

# Potential outcomes and Causal Models

What if?

Fernando Rios-Avila

## Introduction: What if?

From here on, the whole purpose of the methodologies we will discuss is the analysis of causal effects of some:

Policy, treatment, experiment, or otherwise event

But, How is it different from what we did before?

It isn't.

Under Exogeneity assumption  $E(e|X) = 0$ , one can make causal effect claims.

! We seek the truth

How much of the change in outcome is caused by the program alone?

But this is not as easy.

## Examples...Where is the challenge?

A few examples for Causal effect questions:

- Do minimum wages increase unemployment ?
- Do Conditional cash transfers improve health outcomes in children?
- Do Covid Vaccines help reduce the Spread of Covid?

These questions are, however, difficult to answer.

- How do you make sure the “treatment” is the Only factor that explains the difference in outcome across groups??

To do this we need strategies that rule out any other explanation that could “take away” the connection we seek.

We need to close *all* back doors, block all **alternative** explanations, or nuisance factors

## We need to figure out what is what we seek

### Potential Outcomes

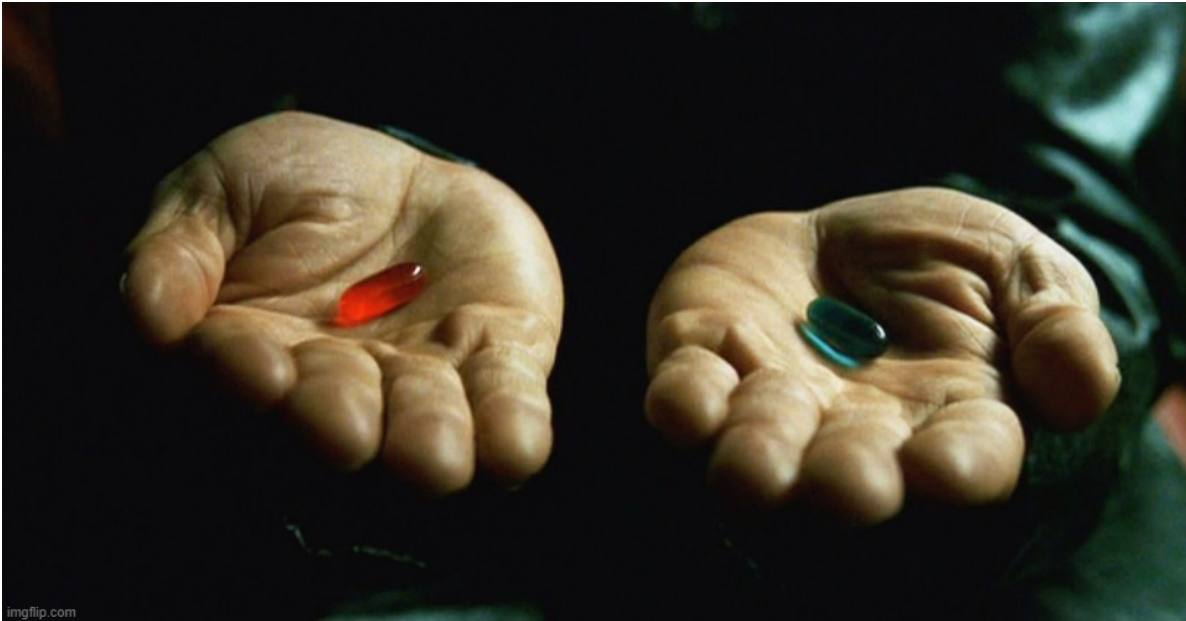


Figure 1: Choice

---

In a world of quantum mechanics, multiverses, parallel worlds, and other infinite number of scenarios, Causal Inference would be extremely simple:

*Path1 : Neo → RedPill → wakesup → SavesWorld*

*Path2 : Neo → BluePill → goestoSleep → Nothing happend*

Then just compare the outcomes: The red Pill Saved the world.

Thus, treatment effect is simply the difference in the potential outcomes: ***Two counterfactuals!***

This difference captures the TRUE treatment effect. Easy enough...

## The Path not taken

The way we express the  $TE_i$ , compares *counterfactuals*. Two possible States of the world, where an allknowing researcher can perfectly identify TE.

Unfortunately, the same person cannot take both paths, and we cannot see both options. A person is either **Treated** or **Untreated**. Thus, the first approach to Casual effects is impossible.

...

But it does provide us with a clue of how to go ahead and analyze Causal effects. We “simply” need to estimate the counterfactual!

But before going deeper into how to estimate the counterfactuals And treatment effects some notation

## Potential Outcomes: Notation

- $i$  will represent a unit. Person, city, country, school, classroom, etc
- $D$  will indicate the treatment Status of a unit.  $D = 1$  means is treated, and  $D = 0$  is untreated.
- Each unit has two potential outcomes  $Y_i(D = 1)$  and  $Y_i(D = 0)$
- All units have only **one** effective or realized outcome:  $Y_i$ , which is what we observe, and depends on their treatment status:

$$Y_i = (1 - D_i) * Y_i(0) + D_i * Y_i(1)$$

- Unit Specific causal effect is the difference between potential outcomes:

$$\delta_i = Y_i(1) - Y_i(0)$$

- $\pi$  is the proportion of treated units

## Parameters of Interest

Assume we can see the true USCE (unit specific casual effect) for all units. In addition to their Treatment Status.

There are three parameters one might be interested in analyzing:

$$\begin{aligned}ATE &= E(\delta_i) \\ATT &= E(\delta_i | D_i = 1) \\ATU &= E(\delta_i | D_i = 0)\end{aligned}$$

In general, this three estimands may tell very different stories.

Lets Put some numbers here

## Simulating effects Stata

```
clear
set linesize 255
set seed 101
set obs 1000
gen y0 = rnormal(5)
gen t = rnormal(0.0,0.5)
gen y1 = y0+t
** Assume only those with t>0 take treatment
gen trt =(t>0)
gen y = y0 * (1-trt) + y1 * (trt)
format y0 y1 y t %4.3f
list in 1/10, sep(0) clean
** For Everyone 100 obs
tabstat t, by(trt)
```

<IPython.core.display.HTML object>

Number of observations (\_N) was 0, now 1,000.

	y0	t	y1	trt	y
1.	5.254	-0.801	4.453	0	5.254
2.	4.997	0.039	5.035	1	5.035
3.	2.608	-0.490	2.118	0	2.608

4.	6.280	-0.579	5.700	0	6.280
5.	5.761	0.090	5.850	1	5.850
6.	6.132	0.211	6.342	1	6.342
7.	4.928	0.153	5.081	1	5.081
8.	4.377	-0.073	4.304	0	4.377
9.	3.142	0.427	3.569	1	3.569
10.	6.050	-0.332	5.718	0	6.050

Summary for variables: t  
Group variable: trt

trt	Mean
0	-.3984745
1	.3748617
Total	-.0187664

## But there is only 1

But, we never see potential outcomes, nor unit specific effects.

The most naive estimator is to just estimate the mean difference in “post-treatment” outcome after treatment was in place. But that would be very biased!

```
gen yy0 = y if trt==0
gen yy1 = y if trt==1
list y yy1 yy0 trt in 1/10
reg y trt
```

(491 missing values generated)

(509 missing values generated)

	y	yy1	yy0	trt
1.	5.254	.	5.254051	0
2.	5.035	5.035442	.	1
3.	2.608	.	2.608155	0
4.	6.280	.	6.27964	0
5.	5.850	5.850478	.	1

	-----				
6.	6.342	6.34237	.	1	
7.	5.081	5.080995	.	1	
8.	4.377	.	4.376629	0	
9.	3.569	3.568727	.	1	
10.	6.050	.	6.049989	0	
	-----				

Source		SS	df	MS	Number of obs	=	1,000
	-----				F(1, 998)	=	49.49
Model		49.4759442	1	49.4759442	Prob > F	=	0.0000
Residual		997.714103	998	.99971353	R-squared	=	0.0472
	-----				Adj R-squared	=	0.0463
Total		1047.19005	999	1.04823829	Root MSE	=	.99986

y		Coefficient	Std. err.	t	P> t	[95% conf. interval]
	-----					
trt		.4449359	.0632467	7.03	0.000	.320824 .5690477
_cons		4.988793	.0443179	112.57	0.000	4.901826 5.07576
	-----					

## Bias Direction

$$\begin{aligned}
 E(Y_i|D_i = 1) - E(Y_i|D_i = 0) &= ATE \\
 &+ E(Y(0)|D = 1) - E(Y(0)|D = 0) \\
 &+ (1 - \pi)(ATT - ATU)
 \end{aligned}$$

Intuition: Simple Difference will be biased because

1. There could be a selection bias (one group baseline outcome is different from the other)
2. Treatment Heterogeneity. Some groups are affected differently from others

From these two problems, the second one is easier to handle (either concentrate on ATT or ATU).

The first one, however, requires using strategies to be able to account for selection bias.

## Independence assumption

While the Simple mean estimator is most likely to be biased, under Independence assumption, it may still work:

$$Y(1), Y(0) \perp D$$

This means that Treatment Status should NOT depend on the potential outcomes.

In other words, there shouldnt be any differneces in the potential outcomes *before* or *after* treatment takes place.

This eliminates the selection bias. And group Heterogeneity.

$$\begin{aligned} ATT - ATU &= E(Y|D = 1) - E(Y(0)|D = 1) \\ &\quad - E(Y(1)|D = 0) - E(Y|D = 0) \end{aligned}$$

## SUTVA

### Stable Unit Treatment value assumption

This is a strong assumption that is still required to estimate of treatment effects.

It assumes that:

- **Treatment is Homogenous.** Same intensity, quality, type of treatment among all treated.
- **There are no spill overs.** Your Treatment Status effects you and you only, and you are only affected by your treatment status. No externalities nor Spillovers.
- Also, there are no general equilibrium effects
- And NO Anticipation.

This assumptions namely guaranties that when a unit is not treated, its/his/her outcome will not change.

## Narrowing down the problem

1. Individual level effects are impossible to identify. We only observe one outcome at a time. (not both)
2. It is possible to identify causal effects on groups (treated, not-treated, kind-of-treated). But..
3. Simple Mean difference will not identify causal effects, unless Independence and Sutva assumptions hold

This however, suggests a path. Constructing good counterfactuals can help identifying the Causal effects.

Goal:

- Identify a control/comparison group that is statistically identical to the treated group, except for the **Treatment Status**

## The Gold Standard: Randomized Control Trials

! To keep in mind

Searching for good controls doesn't require having access to perfect "*clones*". However, in average, we need groups (T vs UT) that are very similar to each other.

In general, research designed is guided by the rules of program or treatment assignment on participants.

When researchers have control on the assignment rules, the best approach is to design a randomized control trial.

In an RCT, randomized assignment, eliminates any selection-bias problems (although SUTVA remains as an assumption)

## RCT and Selection Problems

Consider the example of the Health Impacts of Hospitals.

- Hospitals (or health care) should improve health of individuals. but
- Only unhealthy people will use Health care services. Selection bias
- Thus It may look that Hospitals Hurt people's health because those who used it have lower health than those who don't:

$$E(Y|D = 1) - E(Y|D = 0) = ATT + E(Y(0)|D = 1) - E(Y(0)|D = 0)$$



So even if Health services help those in need (ATT). If the selection bias is large, Naive estimations may suggest Hospitals are harmful.

This is a problem caused because Treatment (going to the Hospital) is affected by pre-conditions or potential treatment outcomes.

---

Under Random assignment, and SUTVA, this can be fixed:

1. If assignment is given as a lottery, potential benefits (Treatment effects) will not be affected by the lottery.
2. Here, ATT will be the same as ATU. which will allow us to estimate ATE as:

$$\begin{aligned}
 ATE &= E(Y|D = 1) - E(Y|D = 0) \\
 &= E(Y|D = 1) - E(Y(0)|D = 1) \\
 &= E(Y(1)|D = 0) - E(Y|D = 0)
 \end{aligned}$$

## Stata Example

```

gen trt2=rnormal(>0
gen yrct = y0 * (1-trt2) + y1 * (trt2)
tabstat t y0 y1 yrct, by(trt2)
reg yrct trt2

```

Summary statistics: Mean  
Group variable: trt2

trt2	t	y0	y1	yrct
0	-.0097507	5.061626	5.051875	5.061626
1	-.027891	4.984309	4.956418	4.956418
Total	-.0187664	5.023199	5.004433	5.009338

Source	SS	df	MS	Number of obs	=	1,000
Model	2.76706003	1	2.76706003	F(1, 998)	=	2.64
				Prob > F	=	0.1046

Residual		1046.22695	998	1.0483236	R-squared	=	0.0026
-----+-----					Adj R-squared	=	0.0016
Total		1048.99401	999	1.05004406	Root MSE	=	1.0239
-----							
yrct		Coefficient	Std. err.	t	P> t	[95% conf. interval]	
-----+-----							
trt2		-.1052076	.0647568	-1.62	0.105	-.2322827	.0218675
_cons		5.061626	.0456524	110.87	0.000	4.97204	5.151212
-----							

## Internal vs External Validity

RCTs are not the only strategy that allows you to control the Rules of Treatment. In *Experimental Economics* this is done quite often

you select your sample (of students), and randomly assigned treatments of interest (Experimental design).

There are, however, further considerations to be taken:

### ! Internal Validity

The estimated Impact is net of all other confounding factors (Random assignment)

### ! External Validity

The estimated impact can be generalized to the population in general. (Random Sample)

### ! How to Randomize

Depends mostly on how reasonable is to maintain SUTVA assumption

## Examples

### **i** CCT and Education in Mexico

Progresa/Prospera is a CCT program for poor mothers based on children school enrollment to support education attainment.

Eligibility was based on Census data on Poverty levels and Baseline Data collection. But during phase-in period, only 2/3 of localities were selected to receive the Transfer

### **i** Water and Sanitation Intervention in Bolivia

In 2012 IADB and Bolivian Government implemented a random assignment of water/sanitation interventions in Small Rural Communities.

From 369 eligible communities, 182 were selected at random for the program implementation, via public lotteries, constraining on Community Size

## How to Analyze the Data

1. **Verify Data Balance:** Even if treatment was assigned at random, it is important to verify if groups remain comparable. (Thus avoid composition effects)
2. **Mean Difference:** Because of Random Assignment, one could use simple mean differences to estimate ATE
3. **Consider using controls:** Under RA, controls will not affect the outcome, but may improve precision:  $y_i = \alpha_0 + \tau D_i + x_i \beta + \varepsilon_i$ .
  - But controls should not be affected by the treatment itself (thus should be pre-treatment)
4. May consider **falsification** tests
5. And Other Robustness test: Outliers, or distributional impacts may be of interest

## Example

School subsidies for the poor:

Evaluating the Mexican Progresa poverty program

Paul Schultz (2004)

## Background

Progresa is a Cash Transfer Program designed to increase School Enrollment among the poor, minimizing disincentives to work.

This program provided Grants to families whose children attended school for at least 85% of the school year, covering between 50/75% of school cost.

While there were 495 localities that were eligible to benefit from the program, only 314 were **randomly** selected to start receiving resources for the first 2 years. With the unselected localities being treated a couple of years later.

The program continued beyond the original scope of the policy, now it's known as Progresa/Oportunidades

## Method

Goal. Estimate the impact of Progresa (P) on Enrollment (S)

$$S_i = a_0 + a_1P_i + a_2E_i + a_3P_iE_i + \delta Enrolled_c + \beta X + e_i$$

$P_i = 1$  if the community is eligible

$E_i = 1$  if the child is Poor

$P_i * E_i$  the impact on Poor Children in eligible communities.

Model can be estimated Separately (5 years), or using pooled data

This is a kind of DID model. However, we could consider it as a simple mean comparison between those Effectively treated and those untreated.

## Differences in Characteristics

Comparison of mean pre-program characteristics of the panel samples of poor children

Variables	“Poor” potentially eligible children age 6–16, in pre-program surveys from panel samples	
	Resident in Progresa locality (treatment)	Resident in non-Progresa locality (control)
Sample size	17,286	10,278
Enrollment rate	0.896	0.891
Mother’s years of education	2.60	2.62
No resident mother	0.039	0.043
Father’s years of education	2.77	2.73
No resident father	0.090	0.096
Variables	“Poor” potentially eligible children age 6–16, in pre-program surveys from panel samples	
	Resident in Progresa locality (treatment)	Resident in non-Progresa locality (control)
Student/teacher ratio in local primary school	17.9	17.6
Distance to secondary school (km)	2.08	2.10
No distance data on school	0.024	0.018
Distance to Cabeceras (km)	9.15	9.70
Distance to metropolitan area (km)	107	105

## Raw Differences

Table 3

Differences between enrollment rates between Progresa and non-Progresa poor children and over time<sup>a</sup> (significance level

Year of schooling completed in previous year	Pre-program difference of poor Progresa–non-Progresa D1			Post-program difference of poor Progresa–non-Progresa D1		
	All	Female	Male	All	Female	Male
0	0.009 (0.351)	0.010 (0.433)	0.007 (0.615)	– 0.002 (0.854)	– 0.010 (0.564)	0.006 (0.742)
1	0.001 (0.410)	– 0.009 (0.816)	0.010 (0.376)	0.022 (0.008)	0.007 (0.418)	0.036 (0.002)
2	– 0.004 (0.276)	– 0.013 (0.386)	0.006 (0.506)	0.020 (0.009)	0.018 (0.796)	0.021 (0.001)
3	0.015 (0.278)	0.025 (0.162)	0.005 (0.882)	0.032 (0.008)	0.013 (0.679)	0.049 (0.001)
4	0.008 (0.500)	– 0.016 (0.836)	0.030 (0.266)	0.041 (0.001)	0.038 (0.261)	0.044 (0.001)
5	0.015 (0.129)	0.005 (0.544)	0.025 (0.125)	0.047 (0.001)	0.055 (0.232)	0.041 (0.000)
6	0.024 (0.345)	0.048 (0.433)	– 0.019 (0.002)	0.111 (0.002)	0.148 (0.001)	0.065 (0.317)
7	– 0.012 (0.894)	– 0.005 (0.854)	– 0.015 (0.958)	0.013 (0.147)	0.025 (0.533)	0.003 (0.006)
8	– 0.030 (0.913)	– 0.051 (0.932)	– 0.016 (0.836)	0.001 (0.162)	0.015 (0.575)	– 0.010 (0.100)
9 or More	0.103 (0.534)	0.327 (0.001)	– 0.156 (0.006)	0.066 (0.317)	0.111 (0.042)	0.026 (0.813)

Figure 2: Poor HH

Table 4

Difference between enrollment inequality between Progresa and non-Progresa localities<sup>a</sup> (significance levels in parentheses)

Years of schooling completed in previous year	Pre-program differences, D2			Post-program differences, D2		
	All	Female	Male	All	Female	Male
0	0.010 (0.609)	0.009 (0.752)	0.011 (0.691)	0.049 (0.063)	0.010 (0.784)	0.094 (0.014)
1	- 0.002 (0.904)	0.010 (0.703)	- 0.013 (0.601)	- 0.032 (0.083)	- 0.034 (0.205)	- 0.030 (0.225)
2	- 0.009 (0.64)	- 0.012 (0.646)	- 0.006 (0.816)	- 0.027 (0.099)	- 0.011 (0.659)	- 0.040 (0.069)
3	- 0.009 (0.637)	- 0.032 (0.243)	0.012 (0.649)	- 0.027 (0.083)	- 0.016 (0.464)	- 0.037 (0.091)
4	0.002 (0.936)	0.026 (0.327)	- 0.022 (0.408)	- 0.043 (0.007)	- 0.044 (0.053)	- 0.038 (0.087)
5	- 0.020 (0.293)	- 0.003 (0.909)	- 0.037 (0.165)	- 0.047 (0.003)	- 0.047 (0.042)	- 0.049 (0.025)
6	0.042 (0.023)	- 0.009 (0.736)	0.124 (0.000)	- 0.035 (0.006)	- 0.119 (0.000)	0.061 (0.001)
7	0.014 (0.627)	0.010 (0.814)	0.015 (0.710)	0.002 (0.910)	0.026 (0.369)	- 0.021 (0.441)
8	0.023 (0.545)	0.024 (0.665)	0.029 (0.577)	0.002 (0.936)	- 0.026 (0.428)	0.025 (0.406)
9 or More	- 0.022 (0.726)	- 0.284 (0.002)	0.266 (0.003)	0.014 (0.551)	0.094 (0.006)	0.110 (0.000)

Figure 3: Non Poor HH

**With Controls**

Estimated derivatives at sample means	Panel matched sample			
	Primary		Secondary	
	Female	Male	Female	Male
Progresa locality $a^*_1$ ( $t$ ratio)	0.0005 (0.12)	- 0.0088 (1.62)	- 0.0232 (0.81)	0.0048 (0.18)
Poor eligible household in Progresa locality $a^*_3$ ( $t$ ratio)	0.0087 (2.17)	0.0168 (3.97)	0.1155 (4.32)	0.0572 (2.51)
Net Progresa impact [significance non-zero based on joint $\chi^2$ test]	0.0092 [0.0030]	0.0080 [0.0015]	0.0923 [0.0003]	0.0620 [0.0027]
Sample size	33,795	36,390	13,872	14,523
Pseudo $R^2$	0.3728	0.3712	0.3116	0.2979

## Other Outcomes

Table 7

Cumulative expected enrollment years for birth cohort of poor children who enroll and complete grade 1

Grade completed	Pre-program rounds 1 and 2		Post-program rounds 3, 4, and 5		Difference in differences	
	Progresa	Non-Progresa	Progresa	Non-Progresa	D1	DD1
1	0.977	0.975	0.975	0.953	0.022	0.020
2	0.936	0.938	0.939	0.899	0.040	0.042
3	0.896	0.884	0.904	0.837	0.067	0.041
4	0.856	0.838	0.866	0.768	0.098	0.080
5	0.816	0.786	0.825	0.695	0.130	0.100
6	0.464	0.428	0.511	0.352	0.159	0.121
7	0.436	0.407	0.484	0.330	0.154	0.125
8	0.414	0.399	0.450	0.306	0.144	0.129
Expected total years						
enrolled for both sexes	6.80	6.66	6.95	6.14	0.81	0.66
Years enrolled females	6.66	6.62	6.95	6.19	0.76	0.72
Years enrolled males	6.93	6.72	6.96	6.11	0.85	0.64

## Next Class: REG and Panel FE