< Back to Columbia University Mailman School of Public Health

Population Health Methods

... Difference-in-Difference Estimation

Difference-in-Difference Estimation

Overview

The difference-in-difference (DID) technique originated in the field of econometrics, but the logic underlying the technique has been used as early as the 1850's by John Snow and is called the 'controlled before-and-after study' in some social sciences.

Description

DID is a quasi-experimental design that makes use of longitudinal data from treatment and control groups to obtain an appropriate counterfactual to estimate a causal effect. DID is typically used to estimate the effect of a specific intervention or treatment (such as a passage of law, enactment of policy, or large-scale program implementation) by comparing the changes in outcomes over time between a population that is enrolled in a program (the intervention group) and a population that is not (the control group).

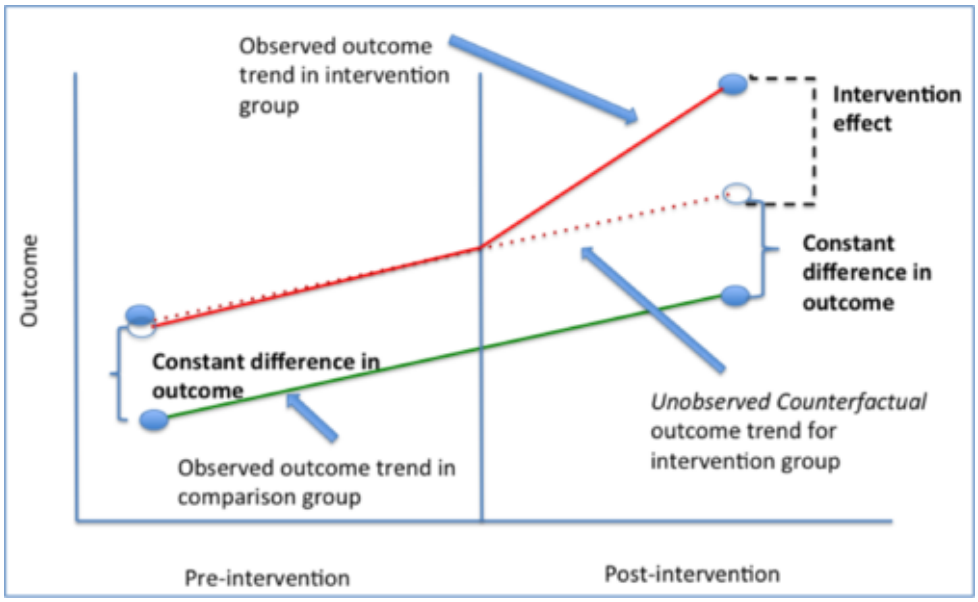


Figure 1. Difference-in-Difference estimation, graphical explanation

DID is used in observational settings where exchangeability cannot be assumed between the treatment and control groups. DID relies on a less strict exchangeability assumption, i.e., in absence of treatment, the unobserved differences between treatment and control groups arethe same overtime. Hence, Difference-in-difference is a useful technique to use when randomization on the individual level is not possible. DID requires data from pre-/post-intervention, such as cohort or panel data (individual level data over time) or repeated cross-sectional data (individual or group level). The approach removes biases in post-intervention period comparisons between the treatment and control group that could be the result from permanent differences between those groups, as well as biases from comparisons over time in the treatment group that could be the result of trends due to other causes of the outcome.

Causal Effects (Ya=1 – Ya=0)

DID usually is used to estimate the treatment effect on the treated (causal effect in the exposed), although with stronger assumptions the technique can be used to estimate the Average Treatment Effect (ATE) or the causal effect in the population. Please refer to Lechner 2011 article for more details.

ASSUMPTIONS

In order to estimate any causal effect, three assumptions must hold: exchangeability, positivity, and Stable Unit Treatment Value Assumption (SUTVA)1

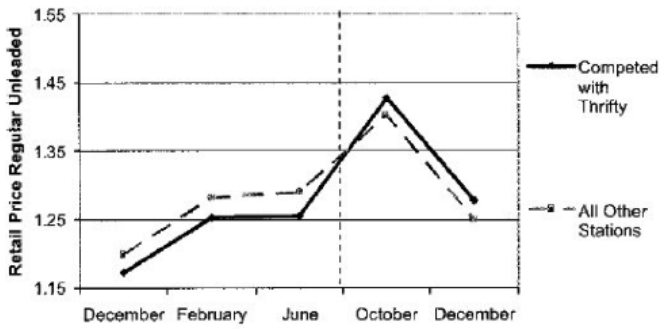
. DID estimation also requires that:

- Intervention unrelated to outcome at baseline (allocation of intervention was not determined by outcome)
- Treatment/intervention and control groups have Parallel Trends in outcome (see below for details)

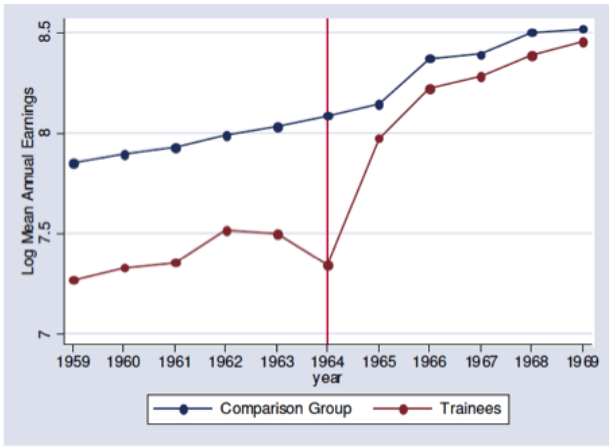
- Composition of intervention and comparison groups is stable for repeated cross-sectional design (part of SUTVA)
- No spillover effects (part of SUTVA)

Parallel Trend Assumption

The parallel trend assumption is the most critical of the above the four assumptions to ensure internal validity of DID models and is the hardest to fulfill. It requires that in the absence of treatment, the difference between the 'treatment' and 'control' group is constant over time. Although there is no statistical test for this assumption, visual inspection is useful when you have observations over many time points. It has also been proposed that the smaller the time period tested, the more likely the assumption is to hold. Violation of parallel trend assumption will lead to biased estimation of the causal effect.



Meeting the Parallel Trend Assumption 2



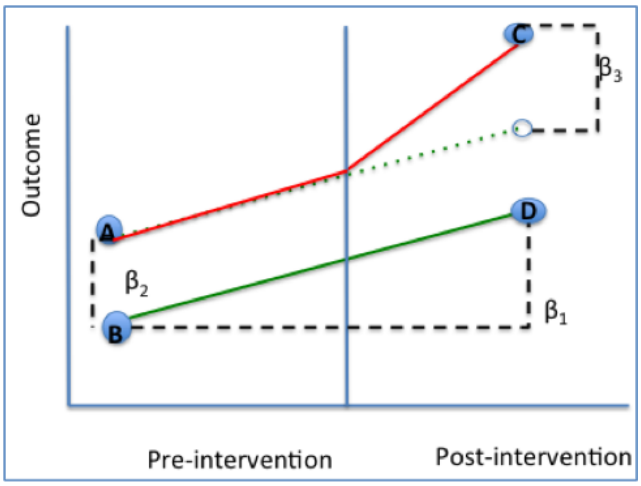
Violation of the Parallel Trend Assumption 3

Regression Model

DID is usually implemented as an interaction term between time and treatment group dummy variables in a regression model.

Y= β0 + β1*[Time] + β2*[Intervention] + β3*[Time*Intervention] + β4*[Covariates]+ε

Coefficient	Calculation	Interpretation
β_0	B	Baseline average
β_1	D-B	Time trend in control group
β_2	A-B	Difference between two groups pre-intervention
β_3	(C-A)-(D-B)	Difference in changes over time



Strengths and Limitations

Strengths

- Intuitive interpretation
- Can obtain causal effect using observational data if assumptions are met
- Can use either individual and group level data
- Comparison groups can start at different levels of the outcome. (DID focuses on changerather than absolute levels)
- Accounts for change/change due to factors other than intervention

Limitations

- Requires baseline data & a non-intervention group
- Cannnot use if intervention allocation determined by baseline outcome
- Cannot use if comparison groups have different outcome trend (Abadie 2005 has proposed solution)
- Cannot use if composition of groups pre/post change are not stable

BEST PRACTICES

- Be sure outcome trend did not influence allocation of the treatment/intervention
- Acquire more data points before and after to test parallel trend assumption
- Use linear probability model to help with interpretability
- Be sure to examine composition of population in treatment/intervention and control groups before and after intervention
- Use robust standard errors to account for autocorrelation between pre/post in same individual
- Perform sub-analysis to see if intervention had similar/different effect on components of the outcome

Epi6 in-class presentation April 30, 2013

1. Rubin, DB. Randomization Analysis of Experimental Data in the Fisher Randomization Test. Journal American Statistical Association.1980.
2. Adapted from "Vertical Relationships and Competition in Retail Gasoline Markets," 2004 (Justine Hastings)
3. Adapted from "Estimating the effect of training programs in earnings, review of economics and statistics", 1978 (Orley Ashenfelter)

Readings

TEXTBOOKS & CHAPTERS

- Mostly Harmless Econometrics: Chapter 5.2 (pg 169-182)

Angrist J., Pischke J.S. 2008. Mostly Harmless Econometrics, Princeton University Press, NJ.
http://www.mostlyharmlesseconometrics.com/
This chapter discusses DID in the context of the technique's original field, Econometrics. It gives a good overview of the theory and assumptions of the technique.

- WHO-Impact Evaluation in Practice: Chapter 6.

http://siteresources.worldbank.org/EXTHDOFFICE/Resources/5485726-1295455628620/Impact_Evaluation_in_Practice.pdf
Accessed on February 9th 2013.
This publication gives a very straightforward review of DID estimation from a health program evaluation perspective. There is also a section on best practices for all of the methods described.

METHODOLOGICAL ARTICLES

- Bertrand, M., Duflo, E., & Mullainathan, S. How Much Should We Trust Differences-in-Differences Estimates? Quarterly Journal of Economics. 2004.

This article, critiquing the DID technique, has received much attention in the field. The article discusses potential (perhaps severe) bias in DID error terms. The article describes three potential solutions for addressing these biases.

- Cao, Zhun et al. Difference-in-Difference and Instrumental Variabels Approaches. An alternative and complement to propensity score matching in estimating treatment effects.CER Issue Brief: 2011.

An informative article that describes the strengths, limitations and different information provided by DID, IV, and PSM.

- Lechner, Michael. The Estimation of Causal Effects by Difference-in-Difference Methods. Dept of Economics, University of St. Gallen. 2011.

  This paper offers an in-depth perspective on the DID approach and discusses some of the major issues with DID. It also provides a substantial amount of information on extensions of DID analysis including non-linear applications and propensity score matching with DID. Applicable use of potential outcome notation included in report.

- Norton, Edward C. Interaction Terms in Logitand Probitmodels. UNC at Chapel Hill. Academy Health 2004.

  These lecture slides offer practical steps to implement DID approach with a binary outcome. The linear probability model is the easiest to implement but have limitations for prediction. Logistic models require an additional step in coding to make the interaction terms interpretable. Stata code is provided for this step.

- Abadie, Alberto. Semiparametric Difference-in-Difference Estimators. Review of Economic Studies. 2005

  This article discusses the parallel trends assumption at length and proposes a weighting method for DID when the parallel trend assumption may not hold.

APPLICATION ARTICLES

Health Sciences

Generalized Linear Regression Examples:

- Branas, Charles C. et al. A Difference-in-Differences Analysis of Health, Safety, and Greening Vacant Urban Space. American Journal of Epidemiology. 2011.
- Harman, Jeffrey et al. Changes in per member per month expenditures after implementation of Florida's medicaid reform demonstration. Health Services Research. 2011.
- Wharam, Frank et al. Emergency Department Use and Subsequent Hospitalizations Among Members of a High-Deductible Health Plan. JAMA. 2007.

Logistic Regression Examples:

- Bendavid, Eran et al. HIV Development Assistance and Adult Mortality in Africa. JAMA. 2012
- Carlo, Waldemar A et al. Newborn-Care Training and Perinatal Mortality in Developing Countries. NEJM. 2010.
- Guy, Gery. The effects of cost charing on access to care among childless adults.Health Services Research. 2010.

- King, Marissa et al. Medical school gift restriction policies and physician prescribing of newly marketed psychotropic medications: difference-in-differences analysis. BMJ. 2013.

- Li, Rui et al. Self-monitoring of blood glucose before and after medicare expansion among meicare beneficiaries with diabetes who do not use insulin.AJPH. 2008.

- Ryan, Andrew et al. The effect of phase 2 of the premier hospital quality incentive demonstration on incentive paymetns to hospitals caring for disadvantaged patients.Health Services Research. 2012.

Linear Probability Examples:

- Bradley, Cathy et al. Surgery Wait Times and Specialty Services for Insured and Uninsured Breast Cancer Patients: Does Hospital Safety Net Status Matter? HSR: Health Services Research. 2012.

- Monheit, Alan et al. How Have State Policies to Expand Dependent Coverage Affected the Health Insurance Status of Young Adults? HSR: Health Services Research. 2011.

Extensions (Differences-in-Differences-in-Differences):

- Afendulis, Christopher et al. The impact of medicare part D on hospitalization rates.Health Services Research. 2011.

- Domino, Marisa. Increasing time costs and co-payments for prescription drugs: an analysis of policy changes in a complex environment.Health Services Research. 2011.

Economics

- Card, David and Alan Krueger. Minimum Wage and Employment: A Case Study of the Fast Food Industry in New Jersey and Pennsylvania. The American Economic Review. 1994.

- DiTella, Rafael and Schargrodsky, Ernesto. Do Police Reduce Crime? Estimates Using the Allocation of Police Forces after a Terrorist Attack. American Economic Review. 2004.

- Galiani, Sebastian et al. Water for Life: The Impact of the Privatization of Water Services on Child Mortality. Journal of Political Economy. 2005.

Websites

Methodological

http://healthcare-economist.com/2006/02/11/difference-in-difference-estimation/

Statistical (sample R and Stata code)

http://thetarzan.wordpress.com/2011/06/20/differences-in-differences-estimation-in-r-and-stata/

Courses

Online

- National Bureau of Economic Research
- What's New in Econometrics? Summer Institue 2007.
- Lecture 10: Differences-in-Differences
- http://www.nber.org/minicourse3.html

Lecture notes and video recording, primarily focused on the theory and mathematical assumptions of difference in differences technique and its extensions.

Join the Conversation

Have a question about methods? Join us on Facebook

JOIN