Day 3 - Part 3: Public Policy Design and Evaluation

New Climate Economy Training Course

World Resources Institute

July 2021

Outline

Randomized Control Trials (RCTs): Why randomize?

Use random assignment of the program to create a comparison group which "mimics" the counterfactual.

In our example, randomly assign hospitalization to hospitalized is the same as the effect of the hospitalization on a randomly chosen patient.

Also known in the literature as:

- Random assignment studies
- Randomized field trials
- Social Experiments
- Randomized experiments

RCTs: Why randomize?

The main reason to randomize, is that it solves the **selection problem**, making D_i independent of potential outcomes.

RCTs: Why randomize?

The main reason to randomize, is that it solves the **selection problem**, making D_i independent of potential outcomes.

$$
\begin{aligned}
E[Y_i|D_i=1]-E[Y_i|D_i=0] &= E[Y_{1i}|D_i=1]-E[Y_{0i}|D_i=0] \\
&= E[Y_{1i}|D_i=1]-E[Y_{0i}|D_i=1]
\end{aligned}
$$

RCTs: Why randomize?

The main reason to randomize, is that it solves the **selection problem**, making D_i independent of potential outcomes.

$$\begin{aligned} E[Y_i|D_i=1]-E[Y_i|D_i=0] &= E[Y_{1i}|D_i=1]-E[Y_{0i}|D_i=0] \\ &= E[Y_{1i}|D_i=1]-E[Y_{0i}|D_i=1] \end{aligned}$$

In fact, given random assignment, previous expression can be simplified to:

$$\begin{aligned} E[Y_{1i}|D_i=1]-E[Y_{0i}|D_i=1] &= E[Y_{1i}-Y_{0i}|D_i=1] \\ &= E[Y_{1i}-Y_{0i}] \end{aligned}$$

RCTs: Are they always feasible?

RCTs is certainly the best way (theoretically) to deal with selection bias. However, in practice, the implementation of an RCT is no at easy as it seems:

- RCTs (in most cases) requires large budgets and a lof of time, which is not always available for policy makers.
- Perfect randomization is almost never possible.
  - There are several departures from that scenario that policy makers should consider when designing RCTs to estimate a causal effect(an interesting topic for another course!)[1]

[1] Maybe the best reference if you want to delve into this subject is JPAL: https://www.povertyactionlab.org/

RCTs: Are they always feasible?

- RCTs compare the difference between treatment and comparison populations in a given area → They are not able to pick up general equilibrium effects.
  - Particularly important for assessing the welfare implications of scaling up a program or a macroeconomic policy!
- The results are not always replicable or generalizable to other contexts.

Practical Case 1: Nyqvist and Jayachandran (2017)

To illustrate how RCTs worked in practice, we are going to use:

Nyqvist, M. B. and S. Jayachandran (2017): “Mothers Care More,But Fathers Decide: Educating Parentsabout Child Health in Uganda”, *American Economic Review: Papers and Proceedings*.

Practical Case 1: Nyqvist and Jayachandran (2017)

- Intrahousehold decision making evidence often finds that fathers have more decision-making power than mothers, buth mothers put more weight on children's well-being.

Practical Case 1: Nyqvist and Jayachandran (2017)

- Intrahousehold decision making evidence often finds that fathers have more decision-making power than mothers, buth mothers put more weight on children's well-being.
- Policy response have been focused on trying to shift decision-making power towards mother → Not always feasible or advisable.

Practical Case 1: Nyqvist and Jayachandran (2017)

- There is a trade-off when targeting policies to improve children's well-being (especially stark in developing countries where women have low bargaining power).
  - Fathers have more power to change household behavior in ways to help children.
  - Mothers have a stronger desire to do so.

Intervention: Implementation of nutrition classes that provided parents of young children with knowledge to improve their children's health in Uganda.

Practical Case 1: Nyqvist and Jayachandran (2017)

Balance

TABLE 1—BASELINE CHARACTERISTICS

	Women	Men
Health knowledge	0.114	–0.118
	(0.968)	(1.018)
Decision-making power	0.295	0.441
	(0.194)	(0.209)

Note: Sample means, with standard deviations in parentheses, are reported for the combined WHN, MHN, and control groups.

Practical Case 1: Nyqvist and Jayachandran (2017)

Results: Impact on attendance and knowledge

TABLE 2—IMPACTS OF HEALTH CLASSES ON ATTENDANCE AND KNOWLEDGE

	Attendance	Participant's health knowledge	Spouse's health knowledge
WHN classes	0.756	0.289	0.018
	(0.012)	(0.039)	(0.042)
MHN classes	0.577	0.220	0.081
	(0.016)	(0.045)	(0.040)
p-value	0.000	0.222	0.243
Observations	4,182	5,258	5,279

Notes: OLS coefficients are reported with standard errors, clustered at the village level, shown in parentheses. The *p*-value reported is for a test of equality between the WHN and MHN coefficients. All regressions include stratum and district fixed effects, and baseline maternal health index, gender norms index, and log HH income. Columns 2 and 3 also control for the baseline knowledge of the participant and spouse, and the respondent's gender. When a control variable has a missing value, we impute with the village mean; flags for missing values for each variable are included as control variables. The standard deviation of attendance in the control group is 0.0; the standard deviations of participant's and spouse's behavior are both 1.0.

Practical Case 1: Nyqvist and Jayachandran (2017)

Problem set

- Replicate balance table from the paper.
- Replicate columns 2 and 3 (impact on health knowledge) from the paper.

Practical Case 2: Blattman and Dercon (2018)

The second paper we are going to review is:

Blattman, C. and Dercon (2018): "The Impacts of Industrial and Entrepreurial Work on Income and Health: Experimental Evidence from Ethiopia", *American Economic Journal: Applied Economics*.

Practical Case 2: Blattman and Dercon (2018)

- Empirically, there is a large body of observational evidence that suggest that formal firms pay premium wages (espcially large, foreign-owned or exporting firms).
- Also, women are commonly employed in low-skill firms, and there is evidence that working in textile factories or other export manufacturers raises women's status in the household.
- In this context, industrial jobs may be attractive only compared to poor people's largely informal alternatives.
- However, it is unclear what opportunities and risks industrial jobs offer to workers relative to their informal opportunities.

Practical Case 2: Blattman and Dercon (2018)

In this context, authors uses a case study and small-scale experiment in Ethiopia to investigate:

- What are the relative qualities of informal and industrial work at this early stage of industrialization?
- Are the benefits of risks to the choice of one occupation over the other?
- How does the quality of self-employment options affect this occupation choice?

Practical Case 2: Blattman and Dercon (2018)

Balance

Table 2—Baseline Summary Means and Test of Randomization Balance

Baseline covariate ($n = 947$)	Control mean ($n = 358$) (1)	Balance test (OLS): Job—Control: Diff. (2)	Job—Control: p value (3)	Entrepreneur—Control: Diff. (4)	Entrepreneur—Control: p value (5)
Female	0.80				
Age	22.02	–0.12	0.68	–0.14	0.63
Unmarried	0.81	–0.06	0.07	–0.04	0.25
Muslim	0.05	–0.00	0.90	0.00	0.98
Household size	4.35	–0.13	0.45	–0.14	0.45
Household head	0.23	0.04	0.25	0.00	0.96
Proportion household dependents	0.43	–0.00	0.98	–0.00	0.96
Total years of education and training	9.31	–0.20	0.34	–0.02	0.92
Executive function, z-score	0.11	–0.18	0.01	–0.13	0.08
Weekly cash earnings (2010 birr)	9.57	0.59	0.81	–1.44	0.57
Durable assets, z-score	0.09	–0.11	0.13	–0.13	0.06
Ever worked in a large firm	0.27	–0.03	0.43	0.05	0.18
Average weekly hours of work	7.52	–0.09	0.94	–0.36	0.80
No work in past 4 weeks	0.68	0.01	0.86	–0.01	0.76
Highest–lowest earnings, past month	181.38	39.44	0.05	15.84	0.33
Could borrow 3,000 birr	0.31	0.04	0.27	–0.00	0.98
Ability to do activities of daily life (0–15)	14.32	0.09	0.40	0.10	0.31
Disability (great difficulty at >1 ADL)	0.01	–0.01	0.26	–0.00	0.77
Risk aversion, z score	0.01	0.05	0.55	0.10	0.20
Future orientation, z-score	0.10	–0.06	0.45	–0.03	0.73
Locus of control index, z-score	–0.04	0.04	0.62	0.13	0.12
Self-esteem index, z-score	–0.05	0.03	0.75	0.06	0.42
Family relations index, z-score	–0.05	–0.02	0.77	0.07	0.35
Friends and neighbors relations index	–0.01	–0.05	0.49	0.00	0.95
Change in subjective well-being, past yr.	0.22	0.20	0.03	0.09	0.33
Symptoms of depression, z-score	–0.02	0.02	0.82	0.01	0.94
Symptoms of anxiety, z-score	–0.04	0.05	0.50	–0.01	0.92
Aggressive or hostile behaviors, z-score	0.04	–0.06	0.44	–0.13	0.11
Conscientiousness index, z-score	–0.00	0.01	0.89	0.04	0.65
Years experience, private firm	0.34	0.17	0.08	0.22	0.02
Years experience, workshop	0.01	0.00	0.73	0.01	0.27
Years experience, government/NGO	0.08	–0.02	0.67	0.02	0.73
Probability of better job, next month	0.68	0.01	0.47	0.01	0.72
Probability of full-time work, next month	0.55	0.01	0.74	0.03	0.17
p-value from F-test of joint significance		0.04		0.01	
Observations		662		643	

Notes: Medians are imputed for baseline variables with missing observations. Treatment and control group differences are calculated using an OLS regression of each covariate on treatment indicators plus block (cohort-gender) fixed effects. Balance tests for the female dummy are omitted because randomization was blocked on gender. Standard errors are heteroskedastic-robust.

Practical Case 2: Blattman and Dercon (2018)

Results: Impacts of the Job Offer and Entrepreneurship Program on Employment and Income

TABLE 6—IMPACTS OF THE JOB OFFER AND ENTREPRENEURSHIP PROGRAM ON EMPLOYMENT AND INCOME

		ITT estimate ($N = 1{,}587$)							
		Job offer			Entrepreneurship program			Job—Entrepreneur	
Outcome	Control mean (1)	Coeff. (2)	SE (3)	Adj. *p*-val. (4)	Coeff. (5)	SE (6)	Adj. *p*-val. (7)	Coeff. (8)	SE (9)
Employment and occupational choice, *z*-score	−0.04	0.078	[0.074]	0.909	0.040	[0.076]	0.849	0.038	[0.079]
Hours work/week, past two weeks	26.39	0.997	[1.895]		3.506	[1.892]		−2.509	[2.010]
Factory labor	7.46	3.017	[1.380]		−4.104	[1.169]		7.122	[1.287]
Farm wage labor	3.07	1.817	[0.914]		−1.480	[0.744]		3.297	[0.865]
Smallholder farming	0.82	−0.258	[0.323]		1.480	[0.398]		−1.738	[0.430]
Petty business	4.04	−0.878	[0.978]		5.381	[1.379]		−6.259	[1.353]
Skilled trades	1.59	−0.736	[0.449]		−0.576	[0.483]		−0.160	[0.408]
Casual nonfarm labor	2.18	−0.954	[0.568]		0.746	[0.770]		−1.700	[0.662]
Low-skill salaried labor	4.19	0.064	[1.095]		−0.412	[0.956]		0.476	[0.984]
Medium-skill salaried labor	1.21	−0.415	[0.420]		1.604	[0.590]		−2.018	[0.545]
Other work	2.27	−0.085	[0.694]		0.489	[0.738]		−0.574	[0.800]
Unemployed in past two weeks	0.34	−0.013	[0.033]		−0.082	[0.032]		0.068	[0.034]
SD of hours/week	16.44	−1.306	[1.342]		3.949	[1.476]		−5.254	[1.458]
Income, *z*-score	−0.01	0.014	[0.052]	0.949	0.150	[0.058]	0.029	−0.135	[0.057]
Weekly earnings, 2010 birr	34.23	3.049	[4.479]		12.005	[5.463]		−8.956	[5.426]
Earnings per hour, 2010 birr	1.46	−0.020	[0.186]		0.153	[0.200]		−0.173	[0.200]
SD of weekly earnings	56.01	4.107	[7.600]		3.769	[8.263]		−0.079	[0.067]
Household-level durable consumption assets, *z*-score†	0.07	−0.071	[0.069]		0.009	[0.067]		−56.289	[33.861]
Household-level nondurable consumption, 2010 birr	664.46	20.548	[34.653]		76.837	[35.492]		−0.282	[0.074]
Household-level durable productive assets, *z*-score†	−0.12	0.049	[0.068]		0.331	[0.077]		−0.282	[0.074]

Notes: Columns 2 to 5 report the results of an OLS regression of each outcome on treatment indicators, baseline covariates, and cohort-gender fixed effects. Survey responses at 11 and 13 months are pooled. Standard errors are robust and clustered by respondent. *p*-values are adjusted for family outcomes using the Westfall-Young approach described in Section V. Some outcomes contain fewer observations than the listed number of observations because a very small number of respondents were not asked certain questions.

†Denotes outcome variables that were measured during only one of the endline surveys.

Non or quasi-experimental methods

Even though RCTs represents the “ideal” on how to get rid of selection bias in impact evaluation, there are other methods that can also be effective. However...

Non or quasi-experimental methods

Even though RCTs represents the “ideal” on how to get rid of selection bias in impact evaluation, there are other methods that can also be effective. However...

- They are effective only if the specific conditions needed for that method's assumption to hold exists.
- In other words, they have the limitation that they need certain conditions to be valid that not always apply.

Matching

The first estimator we are going to study is **matching**

- The ideal comparison group is selected such that matches the treatment group using either a comprehensive baseline survey or time invariant characteristics.
- The matches are selected on the basis of similarities in observed characteristics.
- This assumes no selection bias based on unobserved characteristics!
    - All such differences must be in the data in order for the match to produce a valid estimate of project impacts.

Matching

Methods on Controlling Confounding Variables

- Exact matching method
  - Match each respondent in the treated group with a respondent in the untreated group that have the same values of all variables.
  - It is difficult to find an exact match for respondents when many covariates are involved.
- Stratification
  - Divide the sample into different homogenous strata and conduct analysis with each of them.
  - The analysis results across stratum can be aggregated to obtain the average result of the whole sample.

Matching

Methods on Controlling Confounding Variables

- Stratification
  - Hard to stratify on many covariates simultaneously.
  - It will also reduce sample size.
- Regression adjustment
  - Include possible confounding variables into the analysis.
  - May not work if confounding variables have different functional forms for the treated group than for the untreated group.

Propensity Score Matching (PSM)

What is a propensity score?

- A propensity score is the conditional probability of a unit being assigned to a particular study condition (treatment or comparison) given a set of observed covariates or characteristics X_i.
- Match on the basis of the propensity score:

$$P(X_i) = Pr(D_i|X_i)$$

  - D_i indicates participation in project.
  - Instead of attempting to create a match for each participant with exactly the same value of X_i, we can instead match on the probability of participation.

Propensity Score Matching (PSM)

- Matching on a single index (propensity score), reflecting the probability of participation, could achieve consistent estimates of the treatment effect in the same way as matching on all covariates

Conditions:

- This single index summarises all the relevant information contained in the covariates X_i.
- Matching on this index is equivalent to matching on the X_i, i.e. for a given value of the index the distribution of X_i should be the same for participants and non-participants (This is called the propensity score theorem).

Propensity Score Matching (PSM)

Assumptions

1. **CIA:**
   - There are no systematic differences between participants and non-participants in terms of unobserved characteristics that may affect the outcome.
   - All the variables that affect simultaneously T and Y are observed.
2. **Common support:**
   - In both groups there are individuals with similar propensity scores.
   - Matching is feasible.
3. **Propensity score balances the covariates:**
   - Similar propensity scores are based on similar observed X.

PSM step-by-step

1. Estimation of the propensity score
2. Check the assumptions: common support
3. Match participants with non-participants
4. Check the assumptions: covariate's balance
5. Compute the average treatment effect
6. Compute the standard error of the treatment effect

Pros and cons of matching

PROS

- More focus on the selection process and on the underlying assumptions
- Imposition of the common support ensures comparability
- Versatility:
  - Allows to estimate heterogeneous effects (by sub-group)
  - Allows to put more emphasis on specific variables, on which exact matching can be done (e.g. region, gender)
  - Allows the estimation of multiple treatments: different treatment levels or types of participation can be compared

Pros and cons of matching

CONS

- "Data-hungry" method, more efficient methods under CIA exist .
- Requires strong robustness and sensitivity analysis
- CIA is a Strong assumption:
  - Impossible to verify, so bias stemming from unobservables **can never** be ruled out
  - Matching is only as good as the characteristics used for matching
  - *Matching becomes much better in combination with other techniques, such as: Exploiting baseline data for matching and using difference-in-difference strategy

Practical case 1: Dehejia and Wahba (1999)[2]

In this application used Dehejia and Wahba (1999), we are going to study a training program applied during 1976 and 1977 in USA, and analyse the relation of this program with income

The purpose of this exercise is to:

- Compute the propensity the propensity score (using the command `pscore` or the commands `logit` or `probit`.
- Analyse the distribution of common support (graphically) using the commands: `psgraph` or `twoway kdensity`
- Compute the average treatment effect (att) using nearest neighbor or radius matching (use `psmatch2` or `attnd` commands)
- Test the balance of the common support using the command `pstest`.

[2] Details are explained on the problem set sheet and on the do-file.

Practical case 2: Social Protection Survey (EPS) 2006 of Chile[3]

In this second application, we are going to use the Social Protection Survey (EPS) of 2006 implemented in Chile.

The purpose of this exercise is to:

- Compute the propensity the propensity score (using the command `pscore` or the commands `logit` or `probit`.
- Analyse the distribution of common support (graphically) using the commands: `psgraph` or `twoway kdensity`
- Compute the average treatment effect (att) using nearest neighbor or radius matching (use `psmatch2` or `attnd` commands)
- Test the balance of the common support using the command `pstest`.

[3] Details are explained on the problem set sheet and on the do-file.

Difference in Difference (DiD)

Remember that a limitation of the PSM is related to matching on observables

- If unobserved characteristics are important, we can identify a causal effect using, for example, instrumental variables (hard to find in practice!).
- Difference-in-difference exploits the time or cohort dimension, and allows accounting for unobservable but fixed characteristics.

Difference in Difference (DiD

Time of the intervention

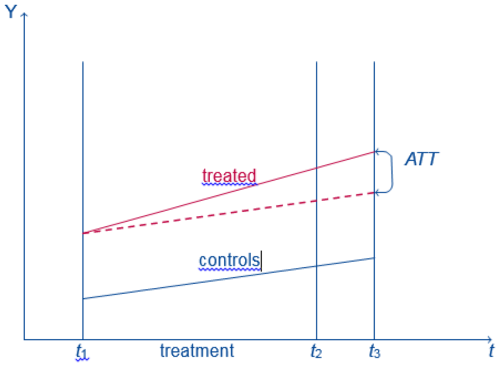
Remember that a limitation of the PSM is related to matching on observables

- Recall $t_1 < t_2 < t_3$
- Intervention between t_1 and t_2.
- Outcome measured at some time t_3 after completion of the intervention.

Difference in Difference (DiD

Time of the intervention

Whenever one has access to **longitudinal data** or **repeated cross-sections**.



Source: EUROSTAT.

Difference in Difference (DiD

Assumptions

- The key assumption for DiD, is that the outcome in treatment and control group would follow the same time trend in the absence of the treatment (parallel worlds).
- This does not mean that they need to have the same average outcome!
- The **Common trend** assumption is difficult to verify but one could uses pre-treatment data to show that the trends are the same before treatment takes place.
- Even if pre-trends are the same one still has to worry about other policies changing at the same time.

Difference in Difference (DiD

Example[4]

- Suppose you are interested in the effect of minimum wages on employment.
- In a competitive labour market, increases in the minimum wage would move us up a downward-sloping labour demand curve $\rightarrow$ Employment would fall!

[4] This example was taken from EUROSTAT.

Difference in Difference (DiD

Example

- Card Krueger (1994) use the change in the minimum wage in the state of New Jersey to check whether an increase in minwage causes a decrease in employment.
- In February 1992 NJ increased the state minimum wage from 4.25 to 5.05 dollars. Pennsylvania's minimum wage stayed at 4.25.
- They surveyed about 400 fast food stores both in NJ and in PA both before and after the minimum wage increase in NJ.

Difference in Difference (DiD)

Example

Average employment per store before and after the New Jersey minimum wage increase.

	PA	NJ	PA-NJ
FTE employment before	23.33	20.44	-2.89
	(1.35)	(0.51)	(1.44)
FTE employment after	21.17	21.03	-0.14
	(0.94)	(0.52)	(1.07)
Change in mean FTE	-2.16	0.59	2.76
employment	(1.25)	(0.54)	(1.36)

Source: EUROSTAT.

Difference in Difference (DiD

Example

- **Remember:** Our assumption is that the employment trends would be the same in both states in absence of the treatment.
- The common trend assumption could be investigated further using data from previous periods (Card and Krueger 2000, repeated their 1992 study collecting payroll data and ran a new experiment in 1996)
- **Conclusion:** too many swings in the data so that PA was not a good comparison group for NJ.

Practical Case 1: Ayres and Donohue III (2003)

- In this study, authors analyse U.S. states that have enacted 'shall-issue' laws which allow citizens to carry concealed weapons.
- Particularly, using the same data they use, we are going to investigate the effect of shall-issue laws on violent crime rates.

Practical Case 1: Ayres and Donohue III (2003)

Variables

Variable Definitions

Variable	Definition
vio	violent crime rate (incidents per 100,000 members of the population)
rob	robbery rate (incidents per 100,000)
mur	murder rate (incidents per 100,000)
shall	= 1 if the state has a shall-carry law in effect in that year = 0 otherwise
incarc_rate	incarceration rate in the state in the previous year (sentenced prisoners per 100,000 residents; value for the previous year)
density	population per square mile of land area, divided by 1000
avginc	real per capita personal income in the state, in thousands of dollars
pop	state population, in millions of people
pm1029	percent of state population that is male, ages 10 to 29
pw1064	percent of state population that is white, ages 10 to 64
pb1064	percent of state population that is black, ages 10 to 64
stateid	ID number of states (Alabama = 1, Alaska = 2, etc.)
year	Year (1977-1999)

Practical Case 1: Ayres and Donohue III (2003)

Exercise[5]

In this exercise, you are asked to:

- Analyse the relation between violent crime rate and shall-carry law using a simple regression
- Set panel data (by state and year) and run fixed effects regressions on the same model of violent crime rate and shall-carry law.

[5]Details on this exercise can be found on the problem set sheet and on the do-file.

Practical Case 2: Engelhardt and Gruber (2011)

In this paper, authors analyse the *Medicare Modernization Act* of 2003

- Better known as the *Part D* prescription-drug benefit to the Medicare program.
- Represents the single most significant expansion of public insurance programs in the United States in the past 40 years.

Practical Case 2: Engelhardt and Gruber (2011)

Objective: Increased the costs of the Medicare program by over 10 percent in order to provide, for the first time, prescription-drug coverage to enrollees.

- Despite the size the of the program, at the time the paper was published, there was no much evidence about its effectiveness.
- Particularly, there was no evidence on the program's success in providing financial security to the nation's elders.

Practical Case 2: Engelhardt and Gruber (2011)

Data: Medical Expenditure Panel Survey (MEPS)

- 2002-2005 and 2007 waves: before and after the implementation of this program.
- Data contains information on insurance coverage, but also on prescription-drug expenditures by source of payment (including out-of-pocket).

Practical Case 2: Engelhardt and Gruber (2011)

Questions:

- Whether the passage of Part D was associated with increased prescription-drug coverage among the elderly, compared to the near-elderly (those just below 65).
- Evaluate the impact of Part D on the prescription-drug spending by payment source among the elderly.
- Evaluate the impact of Part D on the distribution of out-of-pocket prescription-drug spending among the elderly.

Practical Case 2: Engelhardt and Gruber (2011)

Question: Whether the passage of Part D was associated with increased prescription-drug coverage among the elderly, compared to the near-elderly (those just below 65).

Practical Case 2: Engelhardt and Gruber (2011)

Question: Whether the passage of Part D was associated with increased prescription-drug coverage among the elderly, compared to the near-elderly (those just below 65).

Results: Increased by 10 percent points. This suggests that Part D to a large extent crowded out of other forms of prescription-drug coverage.

Practical Case 2: Engelhardt and Gruber (2011)

Question: Evaluate the impact of Part D on the prescription-drug spending by payment source among the elderly.

Practical Case 2: Engelhardt and Gruber (2011)

Question: Evaluate the impact of Part D on the prescription-drug spending by payment source among the elderly.

Results: There was an overall increase of $525 per year spent on drygs as a result of Part D. Yet total public expenditure on prescription drugs rose by $2,100, so that crowd-out was on the order of 75%.

Practical Case 2: Engelhardt and Gruber (2011)

Question: Evaluate the impact of Part D on the distribution of out-of-pocket prescription-drug spending among the elderly.

Practical Case 2: Engelhardt and Gruber (2011)

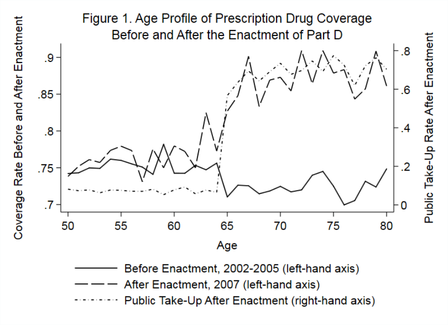
Question: Evaluate the impact of Part D on the distribution of out-of-pocket prescription-drug spending among the elderly.

Results: Part D led to sizeable decline in out-of-pocket drug spending, and this decline was concentrated in the top of the expenditure distribution.

Practical Case 2: Engelhardt and Gruber (2011)

Exercises

Exercise 1[6]**:** Replicate figure 1 of the paper - Graphical evidence on the age of profile of prescription-drug coverage before and after the enanctment of Part D



[6] Details are explained in the problem set sheet and on the do-file.

Practical Case 2: Engelhardt and Gruber (2011)

Exercises

Exercise 2[7]**:** Difference-in-Difference estimates to Medicare Part D Law Change on Prescription-Drug coverage on any source by age group in the 2002-2005 and 2007 MEPS.

TABLE 3—DIFFERENCE-IN-DIFFERENCE ESTIMATES OF MEDICARE PART D LAW CHANGE ON PRESCRIPTION-DRUG COVERAGE FROM ANY SOURCE BY AGE GROUP IN THE 2002–2005 AND 2007 MEPS (*Standard errors in parentheses*)

Group/year	Before Part D (1)	After Part D (2)	Time difference For groups (3)
Panel A. Any coverage			
Age 65–70	0.722 (0.00798)	0.859 (0.0106)	0.137 (0.0132)
Age 60–64	0.750 (0.00795)	0.784 (0.0124)	0.0342 (0.0147)
Difference-in-difference			0.103 (0.0198)
Panel B. Public coverage			
Age 65–70	0.260 (0.00782)	0.657 (0.0144)	0.397 (0.0164)
Age 60–64	0.0796 (0.00499)	0.0760 (0.00781)	−0.00365 (0.00927)
Difference-in-difference			0.401 (0.0188)

Notes: Each cell gives the coverage rate among 60–70 year olds for prescription-drug coverage from any source for each of the table's groups. Standard errors clustered by household and age group (under 65 and 65 and older) are shown in parentheses.

[7] Details are explained in the problem set sheet and on the do-file.

DiD and PSM

One of the advantages of matching estimators, is that they becomes much better in combination with other techniques.

- If we collected baseline data, we can exploit this source for matching and then use a DiD strategy.
- If an assignment rule exists for a project, we can match on this rule.

DiD and PSM

One of the advantages of matching estimators, is that they becomes much better in combination with other techniques.

- If we collected baseline data, we can exploit this source for matching and then use a DiD strategy.
- If an assignment rule exists for a project, we can match on this rule.

DiD and PSM

Implementation

The implementation of this strategy implies:

- Estimate differences using matching techniques **before** treatment (remember using `psmatch2` and `pstest` commands to compute the difference and the validity of the common support.)
- You can assume fixed differences over time and take before-after difference in treatment and control groups as usual in DiD.

DiD and PSM

Exercices

A practical exercise that combine these two techniques will be on the problem set of **day 5**, so be sure you understand how to implement PSM and DiD.

References I

Ayres, I. and J. J. Donohue III (2003). "Shooting Down the 'More Guns Less Crime". In: *Stanford Law Review*, pp. 1193–1312.

Blattman, Christopher and Stefan Dercon (2018). "The Impacts of Industrial and Entrepreneurial Work onIncome and Health: Experimental Evidence from Ethiopia". In: *American Economic Journal: Applied Economics*.

Dehejia, Rajeev H. and Sadek Wahba (1999). "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs". In: *Journal of the American Statistical Association* 94.448, pp. 1053–1062.

Engelhardt, Gary V and Jonathan Gruber (2011). "Medicare Part D and the financial protection of the elderly". In: *American Economic Journal: Economic Policy* 3.4, pp. 77–102.

References II

Nyqvist, Martina BjÖrkman and Seema Jayachandran (2017). “Mothers Care More, But Fathers Decide: Educating Parentsabout Child Health in Uganda”. In: *American Economic Review: Papers and Proceedings*.