CSP · INGENIERÍA INDUSTRIAL UNIVERSIDAD DE CHILE

CENTRO DE SISTEMAS PÚBLICOS

IN4402: Applied statistics for management and economics

Difference in Difference estimator

ANDRÉS FERNÁNDEZ

www.sistemaspublicos.cl

DIFFERENCES IN DIFFERENCES - INTRO

CUASI AND NON EXPERIMENTAL METHODS

- Sometimes treatments (laws, reforms, funding) are randomly assigned and sometimes they are not
  - We would still want to know if they had impact

DIFFERENCES IN DIFFERENCES - INTRO

CUASI AND NON EXPERIMENTAL METHODS

- Sometimes we can “explode” the gradual or “natural” assignment of treatment to evaluate impact
  - Groups should be comparable

DIFFERENCES IN DIFFERENCES - ASSUMPTIONS

CUASI AND NON EXPERIMENTAL METHODS

- How to compare these groups?
  - We could check data before and after the intervention
  - We could compare the *difference* between group in order to take care of *unobservable* differences

DIFFERENCES IN DIFFERENCES - ASSUMPTIONS

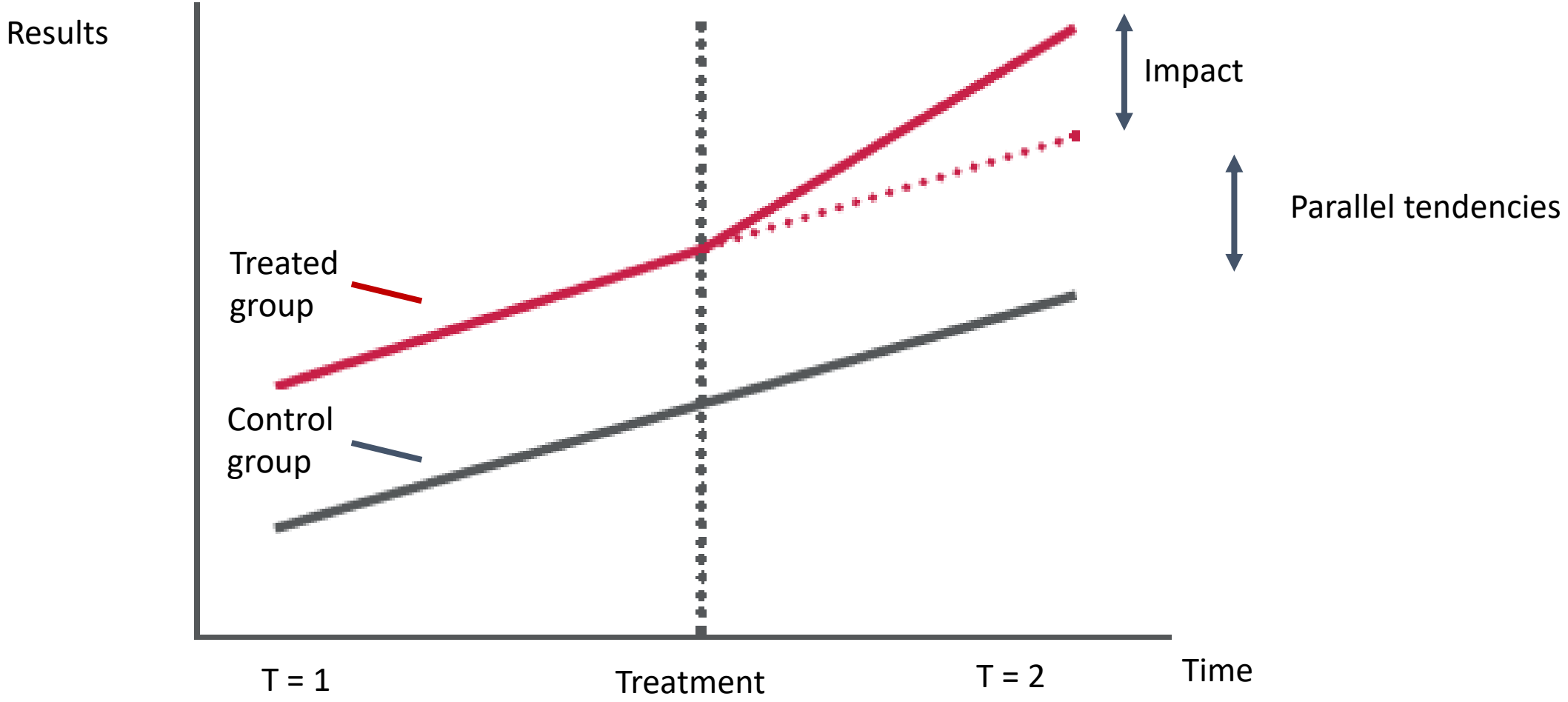
CUASI AND NON EXPERIMENTAL METHODS

- If we think about *growing*, it should be similar among groups: **parallel trends assumption**
  - This means that they don't need to be fully balance groups
  - We are using a linear trend

- We could check empirical data

DIFFERENCES IN DIFFERENCES - ASSUMPTIONS

CUASI AND NON EXPERIMENTAL METHODS



Tomado de
Valenzuela, Ramos y Andrade (2018) Evaluación de Impacto Centro de Desarrollo de Negocios en Chile. *Serie Sistemas Públicos N°15*. Santiago

DIFFERENCES IN DIFFERENCES - MODELS

CUASI AND NON EXPERIMENTAL METHODS

- How would this look in a table?:

$Y_{itd} = Y_i(D_i, T_i)$	Before (T = 0)	After (T = 1)	1 Dif = within
Control (D = 0)	$E(Y_{i00})$	$E(Y_{i01})$	
Treated (D = 1)	$E(Y_{i10})$	$E(Y_{i11})$	
2 Dif = between			

DIFFERENCES IN DIFFERENCES - MODELS

CUASI AND NON EXPERIMENTAL METHODS

- What about a regression?

- Why would I use a regression?

CSP · INGENIERÍA INDUSTRIAL UNIVERSIDAD DE CHILE

CENTRO DE SISTEMAS PÚBLICOS

IN4402: Applied statistics for management and economics

Difference in Difference Application to Energy Consumption

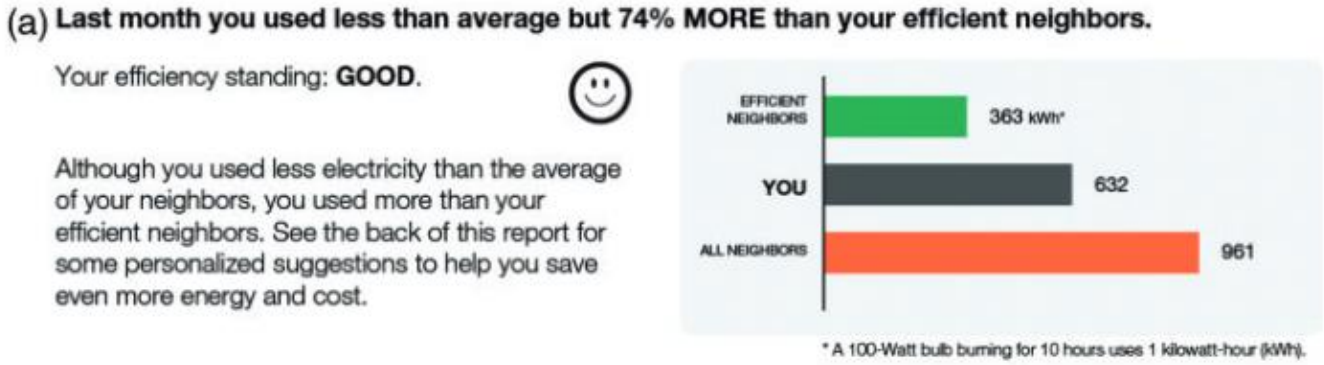
ANDRÉS FERNÁNDEZ

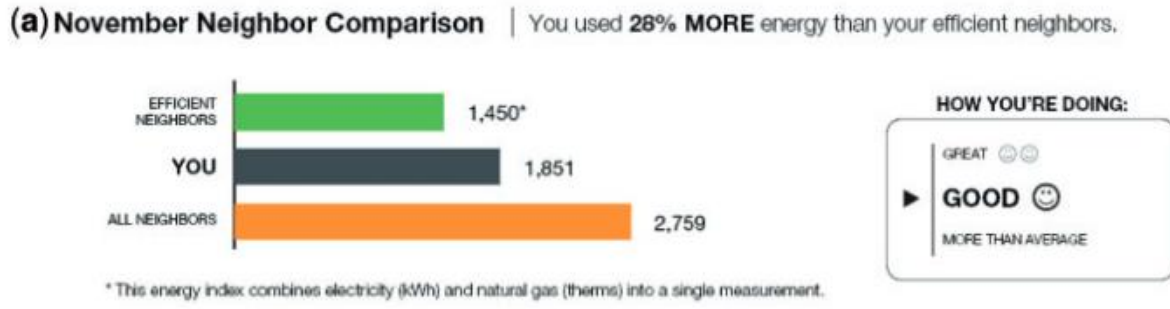
www.sistemaspublicos.cl

DIFFERENCES IN DIFFERENCES - EXAMPLE

CUASI AND NON EXPERIMENTAL METHODS

- An Energy Company was looking of ways of helping customers reduce their energy consumption
  - Randomly assigned *peer comparison* treatment to 170.000 houses
  - Provided monthly reports of energy consumption showing (to assigned people) reference to **neighbors' consumption**





Tomado de Ayres, I., Raseman, S., & Shih, A. (2013)

DIFFERENCES IN DIFFERENCES – EXAMPLE

CUASI AND NON EXPERIMENTAL METHODS

- Are the groups balanced?
- Does the parallel trends assumption hold?

Tomado de Ayres, I., Raseman, S., & Shih, A. (2013)

DIFFERENCES IN DIFFERENCES – EXAMPLE

CUASI AND NON EXPERIMENTAL METHODS

Table 1. SMUD OLS Regression of Log Household Monthly Average kWh/day

	No controls $n=2,262,815$	With controls $n=1,585,490$
Treatment household	−0.0013	−0.0006
Post April 2008 (first mailing)	−0.0183**	0.0784***
Treatment × post April 2008	−0.0199	−0.0215***
Cooling degree days (per billing cycle)		0.0021***
Heating degree days (per billing cycle)		0.0006***
Narrative template		0.0013
6 × 9 envelop		0.0009
Quarterly report recipients		−0.1163***
House sq. ft. (in 100s)		0.0034***
House age		0.0002**
Pool		0.0457***
Spa		−0.0030
House value (in $100,000s)		0.0042***
Gas heat		0.0317**
kWh/day usage in 2006		0.7595***
Tier 2 price in 2006		0.0296***
Length of residence		−0.0006***
Number of residents		0.0079***
Head of household age effects	No	Yes
Income quartile effects	No	Yes
Affluence effects[a]	No	Yes
Proprietary segment effects[b]	No	Yes
Census tracts fixed effects	No	Yes
Monthly fixed effects	No	Yes
Household fixed effects	No	No
R^2	0.0007	0.7061

Tomado de Ayres, I., Raseman, S., & Shih, A. (2013)

DIFFERENCES IN DIFFERENCES – EXAMPLE

CUASI AND NON EXPERIMENTAL METHODS

Table 1. SMUD OLS Regression of Log Household Monthly Average kWh/day

	No controls $n=2,262,815$	With controls $n=1,585,490$
Treatment household	−0.0013	−0.0006
Post April 2008 (first mailing)	−0.0183**	0.0784***
Treatment × post April 2008	−0.0199	−0.0215***
Cooling degree days (per billing cycle)		0.0021***
Heating degree days (per billing cycle)		0.0006***
Narrative template		0.0013
6 × 9 envelop		0.0009
Quarterly report recipients		−0.1163***
House sq. ft. (in 100s)		0.0034***
House age		0.0002**
Pool		0.0457***
Spa		−0.0030
House value (in $100,000s)		0.0042***
Gas heat		0.0317**
kWh/day usage in 2006		0.7595***
Tier 2 price in 2006		0.0296***
Length of residence		−0.0006***
Number of residents		0.0079***
Head of household age effects	No	Yes
Income quartile effects	No	Yes
Affluence effects[a]	No	Yes
Proprietary segment effects[b]	No	Yes
Census tracts fixed effects	No	Yes
Monthly fixed effects	No	Yes
Household fixed effects	No	No
R^2	0.0007	0.7061

Tomado de Ayres, I., Raseman, S., & Shih, A. (2013)

DIFFERENCES IN DIFFERENCES – EXAMPLE

CUASI AND NON EXPERIMENTAL METHODS

- Summary
  - Dif in dif estimator is good to provide causal explanations because reduces unobservable differences between groups
  - It needs to assume that growing trends are parallel
    - This can come from balanced or unbalanced groups
    - It can be assumed (by random assignment) or checked with data
  - Compares time differences assuming a linear behavior
  - Adding covariates help lowering standard errors