Introduction to Difference in Differences (DID) Analysis

Hsueh-Sheng Wu
CFDR Workshop Series

June 15, 2020

BGSU
Center for Family and Demographic Research

Outline of Presentation

- What is Difference-in-Differences (DID) analysis
- Threats to internal and external validity
- Compare and contrast three different research designs
- Graphic presentation of the DID analysis
- Link between regression and DID
- Stata -diff- module
- Sample Stata codes
- Conclusions

What Is Difference-in-Differences Analysis

- Difference-in-Differences (DID) analysis is a statistic technique that analyzes data from a nonequivalence control group design and makes a casual inference about an independent variable (e.g., an event, treatment, or policy) on an outcome variable
- A non-equivalence control group design establishes the temporal order of the independent variable and the dependent variable, so it establishes which variable is the cause and which one is the effect
- A non-equivalence control group design does not randomly assign respondents to the treatment or control group, so treatment and control groups may not be equivalent in their characteristics and reactions to the treatment
- DID is commonly used to evaluate the outcome of policies or natural events (such as Covid-19)

Internal and External Validity

- When designing an experiment, researchers need to consider how extraneous variables may threaten the internal validity and external validity of an experiment
- Internal validity refers to the extent to which an experiment can establish the causal relation between the independent variable and the outcome variable
- External validity refers to the extent to which the causal relation obtained from an experiment can be generalized to other settings

BGSU
Center for Family and Demographic Research

Threats to Internal Validity

- **History**: historical events happened to respondents' lives during the course of the experiment
- **Maturation**: physiological and/or psychological changes among respondents during the course of the experiment
- **Testing**: respondents perform better on a similar test when they take it the second time
- **Instrumentation**: different measuring procedures or measurements are used in the pre-test and the post-test
- **Regression toward the mean**: the ceiling effect or the flooring effect
- **Selection**: the experiment and control groups are not equivalent groups in the first place, which contributes to the differences in the outcome variable later
- **Attrition:** the experiment and control groups differ in the likelihood of dropping out, leading to difference in the outcome variable later

Threats to External Validity

- **Reactive effects of experimental arrangements**: unique features of an experiment lead respondents to have change in the outcome variable
- **Reactive or interaction effect of testing**: unique features of the tests may lead respondents to be sensitive to a certain outcome variable, so the research findings may not be applicable to people that were not exposed to these tests.
- **Interaction effects of selection biases and the experimental variable**: participants of a new trial of an experimental drug
- **Multiple treatment interference**: fatigues from receiving multiple experiments

BGSU
Center for Family and Demographic Research

Compare and Contrast Three Different Research Designs

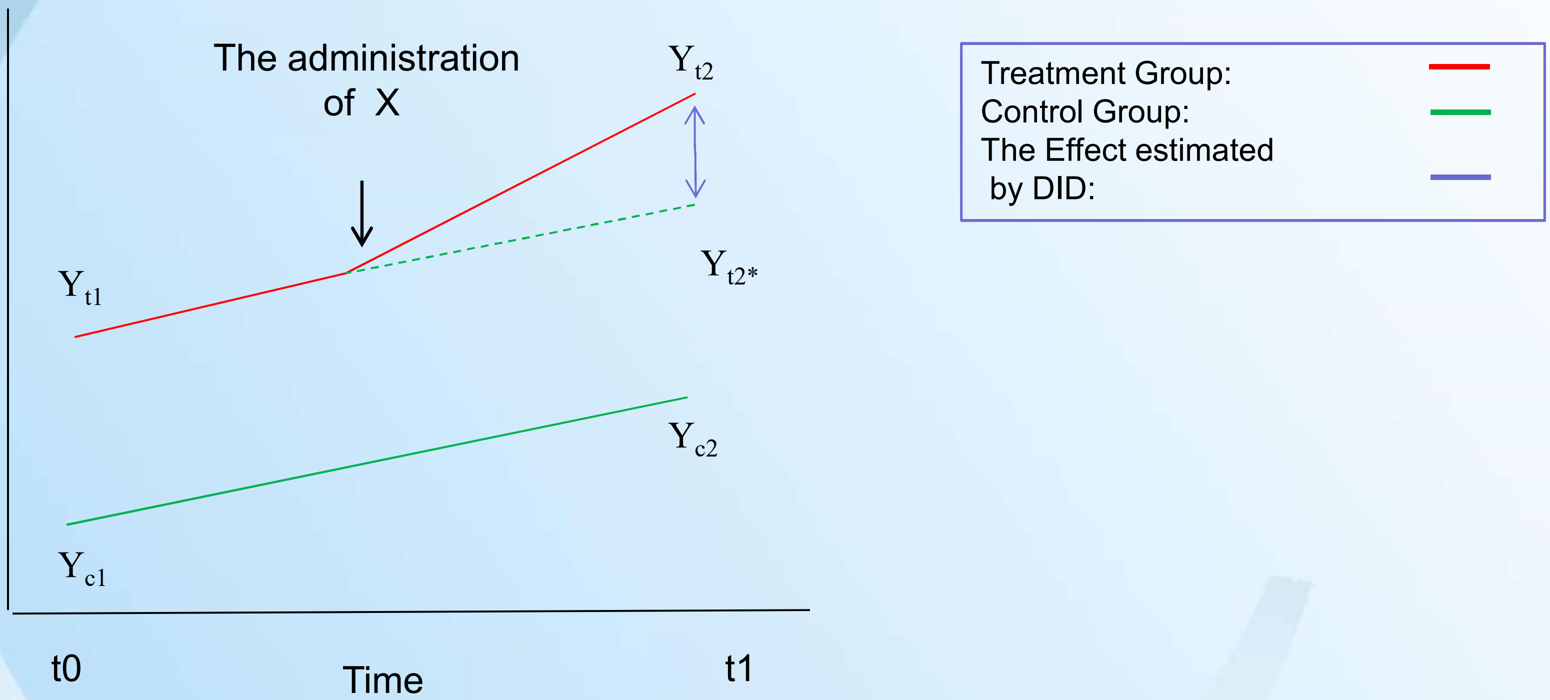
Table 1. Comparisons of an Experiment, a Quasi-Experiment, and a Survey

Sample Design*	The Pretest-Posttest Control Group Design	Nonequivalent Control Group Design*		A Two-Wave Panel Survey**
	Y_{t1} X Y_{t2} R --------------------------- Y_{c1} Y_{t2}	Y_{t1} X Y_{t2} ---------------------- Y_{c1} Y_{t2}		X_{t1} X_{t2} Y_{t1} Y_{t2} X_{c1} X_{c2} Y_{c1} Y_{c2}
Design Characteristics				
Randomization	✓	✗		✗
Manipulation of X	✓	✓		✗
Control for Internal Validity Threats				
History	✓	✓	✓	?
Instrumentation	✓	✓	✓	?
Testing	✓	✓	✓	?
Regression toward the mean	✓	✓	✓	?
Maturation	✓	✓	✓	?
Attrition	✓	✓	✓	?
Selection	✓	✓	✗	?
Interactdions between Selection and other threats	✓	✓	✗	?
Control for External Validity Threats				
Reactive effects of experimental arrangements	✗	✗	✗	✓
Reactive or interaction effect of testing	✗	✗	✗	✗
Interaction effects of selection biases and the experimental variable	?	?	✗	?
Multiple treatment interference	?	?	?	✓

*Difference-in Differences analysis ususally use data collected from this design

** Surveys generally relies on statistical methods, rather than research design, to control for threats to internal validity.

Graphic Presentation of the DID Analysis



Link between Regression and DID

- From the perspective of regression analysis, DID estimates the interaction term of time and treatment

- Y_{ij}= B_0 + B_1* Time + B_2*X + B_3*X*Time, where Y_{ij} is the value of Y for respondents in a treatment (or a control group) at a certain time point; Time is coded as 0 at t0 and 1 at t1; X is coded as 0 for the control group and 1 for the treatment group, so that:

$$Y_{c1} = B_0$$
$$Y_{c2} = B_0 + B_1$$
$$Y_{t1} = B_0 + B_2$$
$$Y_{t2*} = B_0 + B_1 + B_2$$
$$Y_{t2} = B_0 + B_1 + B_2 + B_3$$

- DID estimates the difference between Y_{t2} and Y_{t2*} = $(B_0 + B_1 + B_2 + B_3) - (B_0 + B_1 + B_2) = B3$

BGSU Center for Family and Demographic Research

Strengths and Weaknesses of DID

Strengths:

- DID is intuitive and can be easily understood within the framework of regression
- DID uses a nonequivalent control group design to establish the temporal order between the independent variable and the outcome variable, which is crucial in identify the causal direction of variables
- The incorporation of control group eliminates many threats, except the selection bias to internal validity, so researchers do not need to statistically control every confounding variables in the analysis

Weaknesses:

- In a natural experiment setting, it is difficult to understand what characteristics of experiments leads to change
- It is also unclear how much the experiment resembles the event in real life and raises the question about the external validity of the findings
- The equivalence between the treatment and control group (e.g., selection bias) prevents researchers from making valid casual inference on the treatment and the outcome variable. However, some statistical control (e.g., propensity score matching) can be used along with DID to reduce this problem.

Stata -diff- Module

- Dr. Juan Villa wrote the Stata -diff- module. Users can install this module by typing "ssc install diff" in the Stata command window.
- This module allows researchers to incorporate additional covariates of outcome to adjust for different levels of the outcome between the treatment and control groups
- This module allows researchers to reduce the selection bias problem by calculating the kernel propensity score and use it to match the treatment and control groups. In addition, this module can test whether these two groups are equivalent in covariates after matching is performed.
- This module analyzes quantile outcome variable
- This module conducts triple difference-in-differences analysis
- This module has a bootstrap option to gain a better estimate of the variance of the parameter
- This module can be used to analyze repeated cross-sectional data research design

BGSU Center for Family and Demographic Research

Examples of DID Analysis

See the Stata ado and log files

Conclusions

- Difference-in-Differences (DID) analysis is a useful statistic technique that analyzes data from a nonequivalence control group design and makes a casual inference about an independent variable (e.g., an event, treatment, or policy) on an outcome variable
- The analytic concept of DID is very easy to comprehended within the framework of regression
- Selection bias is the most common threat to the DID analysis. Researchers can reduce this problem by incorporating covariates that may contribute to the outcome variable or by using propensity score matching to make treatment and control groups equivalent.
- The findings of DID analysis may not be generalized to other settings, depending on what the experiment is, how much the experiment mimics the event in real life, and how respondents react to the experiment.
- Sociologists are interested in some constructs that should not or cannot be manipulated for ethical reasons (e.g., change in people's marital status, the occurrence of a pandemic disease or a natural disaster). Thus, if data happen to be collected before and after an event, researchers can use DID to analyze such data and gain a better understanding about the relation between the event and the outcome variable.

Center for Family and Demographic Research