Lab Session: Randomization

Cristian Huse

Introduction

This lab session is the first using the material from IEP (Gertler et al 2016), which uses the fictionalized case study of the HISP dataset. Please see Appendices A and B for details. While the original implementation relies on Stata, in what follows we will conduct our analysis in R. Nevertheless, although there are some departures, alternative analyses or generalizations, please note that the R-based sessions still draw substantially on the original material.

This particular lab session deals with **randomization**. The structure of the session is as follows:

- Preparation;
  - Load libraries, setwd, load data (.dta format)
- Example 1. Estimating **treatment effects with random assignment**
  - Compare treatment and control (at the locality level) **post-treatment**
- Example 2. Testing for **balance in baseline outcome**
  - Compare treatment and control **pre-treatment** (data which usually isn't available)
- Example 3. Testing **balance in baseline covariates**
  - We focus on age, but the general idea is to compare whether treatment and control are comparable pre-treatment for all covariates (there is usually a table reporting these results on a paper)
- Example 4. Estimating treatment effects **controlling for household characteristics**
  - This goes back to the first estimation of treatment effects but offers some advantages

Preparation

```
## Initialize ####

rm(list=ls())

#Load libraries
library(haven) #for read_dta
library(modelsummary)

## Set working directory
#setwd("INSERT PATH OF FOLDER WHERE YOU SAVED THE DATASET WITH / or\\")
setwd("C:/Users/huse-admin/Dropbox/CRISTIAN/Teaching/Cursos_Meus/Teaching_2021/EPE/Lab4_Randomization")
```

```
## open data
#Open the cleaned data set
#set path for data
evaluation <- file.path(getwd(), "Data", "evaluation.dta")

#import .dta file
evaluation.df <- read_dta(evaluation)
```

Randomized Assignment

In this context, the program is randomized at the village level, and you compare follow-up situation of eligible households in treatment and comparison villages. That is, compare only among eligible.

```
#Select the relevant data, i.e., keep if eligible == 1
random.df = subset(evaluation.df, eligible == 1)
```

Example 1. Randomized Assignment in a Regression Framework (Linear Regression)

Under randomized assignment, one compares the effect of treatment P (treatment_locality, treatment is at the locality level) on the outcome of interest Y (health expenditures). To do so, one compares the outcome for individuals in treatment and control groups, as detailed in the lecture. In regression notation:

$$Y_i = \alpha + \delta P_i + \varepsilon_i$$

where the estimated regression coefficient of δ provides an estimate of the ATE (the difference in average outcomes between the treated and non-treated subsamples) and the error term ε captures individual factors that may affect the relationship between the program and the outcome. (Please see the lecture for details.)

To implement this, *"we regress our outcome of interest (health expenditures) on a binary treatment variable (treatment_locality), which is equal to 1 if the household is located in a treatment area. We run this regression on the sample of households that were eligible for HISP and based on data collected after the program has been administered (round = 1). The estimated regression coefficient for δ is -10.14, indicating that eligible households exposed to HISP spent \$10.14 less on health expenditures than eligible households in the comparison group. The standard error of the coefficient is 0.396. The t-statistic calculated in the regression, -25.63, shows that this coefficient is statistically significant at the 1 percent level."*

That is, individuals are compared only once treatment was applied, i.e., follow-up, denoted by **round == 1**. Given the potential within locality cross-correlation, standard errors are clustered at the locality level.

```
ex1_lm <- lm(health_expenditures ~ treatment_locality, round == 1, data = random.df)

#tmp.df = subset(evaluation.df, eligible == 1 & round ==1)
#plot(tmp.df$treatment_locality, tmp.df$health_expenditures)

model1 <- list("(1) TE w/ Randomization" = ex1_lm)

modelsummary(model1, vcov = ~ locality_identifier,
             stars = c("*" = .1, "**" = .05, "***" = .01),
             fmt = 3, gof_omit = "AIC|BIC|Log.Lik.")
```

	(1) TE w/ Randomization
(Intercept)	17.981***
	(0.307)
treatment_locality	−10.140***
	(0.396)
Num.Obs.	5629
R2	0.300
R2 Adj.	0.300
F	2415.822
Std.Errors	C: locality_identifier

* p < 0.1, ** p < 0.05, *** p < 0.01

Example 2. Testing for Balance in a Baseline Outcome

"Independence of potential outcomes is one of the most crucial assumptions to ensure that the difference in average outcomes between program beneficiaries and non beneficiaries provides a consistent estimate of the average treatment effect. While this assumption cannot generally be verified, some falsification tests can be implemented to identify cases when it does not hold."

Assuming pre-programme data is available, there should be **no measurable effect** between treated and non-treated individuals **before** the programme's introduction (**round == 0**). In regression notation, with $t = 0$ being the pre-treatment period:

$$Y_{i,t=0} = \alpha + \delta P_i + \varepsilon_i$$

"When the baseline outcome is regressed on a binary variable capturing exposure to treatment, the estimated coefficient is very small (-0.084) and not statistically significant (p value of 0.693). This indicates that eligible households in the treatment and comparison groups have similar levels of health expenditures prior to the intervention."

This type of analysis is sometimes referred to as a **placebo** – one should not see an effect of the programme prior to its introduction.

```
ex2_lm <- lm(health_expenditures ~ treatment_locality,
             round == 0, data = random.df)

model2 <- list("(2) Balance in Baseline Outcome" = ex2_lm)

modelsummary(model2, vcov = ~ locality_identifier,
             stars = c("*" = .1, "**" = .05, "***" = .01),
             fmt = 3, gof_omit = "AIC|BIC|Log.Lik.")
```

	(2) Balance in Baseline Outcome
(Intercept)	14.574***
	(0.156)
treatment_locality	−0.084
	(0.213)
Num.Obs.	5628
R2	0.000
R2 Adj.	0.000
F	0.537
Std.Errors	C: locality_identifier

* p < 0.1, ** p < 0.05, *** p < 0.01

	(3) Balance in Baseline Covariate (age)
(Intercept)	42.292***
	(0.430)
treatment_locality	−0.635
	(0.532)
Num.Obs.	5628
R2	0.001
R2 Adj.	0.000
F	2.872
Std.Errors	C: locality_identifier

* p < 0.1, ** p < 0.05, *** p < 0.01

Example 3. Testing Balance in a Baseline Covariate

In a similar vein as in Example 2, one can check balancing, i.e., check whether characteristics are similar across treated and non-treated individuals pre-treatment (**round == 0**). Denoting a characteristic by X, balance in this other observed pre-program characteristic can be estimated using the following regression;

$$X_{i,t=0} = \alpha + \delta P_i + \varepsilon_i$$

where the change from Example 2 is the replacement of Y with X.

```
ex3_lm <- lm(age_hh ~ treatment_locality,
             round == 0, data = random.df)

model3 <- list("(3) Balance in Baseline Covariate (age)" = ex3_lm)

modelsummary(model3, vcov = ~ locality_identifier,
             stars = c("*" = .1, "**" = .05, "***" = .01),
             fmt = 3, gof_omit = "AIC|BIC|Log.Lik.")
```

Example 4. Randomized Assignment in a Regression Framework (Multivariate Regression)

Finally, we can control for characteristics of treatment and control groups in a regression framework, pre-treatment (**round == 1**). This is directly comparable to Example 1 above.

The estimated regression coefficient for δ -10.01, implying that households in a treated locality $10.01 less on health expenditures than those in localities not exposed to the program – **holding all the control variables constant**. The t-statistic calculated in the regression shows that this coefficient is statistically significant at the 1 percent level.

```
ex4_lm <- lm(health_expenditures ~ treatment_locality + age_hh + age_sp + educ_hh +
               educ_sp + female_hh + indigenous + hhsize + dirtfloor + bathroom + land +
               hospital_distance,
             round == 1, data = random.df)

model4 <- list("(4) TE w/ Controls" = ex4_lm)

modelsummary(model4, vcov = ~ locality_identifier,
             stars = c("*" = .1, "**" = .05, "***" = .01),
             fmt = 3, gof_omit = "AIC|BIC|Log.Lik.")
```

	(4) TE w/ Controls
(Intercept)	27.565***
	(0.864)
treatment_locality	−10.010***
	(0.341)
age_hh	0.041***
	(0.015)
age_sp	0.003
	(0.017)
educ_hh	−0.039
	(0.047)
educ_sp	−0.022
	(0.049)
female_hh	0.643
	(0.444)
indigenous	−1.905***
	(0.350)
hhsize	−1.603***
	(0.066)
dirtfloor	−1.849***
	(0.278)
bathroom	0.285
	(0.246)
land	0.038
	(0.038)
hospital_distance	−0.003
	(0.004)
Num.Obs.	5629
R2	0.430
R2 Adj.	0.428
F	352.612
Std.Errors	C: locality_identifier

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1) TE w/ Randomization	(4) TE w/ Controls
(Intercept)	17.981***	27.565***
	(0.307)	(0.864)
treatment_locality	−10.140***	−10.010***
	(0.396)	(0.341)
age_hh		0.041***
		(0.015)
age_sp		0.003
		(0.017)
educ_hh		−0.039
		(0.047)
educ_sp		−0.022
		(0.049)
female_hh		0.643
		(0.444)
indigenous		−1.905***
		(0.350)
hhsize		−1.603***
		(0.066)
dirtfloor		−1.849***
		(0.278)
bathroom		0.285
		(0.246)
land		0.038
		(0.038)
hospital_distance		−0.003
		(0.004)
Num.Obs.	5629	5629
R2	0.300	0.430
R2 Adj.	0.300	0.428
F	2415.822	352.612
Std.Errors	C: locality_identifier	C: locality_identifier

* p < 0.1, ** p < 0.05, *** p < 0.01

Comparing Estimates

We now report the results of Examples 1 and 4, but one could have one table for placebo effects, one for covariate balance etc.

```
modelsummary(c(model1,model4), vcov = ~ locality_identifier,
             stars = c("*" = .1, "**" = .05, "***" = .01),
             fmt = 3, gof_omit = "AIC|BIC|Log.Lik.")
```

References

Gertler, Paul J.; Martinez, Sebastian; Premand, Patrick; Rawlings, Laura B.; Vermeersch, Christel M. J. (2016). Impact Evaluation in Practice, Second Edition, Technical Companion (Version 1.0). Washington, DC: Inter-American Development Bank and World Bank.

Appendix A. IEP's Technical Companion

***Impact Evaluation in Practice*, Second Edition**

***Technical Companion*[1]** *(Version 1.0, September 2016)*

Introduction

Impact Evaluation in Practice (second edition) offers a comprehensive and accessible introduction to impact evaluation for policy makers and development practitioners. The book is divided into four parts. Part 1 reviews how to prepare for an impact evaluation, what to evaluate, and why. Part 2 presents the basic concepts of impact evaluation by relying mostly on intuition and graphical representations. Part 3 discusses how to choose an impact evaluation method in a given operational context, and how to manage impact evaluations. Part 4 reviews how to get data, including sampling, power calculations, and data sources for impact evaluation. The presentation in the book is nontechnical, and focuses on the intuition behind technical concepts and impact evaluation methods, as well as sampling and power calculation.

In this technical companion, we include additional material for readers with a background in statistics and econometrics. The technical companion assumes a basic understanding of statistics, in particular key concepts such as regression analysis and hypothesis testing. The technical companion presents an introduction to the analysis of impact evaluation data. It summarizes the basic potential outcome framework that underpins the econometrics of impact evaluation, discusses how to represent the methods in a regression framework, and provides some applications using Stata. The technical companion also provides examples of how to undertake power calculations in Stata and Optimal Design. Applications are based on the case study of the Health Insurance Subsidy Program (HISP) presented in the book. Supplementary data and related do-files can be found on the book website (www.worldbank.org/ieinpractice), and can be used to replicate the results presented in the companion.

While this technical companion presents an introduction to analyzing impact evaluation data, the objective is not to provide an in-depth discussion of the econometrics behind impact evaluation. If you would like additional details or a more comprehensive coverage, you are invited to read Angrist and Pischke (2009) or Angrist and Pischke (2014), on which this companion partly draws. Although we provide some examples and applications using Stata, this online companion is not a thorough empirical guide on how to apply the methods in practice. If you are interested in additional information and practical applications, you can consult, among other relevant material, the applied impact evaluation methods course at the University of California, Berkeley (http://aie.cega.org).

[1] Please cite this technical companion as: *"Gertler, Paul J.; Martinez, Sebastian; Premand, Patrick; Rawlings, Laura B.; Vermeersch, Christel M. J.. 2016. Impact Evaluation in Practice, Second Edition, Technical Companion (Version 1.0). Washington, DC: Inter-American Development Bank and World Bank.*
Marina Tolchinsky provided excellent research assistance in preparing this technical companion and related Stata material. Aakash Mohpal provided useful comments. This is version 1.0 (September 2016) of the technical companion. The authors welcome feedback and suggestions on how to improve this first version in the future.

1

Figure 1: Technical Companion, p. 1

Appendix B. IEP's Technical Companion Read Me: Dataset Description

Please make sure to return to this description also in future labs so you understand how the data is being "cut" according to the method under study.

TECHNICAL COMPANION READ ME

The HISP program

The book on *Impact Evaluation in Practice* (second edition) and related technical companion book uses a fictionalized case, the Health Insurance Subsidy Program (HISP) to illustrate many of the concepts and methods that are presented. While fictionalized, the HISP case is modeled after real-world examples of impact evaluations. One of the primary objectives of HISP is to reduce the burden of health-related out-of-pocket expenditures for low income households.

The HISP Dataset

The HISP dataset has the Stata format and is called "evaluation.dta". It can be downloaded with accompanying do files from the IE in practice website.

The dataset is at the level of household and round. This means that one observation (row) captures information for one household either in the baseline (round 0) or in the follow-up survey (round 1). In other words, every household appears on two rows: one for the baseline and one for the follow-up survey.

The dataset includes the following variables (columns):

Outcome variables	
health_expenditures	Out of pocket health expenditure (per capita per year)
Control variables	
age_hh	Age of the head of the household (in years)
Age_sp	Age of the spouse (years)
educ_hh	Education of the head of household (completed years of schooling)
educ_sp	Education of the spouse (completed years of schooling)
indigenous	Head of household speaks an indigenous language (0=no, 1=yes)
female_hh	Head of the household is a woman (0=no, 1=yes)
hhsize	Number of household members (at baseline)
dirtfloor	Home has a dirt floor at baseline (0=no, 1=yes)
bathroom	Home with private bathroom at baseline (0=no, 1=yes)
land	Number of hectares of land owned by household at baseline
hospital_dist	Distance to closest hospital
Other variables	
locality_identifier	Locality identifier
household_identifier	Unique household identifier
round	Survey round (0 = baseline; 1 = follow-up)
enrolled	Household enrolled in HISP (0=no, 1=yes)
enrolled_rp	Household enrolled in HISP under the randomized promotion

Figure 2: Technical Companion Read Me, p. 1

	scenario (0=no, 1=yes)
eligible	Household eligible to enroll in HISP (0=no, 1=yes)
treatment_locality	Household is located in treatment community (0=no, 1=yes)
promotion_locality	Household is located in locality randomly assigned promotion of HISP (0=no, 1=yes)
poverty_index	Poverty index 1-100 (eligible ≤ 58)
hospital	HH member visited hospital in the past year (0=no, 1=yes), used in power calculations

Figure 3: Technical Companion Read Me, p. 2