ECON 326: Economics of Developing Countries
Problem Set 2

due Monday, February 3 at 11:59 PM

1 RCT Design and Evaluation

In this question, you will design and evaluate an RCT that estimates the effect of Balsakhis, or tutors, on grade-6 students in a large school (1,000 students per year) in Mumbai, India.

(a) Your first goal is to randomly select which students will receive tutoring. How can you *guarantee* that the treatment and control group will be balanced on gender?

You randomize using your strategy above. After 9 months (one school year) of the program, you find that the treatment group has higher test scores on a standardized state exam by .34 standard deviations relative to the control group, and the effect was significant at the 5% level. If a question asks you to use some variable, you may assume that you have the necessary data for it.

Assume each of the following questions is fully independent: **they do not build on each other**.

(b) After obtaining this result, you find out that some of the control group students also attended tutoring sessions during the experiment. Is the estimate of .34 upwardly or downwardly biased relative to the true effect of tutors? Explain.

(c) Suppose you learn that not every student in the treatment group attended the tutoring sessions. Let Y_i = individual i's score on a standardized state exam (i.e. Y_i = standard deviations from the mean), Z_i = a dummy equal to 1 if i was assigned to the treatment group, D_i = a dummy equal to 1 if i attended tutoring, and X_i = all relevant controls.

  (i) Should your estimate of .34 now be interpreted as the ATE, ITT, ATT, or $LATE$?

  (ii) Show with equations, how to estimate ITT, ATT, and $LATE$, and interpret each in words (you may assume nobody in the control group attended the sessions, i.e. no always-takers). Who are compliers in this context? What assumption must be satisfied for the $LATE$ to be properly estimated?

(d) What regressor would you add to the regression to test for heterogeneity of the effect with respect to the natural log of parent's income? Suppose the coefficient on this regressor is γ, and it is statistically significant; interpret γ.

(e) Suppose you observe that control group students also improved their scores relative to schools that did not have this experiment performed on them, even though they did not receive any tutoring. Explain in one sentence why this could have happened based on the education literature. Is the true ATE of the tutoring greater than or less than .34?

2 RCT Paper Interpretation

In this question, we will examine a paper by Glewwe, Kremer, and Moulin (linked here) studying a randomized evaluation in rural Kenya that provided textbooks (written in English) to randomly selected schools. The researchers provide English textbooks to grades 3-7, math textbooks to grades 3, 5, and 7, and science textbooks to grade 8 in 25 randomly chosen schools out of 100. We will be looking at a few tables to understand the main points of the paper. You may read and use the paper if you like, but it is not required.

Moving forward, a "textbook school" is a school that received textbooks, and a "comparison school" is a school that did not receive textbooks. Interpret the following tables:

(a) Before the experiment began, students took pretests in three different subjects: English, Math, and Science. The differences in normalized test scores (i.e. the dependent variable is the z-score on the pretest) between the textbook schools and comparison schools are in Table 1, with the standard errors in parentheses below each coefficient. In at most two sentences, what does this table say (broadly; do not discuss any specific estimates), and why is it necessary?

Additionally, suppose the coefficient on science was .3 instead of .173, and it did not happen by chance. Explain how this could lead to bias in estimating the effect of textbooks, and predict and justify the direction of the bias.

TABLE 1—DIFFERENCES IN NORMALIZED PRETEST SCORES BETWEEN TEXTBOOK SCHOOLS AND 25-SCHOOL COMPARISON GROUP

Subject	English		Math		Science		All subjects combined	
Grade	Grades with texts (3–7)	All grades (3–8)	Grades with texts (3, 5, 7)	All grades (3–8)	Grades with texts (8)	All grades (3–8)	Grades with texts	All grades
Difference between textbook schools and comparison schools	0.046 (0.105)	0.033 (0.101)	0.056 (0.090)	0.054 (0.085)	0.173 (0.105)	−0.017 (0.088)	0.061 (0.091)	0.023 (0.087)
Observations	8,516	9,332	5,069	9,302	816	9,276	14,401	27,910

(b) Table 2 lists the average textbooks per pupil in textbook vs. comparison schools in each year of the program. Note that only certain subjects/grades were given textbooks, even in textbook schools (denoted by "Y" and "N" in the second column of the table). In at most two sentences, what does this table say (broadly), and why is it necessary?

TABLE 2—AVAILABILITY OF TEXTBOOKS PER PUPIL

		School-owned books		Privately-owned books		Total	
Program year	Subject/grade given textbooks?	Textbook schools	Comparison schools[a]	Textbook schools	Comparison schools	Textbook schools	Comparison schools
1	Y	0.65	0.04	0.10	0.18	0.75	0.22
	N	0.03	0.03	0.08	0.10	0.11	0.13
2	Y	0.55	0.04	0.09	0.17	0.64	0.21
	N	0.04	0.03	0.08	0.08	0.12	0.12
3	Y	0.52	0.11	0.09	0.14	0.61	0.25
	N	0.11	0.09	0.09	0.09	0.20	0.19
4	Y	0.43	0.10	0.05	0.11	0.48	0.21
	N	0.10	0.08	0.05	0.06	0.17	0.14

(c) Table 4 reports two estimators of the effect of textbooks on average test scores: a level estimator and a difference-in-differences estimator. Which column(s) show the difference-in-differences estimates, and how? In no more than two sentences, what does Table 4 say (broadly)?

TABLE 4—IMPACT OF TEXTBOOK PROGRAM ON NORMALIZED TEST SCORES

Dependent variable	Normalized test score[a b] (1)	Normalized test score[b] (2)	Normalized test score minus pretest score[c] (3)	Normalized test score minus pretest score[c] (4)
Textbook school	0.023	0.020	0.018	−0.046
	(0.087)	(0.104)	(0.053)	(0.071)
Received a textbook				
Region and sex dummies	Y	Y	Y	Y
Years exposed to textbooks	1	2	1	2
Grades	3–8	4–7	3–8	4–7
Observations	24,132	12,663	11,321	7,354

Notes: Standard errors in parentheses.

[a] Running the same regressions for individual subjects English, math, and science (not shown in this table), yields similar results, with the coefficients on textbooks never statistically significantly different from zero.

[b] Sample includes all children enrolled in January of year 1 who took the relevant October/November test

[c] Sample includes all children who were enrolled in January of year 1 and took the relevant October/November test as well as the pretest in January of year 1.

(d) In Table 8, an interaction term is included between the pretest score and a dummy for a textbook school. Quintiles in Table 9 are ranked by pretest score, with 5 representing the highest quintile. What do these tables suggest about effect heterogeneity? Produce a hypothesis as to why this occurred. What additional data would you collect to provide evidence for this hypothesis?

TABLE 8—NORMALIZED TEST SCORES AS A FUNCTION OF TREATMENT AND PRETEST SCORE

Dependent variable	Normalized test score (year 1)	Normalized test score (year 2)	Normalized test score minus pretest scores (year 1)
Textbook school	0.060	−0.016	0.021
	(0.061)	(0.088)	(0.060)
Pretest score	0.430***	0.342***	−0.338***
	(0.013)	(0.016)	(0.016)
Pretest × textbook school	0.057***	0.061***	0.042**
	(0.018)	(0.022)	(0.021)
Observations	11,342	7,393	11,321

TABLE 9—PROGRAM IMPACT ON NORMALIZED TEST SCORES, BY QUINTILE OF PRETEST SCORES

Years exposed	Quintile 1 (1)	Quintile 2 (2)	Quintile 3 (3)	Quintile 4 (4)	Quintile 5 (5)
1	−0.049	−0.021	0.032	0.142*	0.218**
	(0.064)	(0.069)	(0.073)	(0.079)	(0.096)
2	−0.077	−0.109	−0.089	0.021	0.173
	(0.081)	(0.094)	(0.104)	(0.101)	(0.131)

(e) Why do the authors randomize on the school level? Would you expect the estimated effects to increase (away from zero) or decrease (toward zero) if the authors randomized on the student-level (*within* schools) instead?

3 Labeled Cash Transfers

In Turning a Shove into a Nudge: A "Labeled Cash Transfer" for Education, the authors run a randomized experiment in Morocco (the Tayssir program) in which they offer a small monthly cash transfer to parents of school-aged children in poor rural communities explicitly labeled as an education support program (called labeled cash transfers, or LCTs). They compare these to the effects of standard conditional cash transfers (CCTs), where cash was provided conditional on attendance. I strongly encourage you to read the Background and Experimental Design section of the paper for helpful context.

320 schools were originally selected to be in the study. 60 schools were in the control group (receiving nothing), and the remainder of the 260 schools were split 2x2 by *labeled* vs. *conditional* transfers and *mother* vs. *father* beneficiary:

	Labeled (31%)	Conditional (69%)
Mother (50%)	Labeled for education Given to mothers (40 schools)	Conditional on attendance Given to mothers (89 schools)
Father (50%)	Labeled for education Given to fathers (40 schools)	Conditional on attendance Given to fathers (90 schools)

You'll need to do the following steps for all datasets. First, use existing variables generate a new variable that defines each of the 5 groups (control and 4 treatments). Additionally, the *benef* variable is currently a string (text) variable for the beneficiary (mother or father), but string variables cannot be used in a regression. If you need to regress on *benef* or use it in a mathematical expression, use the *encode* command to turn it into a numerical variable. *You may find the command 'tabulate' and 'tabulate, nolabel' to be helpful in figuring out what value corresponds to mothers vs. fathers.*

First things first: we need to check whether our dataset is balanced. Our randomization occurred on the school level, so we need to make sure the schools in each group are similar on average. Let's start by using the *schools_preliminary.dta* dataset. This dataset contains baseline information on each of the 318 schools remaining in the dataset, with almost all of them having a main campus and a randomly chosen "satellite" campus (just treat satellites as additional schools).

(a) Use a single *table* or *tabstat* function to produce a table of baseline means for the number of classrooms, number of teachers, number of boys, number of girls, and share of girls (you will need to generate this) in each of the 5 groups.

(b) These look balanced at first glance, but we need to check statistically...use the *pwmean* command to do a pairwise comparison of means across every group for the share of girls. Are any differences statistically significant? Feel free to check the other variables too, but your submission only needs the comparison for the share of girls.

Now, let's look at the effects on a few student-level outcomes: attendance and their score on a simple math test. Answer questions (c)-(d) using the *attendance.dta* and *aser_math.dta* datasets separately. In the attendance dataset, the *present* variable takes the value 1 if the student is in attendance, and 0 if not (not in attendance **or** dropped out). The *attendance* variable provides more detail on whether they are missing or dropped out and doesn't need to be used for this problem set, but can be used in case you want to explore further. *I would love to combine them into one dataset, but they don't really match well into each other because the math testing sample was much smaller than the attendance one.*

(c) Generate two dummy variables: one equal to 1 if receiving LCT, another equal to 1 if receiving CCT. Run a regression that shows the effects of the LCT and CCT treatments on attendance (using *present*) and the *total* math score (without separating by mother or father). Are the effects significant at a 10% level?

(d) Generate and use an appropriate interaction term in a regression to determine whether the effect of LCTs and CCTs differs significantly (at a 10% level) between whether the beneficiary was the mother or the father. *Hint: make sure the interaction term uses a* ***dummy*** *variable to define different groups; you may need to slightly adjust your encoded variable to do this.*

For the remaining questions, answer them once each.

(e) The authors note that program take-up was very high: 97% of parents with children enrolled in Tayssir treatment schools had a child enrolled in the program. What does this suggest about the relationship between the ATE, ITT, and $LATE$? What if take-up had been much lower, say 50%?

(f) Suppose the paper concludes that LCTs and CCTs produce the same effects. Explain one reason why this is potentially an important finding (i.e. what is the motivation of this experiment in the first place?). *Hint: read the first two pages of the paper.*

(g) As a preview of the gender lecture, why do you think the gender of the beneficiary *could* matter? *Hint: parents are often the ones making the education decision for their children. Which parent do you think has greater decision making power in a developing country?*