ELSEVIER

Identification for difference in differences with cross-section and panel data

Myoung-jae Lee [a,b,*], Changhui Kang [c]

[a] Department of Economics, Korea University, Anam-dong, Sungbuk-ku, Seoul 136-701, Republic of Korea
[b] School of Economics and Social Sciences, Singapore Management University, 90 Stamford Road, Singapore 178903, Singapore
[c] Department of Economics, National University of Singapore, Singapore 117570, Singapore

Abstract

For difference-in-differences (DD), the identification question in using three types of progressively more informative data is addressed: independent cross-sections, 'mover' panels, and 'no-mover' panels. Although DD identifies 'the average effect on the treated', its meaning and the identification conditions differ across the data types.
© 2006 Elsevier B.V. All rights reserved.

1. Introduction

Difference in differences (DD) is one of the most popular–and often convincing–study designs in finding the effects of a treatment in social sciences. For instance, Bertrand et al. (2004) found as many as 92 papers in six economic journals over 1990–2000. Many DD references and examples can be found in Angrist and Krueger (1999), Besley and Case (2004), Rosenbaum (2002), and Lee (2005).

---

* Corresponding author. Department of Economics, Korea University, Anam-dong, Sungbuk-ku, Seoul 136-701, Republic of Korea. Tel.: +82 16 347 7867; fax: +82 2 926 3601.
  E-mail addresses: myoungjae@korea.ac.kr, mjlee@smu.edu.sg (M.-j. Lee), ecskch@nus.edu.sg (C. Kang).

To briefly introduce DD, consider a treatment given at some time point between time a and b. There are only two regions, $r=0$, 1, and the treatment is given only to region 1. For an individual i with responses y_{ia} and y_{ib}, $E(y_{ib} - y_{ia}|r=0)$ includes only the time effect whereas $E(y_{ib} - y_{ia}|r=1)$ includes both the time and treatment effects. Thus

$$DD = E(y_{ib} - y_{ia}|r = 1) - E(y_{ib} - y_{ia}|r = 0) \qquad (1.1)$$

$$= E(y_b|r = 1) - E(y_a|r = 1) - \{E(y_b|r = 0) - E(y_b|r = 0)\} \qquad (1.2)$$

identifies the desired treatment effect by removing the time effect, where the subscript i is omitted assuming iid across $i=1, \ldots, N$.

Although (1.1) requires panel data, (1.2) does not. In estimating DD, we can think of (at least) three types of data:

- Independent Cross-Sections from a large population, where each subject is observed only once.
- 'Mover' Panel data for two periods without dropouts and new entries, where some subjects are observed in the two different regions.
- 'No-Mover' Panel data without dropouts and new entries, where each subject is observed in the same region twice.

The usual panel data with dropouts and new entries may be regarded as a combination of the three types, because (i) some subjects are observed twice in the same region, (ii) some twice in the two different regions, and (iii) some only once. The main goal of this paper is to show what DD identifies under what conditions for the three data types. Once this task is done, the usual panel data may be analyzed by dividing the data into the three types and then analyzing each type separately. This statement, however, is subject to a caveat, because, for instance, (iii) occurs due to dropouts or new entries in which dropouts may not be representative of the population whereas new entries may be; that is, (iii) in the usual panel data may differ from a single wave in independent cross-sections. The main identification conditions and results are given in the next section.

2. Identification for DD in three data types

Let $j=0$ denote no treatment (or untreated state) and $j=1$ denote the treatment given (or treated state). Also let y_{jit}, $j=0$, 1, denote the 'potential' response at time t ($t=a$, b, $a<b$) when individual i received the treatment j 'exogenously' at some time point before t. Between the two potential responses (or outcomes) at any given time corresponding to the two potential treatments, only one outcome is observed while the other, called 'counter-factual', is not. There are variables other than the treatment that affect the response variable; call them covariates (x_{it}) if observed, and error terms if unobserved. While we want to know the effects of the treatment on the response variable, if the covariates or the error terms are 'unbalanced' across the treatment and control groups, then they can cause biases. Let

$$x_i \equiv \left(x'_{ia}, x'_{ib}\right)'.$$

2.1. Independent cross-sections

Define

δ_{it} = 1 if subject i is observed at time t, and 0 otherwise,
r_{it} = 1 if subject i resides in region 1 at time t, and 0 otherwise.

What is observed is, for $i=1, \ldots, N$,

$$\delta_{ia}, \delta_{ib}, \delta_{ia}r_{ia}, \delta_{ib}r_{ib}, \delta_{ia}x_{ia}, \delta_{ib}x_{ib}, \delta_{ia}y_{0ia}, \delta_{ib}\{(1-r_{ib})y_{0ib} + r_{ib}y_{1ib}\}.$$

Here, the assumption is that any subject residing in region 1 at time b has been treated, as if the treatment was given just before $t=b$. For two periods as here, $\delta_{ia}=1-\delta_{ib}$ should hold. Then the observed response is

$$y_i = (1-\delta_{ib})y_{ia} + \delta_{ib}y_{ib} = (1-\delta_{ib})y_{0ia} + \delta_{ib}\{(1-r_{ib})y_{0ib} + r_{ib}y_{1ib}\}.$$

To ease exposition, we start with DD without any x conditioned on, and then condition on x later.
DD with independent cross-sections is

$$\mathrm{DD} = E(y|r_b = 1, \delta_b = 1) - E(y|r_a = 1, \delta_a = 1)$$

$$- \{E(y|r_b = 0, \delta_b = 1) - E(y|r_a = 0, \delta_a = 1)\}$$

$$= E(y_{1b}|r_b = 1, \delta_b = 1) - E(y_{0a}|r_a = 1, \delta_a = 1)$$

$$- \{E(y_{0b}|r_b = 0, \delta_b = 1) - E(y_{0a}|r_a = 0, \delta_a = 1)\}.$$

Assumption δ. δ_a and δ_b are independent of y_{jt} and x given r_a or r_b for all j and t.

Under Assumption δ, the DD becomes

$$\mathrm{DD} = E(y_{1b}|r_b = 1) - E(y_{0a}|r_a = 1) - \{E(y_{0b}|r_b = 0) - E(y_{0a}|r_a = 0)\}.$$

Assumption y^0. The mean baseline response of those living in region j at time a is the same as the mean baseline response of those living in region j at time b:

$$E(y_{0a}|r_a = j) = E(y_{0a}|r_b = j), \quad j = 0, 1.$$

This is the "baseline uniformity across time for the same region". If the regions attract the "same type" of people over time, then this would hold. If, however, the treatment is substantial enough to change the composition of the residents before and after the treatment, then the condition may not hold for $j=1$. Under Assumptions δ and y^0, r_a in the preceding DD expression can be replaced by r_b to result in

$$\mathrm{DD} = E(y_{1b} - y_{0a}|r_b = 1) - E(y_{0b} - y_{0a}|r_b = 0).$$

Assumption Δy^0. The same time effect holds across the two regions:

$$E(y_{0b} - y_{0a}|r_b = 1) = E(y_{0b} - y_{0a}|r_b = 0).$$

Rewrite the last DD expression as

$$E(y_{1b} - y_{0a}|r_b = 1) - E(y_{0b} - y_{0a}|r_b = 1) + E(y_{0b} - y_{0a}|r_b = 1) - E(y_{0b} - y_{0a}|r_b = 0)$$

$$= E(y_{1b} - y_{0b}|r_b = 1) \text{ under Assumption } \Delta_y^0.$$

Hence, under Assumptions δ, y^0, and Δy^0, DD with independent cross-sections identifies the average treatment effect at time b for those in region 1 at time b.

So far we ignored covariates, but for Assumptions y^0 and Δy^0 to be plausible, covariates should be controlled for.

Assumption y^1. The mean baseline response of those living in region j at time a with its covariate $x_a = x_o$ is the same as the mean baseline response of those living in region j at time b with its covariate $x_b = x_o$:

$$E(y_{0a}|x_a = x_o, r_a = j) = E(y_{0a}|x_b = x_o, r_b = j), \quad j = 0, 1.$$

Assumption Δy^1. The same time–effect holds across the two regions conditioning only on x_b:

$$E(y_{0b} - y_{0a}|x_b, r_b = 1) = E(y_{0b} - y_{0a}|x_b, r_b = 0).$$

Conditioning on x_b instead of x_a or x is somewhat troubling. But, since r_b appears in the conditioning set, only x_b can appear in the conditioning set for independent cross-sections. If $x_a = x_b$ as is often the case in practice, then conditioning on x_b is no longer a problem.

Under Assumption δ, we get

$$DD_{xt} \equiv E(y|x_b = x_o, r_b = 1, \delta_b = 1) - E(y|x_a = x_o, r_a = 1, \delta_a = 1)$$

$$- \{E(y|x_b = x_o, r_b = 0, \delta_b = 1) - E(y|x_a = x_o, r_a = 0, \delta_a = 1)\}$$

$$= E(y_{1b}|x_b = x_o, r_b = 1) - E(y_{0a}|x_a = x_o, r_a = 1)$$

$$- \{E(y_{0b}|x_b = x_o, r_b = 0) - E(y_{0a}|x_a = x_o, r_a = 0)\}.$$

Using Assumption y^1, we can replace "$x_a = x_o$, r_a" with "$x_b = x_o$, r_b" to get

$$DD_{xt} = E(y_{1b} - y_{0a}|x_b, r_b = 1) - E(y_{0b} - y_{0a}|x_b, r_b = 0)$$

$$= E(y_{1b} - y_{0a}|x_b, r_b = 1) - E(y_{0b} - y_{0a}|x_b, r_b = 1)$$

$$+ E(y_{0b} - y_{0a}|x_b, r_b = 1) - E(y_{0b} - y_{0a}|x_b, r_b = 0).$$

Assumption Δy^1 now yields

$$DD_{xt} = E(y_{1b} - y_{0b}|x_b, r_b = 1).$$

Finally, integrating out x_b with the distribution of $x_b|(r_b=1)$, say $F(x_b|r_b=1)$, we get a 'marginal effect'—the effect on the treated:

$$\int DD_{x_t} F(dx_b|r_b = 1) = E(y_{1b} - y_{0b}|r_b = 1).$$

Therefore, under Assumptions δ, y^1 and Δy^1, DD with independent cross-sections identifies the average effect at time b for those in region 1 at time b. Since the treated response is observed only for region 1 at time b, unless some assumptions are imposed, it is natural that only the effect for region 1 at time b is identified. Note that the 'baseline difference' $E(y_{0a}|x_b, r_b=1) \neq E(y_{0a}|x_b, r_b=0)$ is allowed in Assumption Δy^1, although Assumption y^1 imposes the baseline uniformity (given x_a or x_b) across time for the same region.

2.2. Mover panels

The main disadvantage of Assumptions y^1 and Δy^1 is that the time index of x and r should match, because a subject is observed only once. Although Assumption y^1 may not be restrictive as only y_{0a} appears, Assumption Δy^1 is so as both y_{0a} and y_{0b} appear. For this, panel data with movers can help. What is observed is

$$r_{ia}, r_{ib}, x_{ia}, x_{ib}, y_{0ia}, (1 - r_{ib})y_{0ib} + r_{ib}y_{1ib}.$$

The DD conditioned on x is

$$\begin{aligned} DD_x &\equiv E(y_b|x, r_b = 1) - E(y_a|x, r_a = 1) - \{E(y_b|x, r_b = 0) - E(y_a|x, r_a = 0)\} \\ &= E(y_{1b}|x, r_b = 1) - E(y_{0a}|x, r_a = 1) - \{E(y_{0b}|x, r_b = 0) - E(y_{0a}|x, r_a = 0)\}. \end{aligned}$$

Assumption y^2. The mean baseline response of those living in region j at time a with its covariate x is the same as the mean baseline response of those living in region j at time b with its covariate x:

$$E(y_{0a}|x, r_a = j) = E(y_{0a}|x, r_b = j), \quad j = 0, 1.$$

Assumption Δy^2. The same time–effect holds across the two regions conditioning on x:

$$E(y_{0b} - y_{0a}|x, r_b = 1) = E(y_{0b} - y_{0a}|x, r_b = 0).$$

Having both x_a and x_b in the conditioning set, mover panels relax the restrictive nature of Assumptions y^1 and Δy^1 for independent cross-sections.

Under Assumption y^2, r_a in the above DD_x can be replaced by r_b to yield

$$DD_x = E(y_{1b} - y_{0a}|x, r_b = 1) - E(y_{0b} - y_{0a}|x, r_b = 0).$$

Subtract and add $E(y_{0b} - y_{0b}|x, r_b = 1)$ to this, and then use Assumption Δy^2 to get

$$\mathrm{DD}_x = E(y_{1b} - y_{0b}|x, r_b = 1).$$

Integrating x out with the distribution of $(x_a, x_b)|(r_b = 1)$ gives the effect on the treated $E(y_{1b} - y_{0b}|r_b = 1)$. Therefore, under Assumptions y^2 and Δy^2, DD with mover panels identifies the average effect at time b for those in region 1 at time b. Note that the 'baseline difference' $E(y_{0a}|x, r_b = 1) \neq E(y_{0a}|x, r_b = 0)$ is allowed in Assumption Δy^2, although Assumption y^2 imposes the baseline uniformity (given x) across time for the same region.

2.3. No-mover panels

Consider now panel data with no movers across the two regions. Define

r_i	=1 if in region 1, and 0 otherwise,
τ_t	=1 if $t=b$, and 0 otherwise,
	$\Rightarrow d_{it} = r_i \tau_t$: being in region 1 at time b means having received the treatment.

Although it is fine to imagine two real regions, we may as well think of two groups, treated and untreated, in which case d_{it} is not $r_i \tau_t$, but simply a treatment indicator for person i at time t. The two groups in this case should be unchangeable, e.g., groups based on gender or race. The observed response for individual i in no-mover panel is

y_{it}	$= (1 - d_{it})y_{0it} + d_{it}y_{1it}$
	$\Rightarrow y_a = y_{0a}, y_b = y_{0b}$ for $r=0$ and $y_a = y_{0a}, y_b = y_{1b}$ for $r=1$.

Assumption Δy^3. The same time–effect holds across the two regions conditioning on x:

$$E(y_{0b} - y_{0a}|x, r = 1) = E(y_{0b} - y_{0a}|x, r = 0).$$

By construction, $r_a = r_b = r$, and thus Assumption y^2 is not necessary for no-mover panels.
  Observe

$$\mathrm{DD}_x = E(y_b - y_a|x, r = 1) - E(y_b - y_a|x, r = 0)$$

$$= E(y_{1b} - y_{0a}|x, r = 1) - E(y_{0b} - y_{0a}|x, r = 0)$$

$$= E(y_{1b} - y_{0a}|x, r = 1) - E(y_{0b} - y_{0a}|x, r = 1)$$

$$+ E(y_{0b} - y_{0a}|x, r = 1) - E(y_{0b} - y_{0a}|x, r = 0).$$

Under Assumption Δy^3,

$$\mathrm{DD}_x = E(y_{1b} - y_{0a}|x, r = 1) - E(y_{0b} - y_{0a}|x, r = 1) = E(y_{1b} - y_{0b}|x, r = 1).$$

Integrating out x with the distribution of $x|(r=1)$ yields the effect on the treated $(r=1)$. Therefore, under Assumption Δy^3, DD with no-mover panels identifies the average effect at time b for those in region 1. Note that the 'baseline difference' $E(y_{0a}|x, r=1) \neq E(y_{0a}|x, r=0)$ is allowed in Assumption Δy^3, and there is no longer any required baseline uniformity (given x) across time for the same region.

References

Angrist, J., Krueger, A., 1999. Empirical strategies in labor economics. In: Ashenfelter, O., Card, D. (Eds.), The Handbook of Labor Economics, vol. 3A. Elsevier, New York.

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates. Quarterly Journal of Economics 119, 249–275.

Besley, T., Case, A., 2004. Unnatural experiments? Estimating the incidence of endogenous policies. Economic Journal 110, F672–F694.

Lee, M.J., 2005. Micro-Econometrics for Policy, Program, and Treatment Effects. University Press, Oxford.

Rosenbaum, P.R., 2002. Observational Studies, 2nd ed. Springer-Verlag.