University of California, Santa Barbara **Thomas Heckelei**
Bren School of Env. Science & Management **Winter 2024**

EDS 241
Assignment 1: Randomized Controlled Trials and Estimation of Treatment Effects

Due on 1/22/24, 11:59pm
Turn in your Markdown pdf on Canvas in the "Assignments" section.

Start with the provided R-file named "Assignment_1_RMD_template.RMD". Save with your name instead of template and start working with it.

Part 1. Use the small program RCT.R provided with the lecture (but also copied into the template) that generates synthetic potential outcomes without treatment, Yi_0, and with treatment, Yi_1. When reporting findings, report them using statistical terminology (i.e. more than y/n.) Please do the following and answer the respective questions (**briefly**).

(a) Create equally sized treatment and control groups by creating a binary random variable Di where the units with the *1's" are chosen randomly.

(b) Make two separate histograms of Xi for the treatment and control group. What do you see and does it comply with your expectations, explain why or why not?

(c) Test whether Di is uncorrelated with the pre-treatment characteristic Xi and report your finding.

(d) Test whether Di is uncorrelated with the potential outcomes Yi_0 and Yi_1 and report your finding (only possible for this synthetic dataset where we know all potential outcomes).

(e) Estimate the ATE by comparing mean outcomes for treatment and control group. Test for mean difference between the groups and report your findings.

(f) Estimate the ATE using a simple regression of (i) Yi on Di and (ii) Yi on Di and Xi and report your findings and include.

Part 2 is based on Gertler, Martinez, and Rubio-Codina (2012) (article provided with assignment) and covers impact evaluation of the Mexican conditional cash transfer Progresa (later called Oportunidades, now Prospera). Basically, families with low-incomes received cash benefits if they complied to certain conditions, such as regular school attendance for children and regular healthcare visits. You can read more about the program in the Boxes 2.1 (p.10) & 3.1 (p.40) of the Handbook on impact evaluation: quantitative methods and practices by Khandker, B. Koolwal, and Samad (2010). The program followed a randomized phase-in design. You have data on households (hh) from 1999, when treatment hh have been receiving benefits for a year and control hh have not yet received any benefits. You can find a description of the variables at the end of the assignment. Again, briefly report what you find or respond to the questions.

(a) Some variables in the dataset were collected in 1997 before treatment began. Use these variables to test whether there are systematic differences between the control and the

treatment group before the cash transfer began (i.e. test for systematic differences on all 1997 variables). Describe your results. Does it matter whether there are systematic differences? Why or why not? Would it be a mistake to do the same test with these variables if they were collected after treatment began and if so why? Note: If your variable is a proportion (e.g. binary variables), you should use a proportions test, otherwise you can use a t-test.

(b) Estimate the impact of program participation on the household's value of animal holdings (vani) using a simple univariate regression. Interpret the intercept and the coefficient. Is this an estimate of a treatment effect?

(c) Now, include at least 6 independent control variables in your regression. How does the impact of program participation change? Choose one of your other control variables and interpret the coefficient.

(d) The dataset also contains a variable intention_to_treat. This variable identifies eligible households in participating villages. Most of these households ended up in the treatment group receiving the cash transfer, but some did not. Test if the program has an effect on the value of animal holdings of these non-participants (spillover effects). Think of a reason why there might or might not be spillover effects.

References

Gertler, Paul J., Sebastian W. Martinez, and Marta Rubio-Codina (Jan. 2012). "Investing Cash Transfers to Raise Long-Term Living Standards". In: American Economic Journal: Applied Economics 4.1, pp. 164–92. doi: 10.1257/app.4.1.164. https://www.aeaweb.org/articles?id=10.1257/app.4.1.164.

Khandker, Shahidur, Gayatri B. Koolwal, and Hussain Samad (2010). Handbook on impact evaluation: quantitative methods and practices. Washington D.C.: The World Bank.

Table 1: Data description

Variable Name	Variable Label
hhid	Household ID
state	Federal entity
muni	District
local	locality number
female_hh	Head is Female = 1
age_hh	Head's age
educ_hh	Head's education - years
ethnicity_hh	(ind) = 1 if indiv. speaks an indigenous language
educ_sp	Spouse's education - years
age_hh2	Head's age squared
dirtfloor97	(hh) = 1 if dirtfloor in 97
bathroom97	(hh) = 1 if bathroom for exclusive hh use in 97
electricity97	(hh) = 1 if electricity in dwelling in 97
homeown97	(hh) = 1 if home in ownership in 97
primary	(com) = 1 if primary school in the community
telesec	(com) = 1 if telesecundaria in the community
healthcenter	(com) = 1 if health center or medical personnel in town
mobilehealth	(com) = 1 if mobile health services come to town
min_dist	Min dist. between the loc and a large urban center
hhsize97	(hh) Household Size in 97
lnup_cwagepm	(com) log (male commu wage)
lnup_cwagepw	(com) log (female commu wage)
lnup_cwagepc	(com) log (children commu wage)
intention_to_treat	intention to treat = 1, controls = 0
treatment	actual treatment = 1, controls = 0
me	Micro-enterprise = 1
mef	Micro-enterprise (female) = 1
enterprise_m	Micro-enterprise (male) = 1
comuid	(comu) Community Identifier
landpieces	(hh) nb of pieces of land used by the household
land	(hh) = 1 if hh is using any piece of land
ha	(hh) total hectars of land
hairrigation	(hh) total irrigation hectares
haseasonality	(hh) = 1 total seasonality hectares
ani	(hh) = 1 if hh owns animals
ani1	(hh) = 1 if hh owns farm animals (draft)
ani2	(hh) = 1 if hh owns other animals (to sell/sell their products)
no4	(hh) if no agricultural house
small4	(hh) if small farm
big4	(hh) if bif farm (more than 3 ha of land)
landlessmall4	(hh) if small farm without land
landsmall4	(hh) if small farm with land (less than 3 ha)
vani	(hh) value of animals hh owns in 97 prices
vani1	(hh) value of draft animals hh owns in 97 prices
vani2	(hh) value of prod animals hh owns in 97 prices
ani_sales	Sale of any animal in last six months = 1, conditional on having any animal

ani_sales_a	(hh) Total revenue from sales of animals in last 6 months (deflated)
ani_prod	Produced any animal by-product in last month = 1
ani_prod_sales_a	(hh) Total revenue from sales of any animal by-product (cheese, milk, eggs, meat3
crop_sales	(hh) Total revenue from crop sales (corn, beans, other) in last year (deflated)
ag_exp	(hh) Total Agricultural Expenditures for cop production (monthly)
credit	= 1 if anyone in hh has received loan past 6 mths (12 mths 2003 data)
foodexp	Food expenditures
nonfood	Non-Food expenditures
foodexp_pp	Food expenditures PP
nonfoodexp_pp	Non-Food expenditures PP
hhsize	(max) hhsize