

# Taller 3 - Experimento Aleatorio

Maria Camila Caraballo, Laura Sarif Rivera

2025-09-10

## Experimentos Aleatorios

Este trabajo tiene como objetivo replicar parte de los resultados de Banerjee, Duflo, Glennerster y Kinnan (2015) sobre el impacto del microcrédito en Hyderabad, India. A través de esta replicación buscamos poner en práctica lo aprendido en clase y comprender cómo la evidencia empírica respalda (o cuestiona) las conclusiones sobre el papel de las microfinanzas en la reducción de la pobreza.

### 1. Análisis introductorio

Al introducir el artículo, los autores analizaron las limitaciones y lineamientos de la investigación de microcrédito publicada.

**¿Cuáles son los problemas en la identificación del impacto del microcrédito si uno simplemente compara los que tienen microcrédito y los que no tienen microcrédito? ¿Qué método utilizan los autores para identificar tal impacto? ¿Cómo resuelve ese método los problemas antes mencionados?**

Los principales problemas de identificación en este caso radican en que, si se comparan de manera ingenua los hogares que recibieron microcrédito con aquellos que no lo hicieron, surge un sesgo de selección. Esto ocurre porque pueden existir múltiples dimensiones observables y no observables que influyen tanto en la probabilidad de recibir un préstamo como en los resultados que se quieren evaluar. Por ejemplo, los hogares que buscan y obtienen microcrédito suelen ser más ambiciosos y tener mayores aspiraciones de emprender, expandir negocios o invertir en educación, lo que los hace sistemáticamente distintos de aquellos que no buscan crédito. Además, los prestatarios pueden contar con más activos, mejores redes sociales o un patrimonio inicial que los hace más aptos para acceder y aprovechar un préstamo. A nivel contextual también existe un sesgo de ubicación del programa, ya que las instituciones tienden a abrir sucursales en barrios donde esperan una demanda mínima y donde existen condiciones estructurales más favorables, dejando fuera barrios con características distintas.

Para superar estos problemas, los autores implementaron un experimento aleatorizado a nivel de barrio en Hyderabad, India. Partieron de 104 barrios identificados por Spandana como posibles lugares de expansión y los emparejaron según sus características socioeconómicas. Dentro de cada par, uno fue asignado aleatoriamente al tratamiento, es decir recibir una sucursal, y el otro al control. Como resultado, 52 barrios obtuvieron acceso al microcrédito y 52 quedaron como grupo de comparación.

La asignación aleatoria asegura que, en promedio, los barrios tratados y los de control sean equivalentes en sus condiciones previas tanto observables como no observables, de modo que la única diferencia sistemática entre ambos grupos es la intervención. Así, cualquier diferencia en los resultados posteriores puede atribuirse causalmente al acceso al microcrédito. Además, este diseño resuelve el problema de program placement al desligar la decisión de abrir sucursales de las condiciones del barrio, ya que no es la institución la que decide estratégicamente dónde expandirse, sino el azar. Con ello se elimina la posibilidad de que los mejores barrios sean los que reciban primero el programa, lo cual garantiza que la comparación refleje el verdadero impacto del microcrédito y no diferencias preexistentes.

## 2. Estimaciones originales

Los autores reportan los resultados de estimar la siguiente ecuación:

$$y_{ia} = \alpha + \beta Treat_{ia} + X'_a \gamma + \epsilon_i$$

Donde  $y_{ia}$  es un resultado para el hogar  $i$  en el área  $a$ ,  $Treat_{ia}$  es una dicótoma que toma el valor de 1 si el hogar está ubicado en un área tratada, mientras que  $\beta$  es el efecto de  $ITT$ .  $X_a$  es el vector con dimensión  $K \cdot 1$  de variables de control

### ¿Cuál es la unidad de análisis en el estudio?

Si bien los hogares son la unidad de análisis empleada para medir los resultados, la aleatorización del tratamiento se llevó a cabo a nivel de barrio, de modo que la apertura de la sucursal se asignó colectivamente y no de manera individual.

Replique la tabla 2 panel A, columnas 1 y 3 para mostrar el efecto de microfinanzas (ser tratado por el programa) en la obtención de más préstamos (acceso a créditos). Explique la intuición detrás de los resultados. Presente sus resultados con 4 cifras significativas.

```
# Cargar librerías y paquetes

#install.packages("haven") cual
#install.packages("plm")
#install.packages("skimr")
#install.packages("dplyr")
#install.packages("ggplot2")
#install.packages("modelsummary")
#install.packages("knitr")
#install.packages("tinytex")
#tinytex::install_tinytex()
#install.packages("survey")
library(knitr)
library(haven)
library(plm)
library(scales)
library(skimr)
library(dplyr)
library(ggplot2)
library(AER)
library(fixest)
library(lmtest)
library(sandwich)
library(tidyverse)
library(stargazer)
library(dplyr)
library(survey)

# Leer archivo .dta
data <- read_dta("data_endline_1and2.dta")

# Ver base
str(data)
```

```
## tibble [6,863 x 44] (S3: tbl_df/tbl/data.frame)
```

```
## $ hhid : num [1:6863] 1 2 3 4 5 6 7 8 9 10 ...  
## .. attr(*, "label")= chr "Household ID"  
## .. attr(*, "format.stata")= chr "%10.0g"  
## $ areaid : num [1:6863] 1 1 1 1 1 1 1 1 1 1 ...  
## .. attr(*, "label")= chr "Area ID"  
## .. attr(*, "format.stata")= chr "%8.0g"  
## $ treatment : dbl+lbl [1:6863] 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1  
## ..@ label : chr "Treatment area"  
## ..@ format.stata: chr "%9.0g"  
## ..@ labels : Named num [1:2] 0 1  
## ..- attr(*, "names")= chr [1:2] "Control" "Treatment"  
## $ w : num [1:6863] 0.82 1 1 1 1 ...  
## .. attr(*, "label")= chr "Raw weight"  
## .. attr(*, "format.stata")= chr "%9.0g"  
## $ w1 : num [1:6863] 0.777 1 1 1 1 ...  
## .. attr(*, "label")= chr "Weight for endline 1 results"  
## .. attr(*, "format.stata")= chr "%9.0g"  
## $ w2 : num [1:6863] 0.82 1 1 1 1 ...  
## .. attr(*, "label")= chr "Weight for endline 2 results"  
## .. attr(*, "format.stata")= chr "%9.0g"  
## $ sample1 : dbl+lbl [1:6863] 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1  
## ..@ label : chr "Surveyed in endline 1"  
## ..@ format.stata: chr "%9.0g"  
## ..@ labels : Named num [1:2] 0 1  
## ..- attr(*, "names")= chr [1:2] "No" "Yes"  
## $ area_pop_base : num [1:6863] 272 272 272 272 272 272 272 272 272 272 ...  
## .. attr(*, "label")= chr "Area population, baseline"  
## .. attr(*, "format.stata")= chr "%9.0g"  
## $ area_debt_total_base : num [1:6863] 81050 81050 81050 81050 81050 ...  
## .. attr(*, "label")= chr "Total outstanding debt in area, baseline"  
## .. attr(*, "format.stata")= chr "%12.0g"  
## $ area_business_total_base: num [1:6863] 11 11 11 11 11 11 11 11 11 11 ...  
## .. attr(*, "label")= chr "Total number of businesses in area, baseline"  
## .. attr(*, "format.stata")= chr "%9.0g"  
## $ area_exp_pc_mean_base : num [1:6863] 1335 1335 1335 1335 1335 ...  
## .. attr(*, "label")= chr "Area mean monthly per-capita expenditure, baseline"  
## .. attr(*, "format.stata")= chr "%9.0g"  
## $ area_literate_head_base : num [1:6863] 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 ...  
## .. attr(*, "label")= chr "Area literacy rate (HH heads only), baseline"  
## .. attr(*, "format.stata")= chr "%9.0g"  
## $ area_literate_base : num [1:6863] 0.534 0.534 0.534 0.534 0.534 ...  
## .. attr(*, "label")= chr "Area literacy rate, baseline"  
## .. attr(*, "format.stata")= chr "%9.0g"  
## $ spandana_1 : dbl+lbl [1:6863] 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0  
## ..@ label : chr "Has outstanding loan from Spandana, endline 1"  
## ..@ format.stata: chr "%9.0g"  
## ..@ labels : Named num [1:2] 0 1  
## ..- attr(*, "names")= chr [1:2] "No" "Yes"  
## $ othermfi_1 : dbl+lbl [1:6863] 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0  
## ..@ label : chr "Has outstanding loan from non-Spandana MFI, endline 1"  
## ..@ format.stata: chr "%9.0g"  
## ..@ labels : Named num [1:2] 0 1  
## ..- attr(*, "names")= chr [1:2] "No" "Yes"  
## $ anymfi_1 : dbl+lbl [1:6863] 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0
```

```

## ..@ label      : chr "Has outstanding loan from any MFI, endline 1"
## ..@ format.stata: chr "%9.0g"
## ..@ labels     : Named num [1:2] 0 1
## .. ..- attr(*, "names")= chr [1:2] "No" "Yes"
## $ anybank_1     : dbl+lbl [1:6863] 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0
## ..@ label      : chr "Has outstanding loan from bank, endline 1"
## ..@ format.stata: chr "%9.0g"
## ..@ labels     : Named num [1:2] 0 1
## .. ..- attr(*, "names")= chr [1:2] "No" "Yes"
## $ anyinformal_1 : dbl+lbl [1:6863] 1, 0, 1, 1, 0, 1, 0, 0, 0, 0, 0, 1, 1, 0, 0, 1
## ..@ label      : chr "Has outstanding informal loan, endline 1"
## ..@ format.stata: chr "%9.0g"
## ..@ labels     : Named num [1:2] 0 1
## .. ..- attr(*, "names")= chr [1:2] "No" "Yes"
## $ spandana_amt_1 : num [1:6863] 18000 0 0 0 0 15000 0 15000 0 0 ...
## ..- attr(*, "label")= chr "Total outstanding loan amount, Spandana (Rs.), endline 1"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ othermfi_amt_1 : num [1:6863] 0 0 0 0 0 0 0 0 0 0 ...
## ..- attr(*, "label")= chr "Total outstanding loan amount, non-Spandana MFI (Rs.), endline 1"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ anymfi_amt_1   : num [1:6863] 18000 0 0 0 0 15000 0 15000 0 0 ...
## ..- attr(*, "label")= chr "Total outstanding loan amount, all MFIs (Rs.), endline 1"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ bank_amt_1     : num [1:6863] 0 0 0 0 0 0 0 0 0 30000 ...
## ..- attr(*, "label")= chr "Total outstanding loan amount, banks (Rs.), endline 1"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ informal_amt_1 : num [1:6863] 93540 0 60000 60000 0 ...
## ..- attr(*, "label")= chr "Total outstanding loan amount, informal (Rs.), endline 1"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ anyloan_amt_1  : num [1:6863] 115780 0 0 51700 23000 ...
## ..- attr(*, "label")= chr "Total outstanding loan amount (Rs.), endline 1"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ bizassets_1    : num [1:6863] 0 0 2000 0 31700 0 0 0 0 12000 ...
## ..- attr(*, "label")= chr "Business assets (stock, Rs.), endline 1"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ bizinvestment_1 : num [1:6863] 0 0 0 0 0 0 0 0 0 0 ...
## ..- attr(*, "label")= chr "Business investment (last 12 months, Rs.), endline 1"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ bizprofit_1    : num [1:6863] 0 0 1595 4795 3650 ...
## ..- attr(*, "label")= chr "Business profits (last 30 days, Rs.), endline 1"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ any_new_biz_1  : dbl+lbl [1:6863] 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0
## ..@ label      : chr "Started a business in last year, endline 1"
## ..@ format.stata: chr "%9.0g"
## ..@ labels     : Named num [1:2] 0 1
## .. ..- attr(*, "names")= chr [1:2] "No" "Yes"
## $ bizassets_2    : num [1:6863] 0 0 0 2915 34902 ...
## ..- attr(*, "label")= chr "Business assets (stock, Rs.), endline 2"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ bizinvestment_2 : num [1:6863] 0 0 0 0 0 ...
## ..- attr(*, "label")= chr "Business investment (last 12 months, Rs.), endline 2"
## ..- attr(*, "format.stata")= chr "%9.0g"
## $ bizprofit_2    : num [1:6863] 0 0 2049 2082 NA ...
## ..- attr(*, "label")= chr "Business profits (last 30 days, Rs.), endline 2"

```

```
## -- attr(*, "format.stata")= chr "%9.0g"
## $ any_new_biz_2 : dbl+lbl [1:6863] 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
## ..@ label : chr "Started a business in last year, endline 2"
## ..@ format.stata: chr "%9.0g"
## ..@ labels : Named num [1:2] 0 1
## .. -- attr(*, "names")= chr [1:2] "No" "Yes"
## $ total_exp_mo_pc_1 : num [1:6863] 769 1371 1246 1133 982 ...
## -- attr(*, "label")= chr "Total expenditure per capita (monthly, Rs.), endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ durables_exp_mo_pc_1 : num [1:6863] NA 9 50.8 38.3 NA ...
## -- attr(*, "label")= chr "Durables expenditure per capita (monthly, Rs.), endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ nondurable_exp_mo_pc_1 : num [1:6863] NA 1362 1195 1095 NA ...
## -- attr(*, "label")= chr "Non-durables expenditure per capita (monthly, Rs.), endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ temptation_exp_mo_pc_1 : num [1:6863] 60.7 0 23.9 0 0 ...
## -- attr(*, "label")= chr "Temptation goods expenditure per capita (monthly, Rs.), endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ girl515_school_1 : num [1:6863] NA 1 NA 1 NA 0.5 NA NA NA NA ...
## -- attr(*, "label")= chr "Share of females 5-15 in school, endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ boy515_school_1 : num [1:6863] NA 1 1 1 1 1 NA 1 NA 1 ...
## -- attr(*, "label")= chr "Share of males 5-15 in school, endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ girl515_workhrs_pc_1 : num [1:6863] NA 0 NA 0 NA 0 NA NA NA NA ...
## -- attr(*, "label")= chr "Hours worked per female 5-15 in last 7 days, endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ boy515_workhrs_pc_1 : num [1:6863] NA 0 0 0 0 0 NA 0 NA 0 ...
## -- attr(*, "label")= chr "Hours worked per male 5-15 in last 7 days, endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ girl1620_school_1 : num [1:6863] 0 NA NA NA NA NA 0 NA 1 NA ...
## -- attr(*, "label")= chr "Share of females 16-20 in school, endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ boy1620_school_1 : num [1:6863] 1 NA NA NA NA 0 0 NA 1 NA ...
## -- attr(*, "label")= chr "Share of males 6-20 in school, endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ women_emp_index_1 : num [1:6863] -0.4154 0.5629 -0.0623 -0.368 -0.3653 ...
## -- attr(*, "label")= chr "Index of women's independence/empowerment, endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## $ bizexpense_1 : num [1:6863] 0 0 205 205 8750 ...
## -- attr(*, "label")= chr "Business expenses (last 30 days, Rs.), endline 1"
## -- attr(*, "format.stata")= chr "%9.0g"
## - attr(*, "label")= chr "Endline data for \"The miracle of microfinance?\" (Banerjee et al.), AEJ 2005"
## - attr(*, "notes")= chr [1:2] "1" "householdid identifies panel HHs"
```

##	hhid	areaid	treatment	w
##	0	0	0	0
##	w2	sample1	area_pop_base	area_debt_total_base
##	0	0	0	0
##	area_exp_pc_mean_base	area_literate_head_base	area_literate_base	spandana_1
##	0	0	0	52

##	anymfi_1	anybank_1	anyinformal_1	spandana_amt_1
##	52	52	52	52
##	anymfi_amt_1	bank_amt_1	informal_amt_1	anyloan_amt_1
##	52	52	52	1
##	bizinvestment_1	bizprofit_1	any_new_biz_1	bizassets_2
##	63	624	106	721
##	bizprofit_2	any_new_biz_2	total_exp_mo_pc_1	durables_exp_mo_pc_1
##	773	721	36	82
##	temptation_exp_mo_pc_1	girl515_school_1	boy515_school_1	girl515_workhrs_pc_1
##	36	3828	3790	3828
##	girl1620_school_1	boy1620_school_1	women_emp_index_1	bizexpense_1
##	4689	4997	1	178

*# Eliminar NA*

```
data_inicial <- data %>% filter(!is.na(spandana_1))
```

*# Filtrar solo hogares de Endline 1*

```
data_inicial <- subset(data_inicial, sample1 == 1)
```

*# Definir diseño muestral con pesos y clustering por área*

```
design1 <- svydesign(
  id = ~areaid,
  weights = ~w1,
  data = data_inicial
)
```

*# Regresiones*

```
spandana <- svyglm(
  spandana_1 ~ treatment + area_debt_total_base + area_pop_base + area_business_total_base + area_exp_pc_1,
  design = design1,
  family = gaussian()
)
```

```
mfi <- svyglm(
  anymfi_1 ~ treatment + area_debt_total_base + area_pop_base + area_business_total_base + area_exp_pc_1,
  design = design1,
  family = gaussian()
)
```

*# 6. Mostrar tabla de resultados*

```
stargazer(spandana, mfi, type = "text",
  keep = "treatment",
  covariate.labels = c("Tratamiento"),
  dep.var.labels = c("Spandana",
    "Cualquier MFI"),
  digits = 4,
  title = "Tabla 2- Crédito")
```

```
##
## Tabla 2- Crédito
## =====
##               Dependent variable:
##      -----
##               Spandana      Cualquier MFI
```

```
##              (1)              (2)
## -----
## Tratamiento      0.1274***      0.0835***
##              (0.0197)      (0.0270)
## -----
## Observations      6,811      6,811
## Log Likelihood    -1,618.3450  -3,601.7220
## Akaike Inf. Crit.  3,252.6900  7,219.4450
## =====
## Note:              *p<0.1; **p<0.05; ***p<0.01
```

Los resultados muestran que, en el caso de *Spandana*, vivir en un barrio tratado incrementó en aproximadamente 12,7 puntos porcentuales la probabilidad de que un hogar tuviera un préstamo activo con esta institución, efecto que resulta estadísticamente significativo y refleja el impacto positivo y directo de la apertura de la sucursal sobre el uso de sus créditos.

En cuanto a la variable *Any MFI*, también se observa un efecto positivo y significativo: la probabilidad de contar con un préstamo de cualquier institución de microfinanzas aumentó en 8,3 puntos porcentuales en los barrios tratados. Este resultado confirma que la entrada de una nueva institución amplía la oferta y mejora el acceso, aunque el incremento es menor al observado para Spandana, lo cual sugiere que parte del aumento corresponde a un desplazamiento de clientes desde otras MFIs.

**Replique la tabla 3 panel A, columnas 3 y 4. Explique la intuición detrás de los resultados.**

```
# Revisar NA de la base
colSums(is.na(data))
```

```
##              hhid              areaid              treatment              w
##              0              0              0              0
##              w2              sample1              area_pop_base      area_debt_total_base
##              0              0              0              0
##      area_exp_pc_mean_base  area_literate_head_base      area_literate_base      spandana_1
##              0              0              0              52
##              anymfi_1              anybank_1              anyinformal_1      spandana_amt_1
##              52              52              52              52
##              anymfi_amt_1              bank_amt_1              informal_amt_1      anyloan_amt_1
##              52              52              52              1
##              bizinvestment_1              bizprofit_1              any_new_biz_1      bizassets_2
##              63              624              106              721
##              bizprofit_2              any_new_biz_2              total_exp_mo_pc_1      durables_exp_mo_pc_1
##              773              721              36              82
##      temptation_exp_mo_pc_1      girl1515_school_1      boy515_school_1      girl1515_workhrs_pc_1
##              36              3828              3790              3828
##              girl1620_school_1      boy1620_school_1      women_emp_index_1      bizexpense_1
##              4689              4997              1              178
```

```
# Eliminar NA
data_gasto <- data %>% filter(!is.na(bizexpense_1))
data_ganancia <- data %>% filter(!is.na(bizprofit_1))

# Definir diseño muestral con pesos y clustering por área en cada base
design2 <- svydesign(
  id = ~ areaid, # cluster nivel área
```

```

weights = ~ w1, # pesos
data = data_gasto
)

design3 <- svydesign(
  id = ~ areaid, # cluster nivel área
  weights = ~ w1, # pesos
  data = data_ganancia
)

# Columna 3
gasto <- svyglm(
  bizexpense_1 ~ treatment + area_debt_total_base + area_pop_base + area_business_total_base + area_exp_1,
  design = design2,
  data = data_gasto
)

# Columna 4
ganancia <- svyglm(
  bizprofit_1 ~ treatment + area_debt_total_base + area_pop_base + area_business_total_base + area_exp_1,
  design = design3,
  data = data_ganancia
)

# Tabla
stargazer(gasto, ganancia, type = "text",
  keep = "treatment",
  covariate.labels = c("Treatment"),
  dep.var.labels = c("Gasto",
    "Ganancias"),
  digits = 0,
  title = "TABLA 3 -(Panel A)")

```

```

##
## TABLA 3 -(Panel A)
## =====
##                               Dependent variable:
##                               -----
##                               Gasto      Ganancias
##                               (1)        (2)
## -----
## Treatment                    255        354
##                               (1,056)    (313)
## -----
## Observations                 6,685      6,239
## Log Likelihood               -78,063    -67,434
## Akaike Inf. Crit.            156,142    134,884
## =====
## Note:                        *p<0.1; **p<0.05; ***p<0.01

```

La Tabla 3 muestra que, aunque los efectos del tratamiento sobre los gastos y las ganancias de los negocios son positivos, ninguno de ellos resulta estadísticamente significativo. Esto significa que, en promedio, el



acceso al microcrédito no generó un aumento claro ni en la inversión en insumos ni en las utilidades de los hogares. Como los datos incluyen ceros para aquellos que no tienen un negocio, los resultados sugieren que la disponibilidad de crédito no necesariamente se traduce en un incremento de ingresos para los hogares pobres a través de la expansión de sus actividades empresariales.

En este sentido, los hallazgos cuestionan la idea de que el financiamiento mediante microcrédito sea un determinante suficiente para mejorar las condiciones económicas de los hogares o para incentivar de manera sistemática la creación de nuevos emprendimientos.

**Usando “any MFI” como definición del tratamiento y “Treated Area” como asignación aleatoria inicial, calcule la tasa de cumplimiento (*compliance rate*) ¿Parece ser alta o baja?**

Con base en los resultados, la tasa de participación efectiva en microcrédito es del 25 % en los hogares de áreas tratadas y del 18 % en los de control. La diferencia entre ambos grupos, equivalente a 7 puntos porcentuales, constituye la tasa de cumplimiento. Este valor es relativamente bajo, lo que indica que la asignación a un área tratada solo aumentó de manera limitada la probabilidad de acceder a un préstamo. En consecuencia, el instrumento derivado de la asignación aleatoria resulta válido pero débil, ya que no genera un cambio sustancial en la participación en instituciones de microcrédito.

```
# Eliminar NA
data_clean <- data %>% filter(!is.na(anymfi_1), !is.na(treatment), !is.na(w1))

# tasas de cumplimiento por grupo
comp_rates <- data_clean %>%
  group_by(treatment) %>%
  summarise(tasa_cumplimiento = weighted.mean(anymfi_1, w = w1, na.rm = TRUE))

# calcular la diferencia (first stage)
first_stage <- comp_rates$tasa_cumplimiento[comp_rates$treatment == 1] -
  comp_rates$tasa_cumplimiento[comp_rates$treatment == 0]

p <- first_stage
```

### 3. Cálculo de impacto

Ahora quieres estimar el impacto de recibir efectivamente un microcrédito sobre 2 variables: i) el consumo total (*total\_exp\_mo\_pc\_1*) y ii) los beneficios del negocio (*bizprofit\_2*)

Para ello, se toma como variable de tratamiento efectivo a la dicótoma *anymfi\_1* que toma el valor de uno si el hogar recibió algún préstamo de cualquier IMF para la línea final 1 y cero en caso contrario. Podemos tomar como instrumento el hecho de que la ubicación de los bancos fue aleatoria. Para esto, use como instrumento en sus estimaciones la asignación aleatoria.

Finalmente, no incluya ningún control en sus estimaciones y responda las siguientes preguntas teniendo en cuenta estas nociones y variables:

**¿Cuáles son los supuestos necesarios para que los estimadores de VI sean válidos (en el contexto del artículo)?**

En el artículo, la estrategia empírica busca estimar el efecto causal de recibir un préstamo de microcrédito de instituciones financieras sobre variables como el consumo y las utilidades del negocio. Dado que el acceso al crédito no fue asignado de manera estrictamente aleatoria, sino que surgió como un spillover derivado de la apertura de sucursales, los autores emplean como instrumento la ubicación aleatoria de los bancos. Para que esta estrategia de variables instrumentales sea válida, deben cumplirse ciertos supuestos fundamentales.

En primer lugar, el instrumento debe ser relevante, lo que significa que la ubicación aleatoria de las sucursales debe estar correlacionada de manera significativa con la probabilidad de que un hogar reciba un microcrédito.

En términos prácticos, vivir en un barrio tratado debería aumentar de forma sustancial la probabilidad de acceder a un préstamo.

En segundo lugar, el instrumento debe ser exógeno, es decir, independiente de los factores no observados que influyen directamente en las variables de interés, como el consumo o las condiciones del hogar. En este caso, la exogeneidad se justifica porque la asignación de las sucursales fue producto de un procedimiento aleatorio; por lo tanto, residir en un área tratada no debería estar relacionado con características previas de los hogares. De igual manera, debe cumplirse la condición de exclusión, que establece que la apertura de la sucursal solo puede afectar el consumo y las condiciones del hogar a través de su impacto sobre la probabilidad de recibir un microcrédito, y no mediante otros mecanismos alternativos como cambios en expectativas, precios locales o salarios.

Finalmente, se requiere la condición de monotonicidad, que implica que no debe existir un grupo de hogares para quienes la apertura de la sucursal reduzca la probabilidad de acceder a microcrédito. En otras palabras, la presencia de una sucursal no debería hacer que un hogar que habría solicitado un préstamo en ausencia del tratamiento deje de hacerlo.

Para cada una de las dos variables, estime el efecto *Local Average Treatment Effect* de obtener un préstamo, a través del método de Variable Instrumental. Estime también el ITT. Para cada tipo de efecto (LATE e ITT) presente una tabla con sus estimaciones y sus respectivos errores estándar. Interprete sus resultados, ¿Cuál es la diferencia en la interpretación de los dos tipos de efecto?

```
# Eliminar NA
data_exp <- data %>% filter(!is.na(total_exp_mo_pc_1))
data_prof <- data %>% filter(!is.na(bizprofit_1))

# Intention-to-Treat (ITT) esto compara todos los municipios tratados vs. control.
itt1_consumo <- lm(total_exp_mo_pc_1 ~ treatment, data = data_exp, weights = w1)
itt2_profit <- lm(bizprofit_1 ~ treatment, data = data_prof, weights = w1)

# Guardar coeficientes ITT
itt_consumo_hat <- coef(itt1_consumo)["treatment"]
itt_profit_hat <- coef(itt2_profit)["treatment"]

# LATE (Local Average Treatment Effect) es el efecto promedio para los individuos cuyo estatus de trata
late1_consumo <- ivreg(total_exp_mo_pc_1 ~ anymfi_1 | treatment, weights = w1, data = data_exp)
late2_profit <- ivreg(bizprofit_1 ~ anymfi_1 | treatment, weights = w1, data = data_prof)

# Guardar coeficientes LATE
late_consumo_iv <- coef(late1_consumo)["anymfi_1"]
late_profit_iv <- coef(late2_profit)["anymfi_1"]

# Agrupar errores por barrio ITT
coeftest(itt1_consumo, vcov = vcovCL(itt1_consumo, cluster = ~ areaid))

##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1419.157    31.061  45.6898  <2e-16 ***
## treatment    37.980    46.215   0.8218   0.4112
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
coeftest(itt2_profit, vcov = vcovCL(itt2_profit, cluster = ~ areaid))
```

```
##
## t test of coefficients:
##
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)   743.65     178.89  4.1571 3.267e-05 ***
## treatment     422.92     318.85  1.3264  0.1847
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Agrupar errores por barrio LATE
```

```
coeftest(late1_consumo, vcov = vcovCL(late1_consumo, cluster = ~ areaid))
```

```
##
## t test of coefficients:
##
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1336.20     143.36  9.3204 <2e-16 ***
## anymfi_1     460.09     654.44  0.7030  0.4821
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
coeftest(late2_profit, vcov = vcovCL(late2_profit, cluster = ~ areaid))
```

```
##
## t test of coefficients:
##
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -205.31     936.69 -0.2192  0.8265
## anymfi_1     5490.60    4626.33  1.1868  0.2353
```

```
# Matrices para cluster
```

```
matriz_ITT1  <- vcovCL(itt1_consumo, cluster = ~ areaid)
matriz_ITT2  <- vcovCL(itt2_profit, cluster = ~ areaid)
matriz_LATE1 <- vcovCL(late1_consumo, cluster = ~ areaid)
matriz_LATE2 <- vcovCL(late2_profit, cluster = ~ areaid)
```

```
# Errores estandar
```

```
se_itt_consumo  <- sqrt(diag(matriz_ITT1))
se_itt_profit   <- sqrt(diag(matriz_ITT2))
se_late_consumo <- sqrt(diag(matriz_LATE1))
se_late_profit  <- sqrt(diag(matriz_LATE2))
```

```
# Tabla
```

```
stargazer(itt1_consumo, itt2_profit, late1_consumo, late2_profit,
  type = "text",
  dep.var.labels.include = FALSE,
  dep.var.labels = NULL,
  se = list(se_itt_consumo, se_itt_profit, se_late_consumo, se_late_profit),
  column.labels = c("ITT consumo", "ITT Utilidad",
                    "LATE consumo", "LATE Utilidad"),
  title = "Impacto del microcrédito con errores agrupados por barrio")
```

```
##
## Impacto del microcrédito con errores agrupados por barrio
## =====
##                               Dependent variable:
##                               -----
##                               OLS          OLS          instrumental    instrumen
##                               ITT consumo  ITT Utilidad  variable         variable
##                               (1)          (2)          LATE consumo    LATE Util.
##                               -----
## treatment                    37.980      422.924
##                               (46.215)     (318.846)
##
## anymfi_1                     460.094      5,490.6
##                               (654.441)     (4,626.3
##
## Constant                    1,419.157***  743.653***    1,336.201***    -205.3
##                               (31.061)     (178.889)    (143.364)     (936.68
## -----
## Observations                 6,827        6,239        6,775        6,190
## R2                           0.0003       0.0003       -0.033       -0.03
## Adjusted R2                  0.0001       0.0002       -0.033       -0.03
## Residual Std. Error 1,099.659 (df = 6825) 11,807.920 (df = 6237) 1,113.639 (df = 6773) 12,031.210 (d
## F Statistic                 1.987 (df = 1; 6825) 1.954 (df = 1; 6237)
## =====
## Note:                               *p<0.1; **p<0.05; ***p<0.01
```

El efecto de intención de tratar (ITT) mide el impacto promedio de vivir en un barrio donde se abrió una sucursal de microfinanzas, independientemente de que los hogares tomaran o no un crédito. En los resultados, los coeficientes de consumo y utilidades son positivos, pero no estadísticamente significativos, lo que sugiere que la mera presencia de la sucursal y la posibilidad de acceder a un préstamo no generaron cambios robustos en las condiciones económicas de los hogares.

Por su parte, el efecto local promedio del tratamiento (LATE) estima el impacto del microcrédito sobre los hogares cuya decisión de endeudarse dependió de la apertura de la sucursal (los *compliers*). Aunque los resultados muestran efectos de mayor magnitud en consumo y, especialmente, en utilidades, estos no alcanzan significancia estadística, lo que indica que incluso entre quienes tomaron crédito inducidos por la sucursal no hay evidencia concluyente de mejoras sostenidas en sus ingresos o consumo.

En conjunto, los resultados sugieren que, si bien el acceso al microcrédito incrementa el endeudamiento, su impacto causal sobre el consumo y las utilidades de los hogares no es estadísticamente significativo, lo que pone en cuestión la idea de que el microcrédito, por sí solo, sea un motor determinante de mejoras económicas a corto plazo.

**¿Cuál debería ser la relación matemática entre la tasa de cumplimiento, el *LATE* y el *ITT*? ¿Parece mantenerse esta relación en sus resultados? En una tabla presente sus cálculos de *LATE* matemáticos manuales y discuta.**

```
# LATE manual
late_consumo_manual <- itt_consumo_hat / p
late_profit_manual  <- itt_profit_hat  / p

# Tabla final
tabla_resultados <- data.frame(
```

```

Outcome = c("Consumo", "Utilidad"),
ITT = c(itt_consumo_hat, itt_profit_hat),
Tasa_cumplimiento = c(p, p),
LATE_manual = c(late_consumo_manual, late_profit_manual),
LATE_IV = c(late_consumo_iv, late_profit_iv)
)
tabla_resultados

```

```

##      Outcome      ITT Tasa_cumplimiento LATE_manual  LATE_IV
## 1  Consumo  37.97957      0.07421012    511.7842  460.094
## 2 Utilidad 422.92353      0.07421012   5699.0004 5490.600

```

El LATE se obtiene dividiendo el estimador ITT por la tasa de cumplimiento. Esta identidad refleja que el ITT mide el efecto promedio de la asignación al tratamiento en toda la población, mientras que el LATE aísla el impacto del microcrédito únicamente en el subgrupo de hogares cuya decisión de endeudarse depende de dicha asignación (los *compliers*).

Al comparar los valores, se observa que los LATE calculados manualmente son muy similares a los estimados mediante el método de variables instrumentales, lo que confirma la coherencia de la estrategia. Las pequeñas diferencias entre ambos procedimientos se explican porque el cálculo manual no incorpora la estimación de errores estándar, covarianzas ni otros elementos estadísticos que sí considera la técnica de IV.

En conclusión, la relación matemática entre ITT, tasa de cumplimiento y LATE se cumple en la práctica, lo que respalda la validez de la estrategia de identificación utilizada y permite interpretar los resultados como una estimación confiable del efecto local del microcrédito sobre los hogares cuya participación fue inducida por la apertura de las sucursales.