, with correlation within school it is not *iid any more.*

$$E[\epsilon\epsilon'] = \Omega$$

where Ω is block diagonal, with the intra-school correlation in each block.

So the right standard error is:

$$VAR\hat{\beta} = (X'X)^{-1}X'\Omega X(X'X)^{-1}$$

Which will typically be larger or smaller?

Ways to compute and estimate for Ω:

-Estimate the intra-school correlation and plug it into Omega (Moulton).

-Estimate generalized "White" covariance matrix, where you compute $E[\epsilon_g \epsilon_g']$ within each school (this is easily implemented in Stata using the "cluster" option).

Examples where it will be relevant in your research:

-if your treatment varies at the "session level" (and you collect and analyze individual data).

-if your treatment varies at the "recitation level."

This is important to keep in mind when computing the power of a study: obviously, the power will greatly depend on whether the outcomes are correlated or not. The Optimal Design software allows you to take this into account.

2 Incentives for teachers

2.1 Incentives based on presence

Now let's turn to incentives for teachers. Suppose you could give very strict incentives for teachers to show up at school: they will be paid if they show up to school, and not otherwise. On the basis of what we have learned in previous lectures, and your intuition, what are the pros and cons?

PROS:

-

-

-

CONS:

-

-

-

Experiment: Teachers working for informal NGO school (Seva Mandir). 60 schools randomly selected out of 120.

Each teacher was given a camera and instructed to take a picture of himself and the children twice a day.

Required at least 8 children (or $\frac{1}{2}$ of the number of registered children).

Rolls of film were collected and developed.

Payments:

-Comparison schools: paid Rs1,000/month–required to attend 20 days/month

-Program schools: base salary Rs 1,000 for 20 days/month; Rs 50 bonus for each day above; Rs 50 fine for each day below; fine capped at Rs 500.

Results: see tables.

Conclusions:

-

-

-

Extension to government schools??? What would make it hard?

Let's look at other things that *may* work in government schools.

2.2   Multitasking: incentives based on test scores

It is more common to provide incentives based on test scores. (e.g. spirit of "no child left behind," and many school reforms in the U.S.).

PROS:

-

-

-

CONS:

-

-

-

One particular worry is the fear of "multitasking" (Holmstrom and Milgrom). When an agent is given very strong incentives along one dimension, he may decide to entirely focus on this

dimension and ignore the others. In the case of test scores this may lead to:

-

-

-

-

Glewwe, Illias, and Kremer evaluate a teacher incentives program based on test scores. The program is based on district test scores in the upper grades. Prizes are offered in two categories: "top scoring schools" and "most improved schools."

Results: see tables. What do we conclude?

Methodological aside: Calculating standard error (2)

A second problem with calculating standard errors arises when we have many outcomes, and some are significant and others are not. How do we know that this is not due to chance? Suppose that a researcher runs 100 regressions with 100 different variables, and reports 5 coefficients that happen to be significant: can this be interpreted? So in a paper like this one, where nothing is significant but one or two variables, do we want to believe these one or two variables, or are we worried this is a fluke?

Likewise, you may be tempted when analyzing the results of an experiment to look at the effect for boys, girls, boys in grade 1, boys in grade 2, etc.... What is the risk when running this? If you do 100 subgroups and run the regressions, how many results do you expect to be significant at the 5% level, even if the null hypothesis of no effect is true?

The problem is that our standard errors are valid when we assume that we run one regression and stand by the result: conventional statistical theory is not equipped for the possibility that you may "pick and choose."

Solutions:

-Need to commit to a design (what regressions you will run is specified in your proposal)

-Need a theoretical model (chain of outcomes) to interpret the results

-Can statistically adjust for multiple outcomes and multiple groups:

-(a) Multiple outcomes can be addressed by looking at average effects across a range of categories.

-(b) Bonferoni bounds compute standard errors across outcomes that cannot be averaged, correcting for the possibility of "picking and choosing"

So: the teachers program seems to have led to multitasking: short term strategies to do well on

the exam, but not real effort to increase learning in those schools.

Something that is a bit surprising, though, is that teachers did not adopt the same strategies in the case of the girls' incentives. We don't fully understand what is going on.

3 How to structure the system to provide incentives?

(If you are interested, Prof. Banerjee and I have a summary paper that covers many of these issues. It is called "Adressing Absence" and can be found on my web page).

3.1 Can schools monitor schools effectively?

Kremer and Chen (2003) tried to provide incentives based on presence for teachers in government schools. The headmaster was asked to mark teacher absence, and the teachers that were present at least some amount of the time received a prize (a bicycle).

Kremer and Chen independently measured absence using random checks.

What do you think happened?

3.2 Can parents monitor schools effectively?

All the rage right now. Idea is to provide "voice" to parents (see World Development Report (2004)). Many countries try to improve the power of school committees, etc.

Kremer and Vermeersh evaluated a project that sought to "empower" school committees.

-NGO, ICS Africa, facilitated a meeting between the school committee and the school administration at the sub-district level to ensure that the information the school committee had on the functioning of the school was transmitted to the hierarchy, who could then act on it.

-ICS financed a prize to be given to the best teacher in the school.

Program was implemented in 36 schools out of 72. Disappointing results: no difference in absence, no difference in test scores!

What could be the reasons why the monitoring is disappointing? (in particular, in contrast with the other successful approaches we saw in the class).

-

-

-

3.3 Should education be privately provided? Vouchers

Let's step aside and think:

-Why should education be publicly funded?

-Why should education be publicly *provided*?

Those two questions are clearly different....

If teachers cannot be disciplined in government schools (for historical or institutional reasons), the question of whether schools should be publicly financed, but privately provided, does arise.

There are two ways of potentially providing publicly funded, privately provided education (or other service):

-Contracting out: i.e. NGOs are given money to run schools, or communities are given money to run a school or hire extra teachers for primary school (the latter is more akin to de-centralization)

-Vouchers for private schools.

Vouchers are often run on a lottery basis (due to a limited budget for the programs). Such a program was run in Columbia (PACES program). Voucher recipients were selected by lottery among the applicants to the program. This made it an (unintended) randomized experiment. For the analysis, we can compare voucher winners and losers.

There is an important feature of the program: students would lose it if they repeated a grade. This means that the effect of the program is a combination of incentive and choice (and not the "pure" effect of school choice).

First paper (Angrist, Bettinger, Kremer, and others): looked at short-term effects of the program.

Second paper (Angrist, Bettinger, Kremer): looked at longer-term effects.

Both papers show positive effects, with winners more likely than winners to:

-be in private school

-score highly on tests

-graduate from secondary school

-do well at secondary school exam

and less likely to:

-be married or cohabiting as a teenager

-be working

Aside: What can we be worried about when just comparing lottery winners and losers (this is an external validity problem)? How could we address this problem?