Alternatives to polynomial trend-corrected differences-in-differences models

Vincent Vandenberghe

Routledge
Taylor & Francis Group

Check for updates

ARTICLE

Alternatives to polynomial trend-corrected differences-in-differences models

Vincent Vandenberghe

IRES-IMMAQ-UCL, Economics School of Louvain (ESL), Louvain-la-Neuve, Belgium

ABSTRACT

A common problem with differences-in-differences (DD) estimates is the failure of the parallel-trend assumption. To cope with this, most authors include polynomial (linear, quadratic...) trends among the regressors, and estimate the treatment effect as a once-in-a-time trend shift. In practice, that strategy does not work very well, because *inter alia* the estimation of the trend uses post-treatment data. An extreme case is when sample covers only one period before treatment and many after. Then the trend's estimate relies almost completely on post-treatment developments, and absorbs most of the treatment effect. What is needed is a method that *i)* uses pretreatment observations to capture linear or nonlinear trend differences, and *ii)* extrapolates these to compute the treatment effect. This article shows how this can be achieved using a fully flexible version of the canonical DD equation. It also contains an illustration using data on a 1994–2000 EU programme that was implemented in the Belgian province of Hainaut.

I. Introduction

When the parallel-trend assumption fails, most authors (e.g. Friedberg 1998; Autor 2003; Besley and Burgess 2004) resort to a polynomial (linear,...) trend-augmented version of the canonical differences-in- differences (DD) model (Angrist and Pischke 2009).

$$Y_{it} = \alpha + \sum_{\tau=t_2}^{T} \alpha^{\tau} I_{\tau,t} + \alpha^D D_i + \eta AFTER_t *$$
$$D_i \left[+\theta \, time * D_i \right]$$

(1)

with *t=1,...,T* and $I_{\tau,t} = 1$ if $t = \tau$ and 0 otherwise covering before and after treatment periods, and where Y_{it} is entity *i*'s outcome in time *t*, *D* the treatment dummy, *AFTER* the after-treatment dummy, and here *time* is a continuous variable.

Coefficient θ captures the linear trend characterizing the treated entities. And η – a trend shift around *time = 0* – measures the treatment effect. As suggested by Wolfers (2006), the problem with this strategy is that it uses post-treatment observations, and that the treatment outcome takes the form of a once-in-a-time trend shift. A case in

point is visible in Figure 1. The latter describes the evolution of income per head in the Belgian province of Hainaut (in deviation to the rest of Belgium), before and after it benefited from EU money.[1] That treatment began in 1994 and lasted until 2000. The trend is clearly negative prior to treatment, and still so after. The estimation of η, using the canonical DD model 10 years after treatment, delivers a negative value, in the range of −300 €. A 'placebo' estimation of that model evidently reveals that there was no parallelism before the treatment started. So, the −300€ figure is not trustworthy. This justifies estimating the trend-augmented Equation (1). The short dashed grey line in Figure 1 depicts the result. After treatment, the income handicap tends to stabilize, and this explains the moderately negative estimated trend ($\theta < 0$). By construction, this trend applies to the pretreatment period. Being negative, it delivers 'corrected' DD estimates that are less negative than the traditional ones (−245.8€>-297.3€). Also, η corresponds to the trend shift just after *time = 0*.[2] And as income handicap after treatment is larger, that shift is still negative, suggesting that the EU

CONTACT Vincent Vandenberghe ✉ Vincent.vandenberghe@uclouvain.be ⬤ IRES-IMMAQ-UCL, Economics School of Louvain (ESL), 3 place Montesquieu, B-1348, Louvain-la-Neuve, Belgium
[1]See Vandenberghe (2018) for more details about EU-Objective 1-Hainaut.
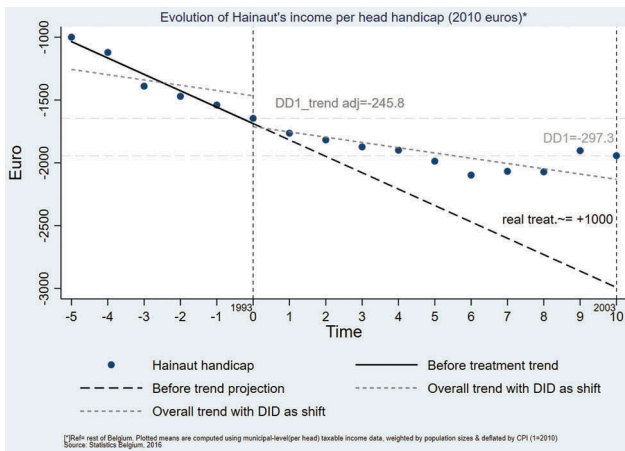[2]Defining *time = year-1993*.

Figure 1. The limitations of trend-augmented DD[$].

[$] Plotted values are (municipal)-population-weighted mean difference between Hainaut and rest of Belgium. These are used to estimate a linear trend-adjusted DD model.

policy failed (it 'caused' approx.: – 245€ of additional income handicap). Yet, θ underestimates the actual pretreatment trend (solid black line in Figure 1). Before treatment, the handicap was growing faster than after. Prolonging the initial trend up to $t = 10$ suggests that, *ceteris paribus*, the income handicap might have reached −3,000€, while it ended being less than −2,000€. The tentative conclusion is that the real treatment outcome was positive (in the range of +1,000€). What we propose hereafter is an alternative way of correcting DD estimation that solely uses pretreatment observations.

II. Beyond polynomial trend-corrected DD

Mora and Reggio (2012) suggest that DD analysis can be done by estimating a generalized fully flexible equation, where the right-hand part only consists of time, treatment and timeXtreatment dummies:

$$Y_{it} = \gamma + \sum_{\tau=t_2}^{T} \gamma_\tau I_{\tau,t} + \gamma^D D_i + \sum_{\tau=t_2}^{T} \gamma^D_\tau I_{\tau,t} D_i \quad (2)$$

with $t = t_1, \ldots. T$ and $I_{\tau,t} = 1$ if $t = \tau$ and 0 otherwise, covering before and after treatment periods.

The advantages to this equation are manifold. First, conditional on the availability of many pretreatment periods in the data, the OLS-estimated coefficients can be used to compute a whole family of DD estimators $DD_{[p]}$, where $p = 1, 2\ldots q$ is the degree of parallelism underpinning identification. The canonical DD model is noted $DD_{[1]}$, and rests on parallelism of degree 1 (*Parallel*$_{[1]}$ hereafter).[3] Without *Parallel*$_{[1]}$, one should estimate $DD_{[2]}$ that rests on *Parallel*$_{[2]}$ (i.e. outcome growth rate parallelism).[4] If *Parallel*$_{[2]}$ fails, one should turn to $DD_{[3]}$ which requires *Parallel*$_{[3]}$ or outcome acceleration[5] parallelism…and so on up to degree $p = q$, if data permit. Second, Equation (2), unlike Equation (1) can capture dynamic (i.e. lagged) responses to treatment.[6] Third, – and this is something we particularly stress in the context of this article as it brings a solution to Wolfer's trend and shift problem – corrections for the violation of *Parallel*$_{[p]}$ rests solely on pretreatment observations.

Consider the canonical $DD_{[1]}$/*Parallel*$_{[1]}$ estimator, with just before-and-after observations t^* and t^*+1. Treatment effect writes,[7,8]

$$DD_{[p=1]}^{t^*+1;t^*} = (\gamma^D_{t^*+1} + \gamma^D) - (\gamma^D_{t^*} + \gamma^D)$$
$$= \gamma^D_{t^*+1} - \gamma^D_{t^*}$$

$$(3)$$

Also, Equation (2) can be used to assess *Parallel*$_{[1]}$ prior to treatment. Using pretreatment periods t^*-2, t^*-1, one can compute 'placebo' $DD_{[1]}$ capturing the deviation from *Parallel*$_{[1]}$ prior to treatment. For instance, $DD_{[1]}^{t^*;t^*-1} = \gamma^D_{t^*} - \gamma^D_{t^*-1}$. should not be statistically different from zero. If not, then treated and control trends diverge before treatment (as illustrated in Figure 1 or its stylized equivalent Figure 2). And identification should rest on *Parallel*$_{[2]}$. The point is that this can be easily done by computing

---

[3] If outcome level change by unit of time (i.e. 1st derivate) is 'speed', then *Parallel*$_{[1]}$ means stable level differences due to identical speeds.

[4] If outcome growth rate change by unit of time (2nd derivative) is 'acceleration', then *Parallel*$_{[2]}$ means stable growth rate differences due to same accelerations.

[5] If outcome acceleration change by unit of time (3rd derivative) is 'surge', then *Parallel*$_{[3]}$ corresponds to a situation where acceleration differences remain stable due to identical surges.

[6] The pattern of lagged effects is usually of substantive interest (e.g. if treatment effect should grow or fade as time passes).

[7] When estimating Equation (2) with only two periods, $\gamma^D_{t^*}$ is subsumed into the constant γ^D and $DD_{[1]}$ is directly captured by the time X treatment coefficient.

[8] Treatment effect's SE must account for the fact that it consists of a linear combination of estimated coefficients, and thus of the covariance between variables. That is automatically done by STATA test or lincom commands used hereafter, that exploit the variance-covariance matrix of the estimated coefficients.
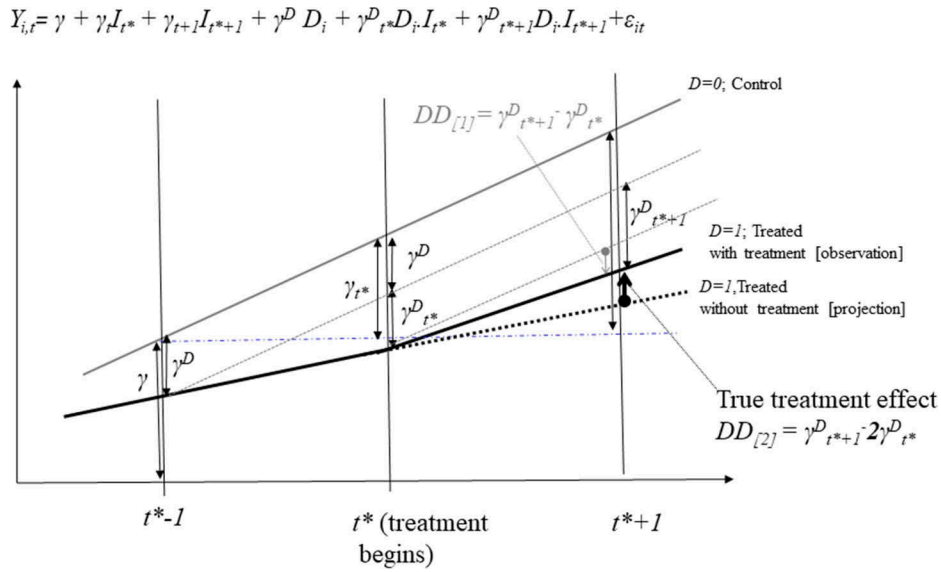
$$Y_{i,t} = \gamma + \gamma_t I_{t*} + \gamma_{t+1} I_{t*+1} + \gamma^D D_i + \gamma^D_{t*} D_i . I_{t*} + \gamma^D_{t*+1} D_i . I_{t*+1} + \varepsilon_{it}$$

Figure 2. [$] How $DD_{[2]}$ copes with failure of $Parallel_{[1]}$.

[$] On this figure, $t*-1$ is the first period observed in the data. Hence, γ^D_{t*-1} is subsumed into γ^D and, in contrast with Equation (4), $DD_{[2]}$ is computed using only two coefficients.

$$DD_{[p=2]}{}^{t*+1; \, t*-1} = DD_{[1]}{}^{t*+1; \, t*} - DD_{[1]}{}^{t*;t*-1}$$
$$= \left(\gamma^D_{t*+1} - \gamma^D_{t*}\right)$$
$$- \left(\gamma^D_{t*} - \gamma^D_{t*-1}\right)$$
$$= \gamma^D_{t*+1} - 2\gamma^D_{t*} + \gamma^D_{t*-1} \qquad (4)$$

which is the difference between the observed $t*+1$ outcome level handicap[9] γ^D_{t*+1} and its prediction $\gamma^D_{t*} + DD_{[1]}{}^{t*;t*-1}$ given the handicap in $t*$ and its expected rise due to growth-rate difference between $t*$ and $t*-1$. This prediction uses only regression coefficients driven by pretreatment observations; a major difference with the trend-augmented method of Equation (1). Note finally that the aforementioned logic can be generalized in many ways: to the case of lagged/dynamic treatment effects, or to $DD_{[p=q]}/Parallel_{[p=q]}$. where $q > 2$ (Vandenberghe 2018).

III. Application to Hainaut data

To illustrate the properties of the Equation (2)-based generalized fully flexible DD estimator, we use municipal data on taxable income per head. Time series are available for each of Belgium's 589 municipalities, from 1988 to 2003, covering years before 1994 (start of EU policy) and after 2000 (end of the policy). The treated entities are the 69

Table 1. Municipality count.

Rest of Belgium	520
Hainaut	**69**
Total	589

municipalities of Hainaut (Table 1 and Figure 3). The 520 other ones form the control group. All reported estimates are obtained using data that are weighted by municipal population sizes and deflated by 2010 consumer-price index.
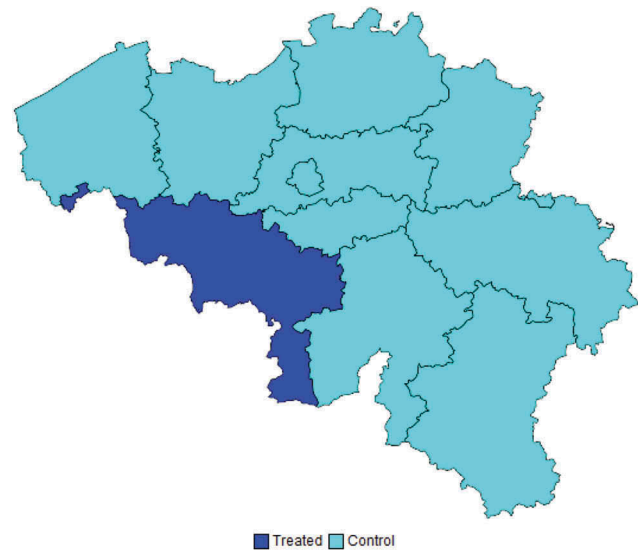

Figure 3. Hainaut vs rest of Belgium.

---

[9]Net of the initial handicap in $t*-1$: γ^D

Table 2. $DD_{[1]}$ estimation + $DD_{[1]}$ placebo estimations.

	$DD_{[1]}$	Placebo $DD_{[1]}$				
DD	−329.88*	−121.29***	−268.83***	−80.24***	−68.37*	−106.81***
prob DD = 0	0.017	0.000	0.000	0.000	0.035	0.000
Post-treat. year	2003	1989	1990	1991	1992	1993
Pretreat. year	1993	1988	1989	1990	1991	1992
Nobs	21,832	21,832	21,832	21,832	21,832	21,832
R^2	0.92	0.92	0.92	0.92	0.92	0.92

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4. $DD_{[2]}$ 'placebo' estimation.

	$DD_{[2]}$ 1990	$DD_{[2]}$ 1991	$DD_{[2]}$ 1992	$DD_{[2]}$ 1993
DD	−147.54	188.59	11.87	−38.43
prob DD = 0	0.593	0.448	0.959	0.843
Post-treat. year	1990	1991	1992	1993
Pretreat. year 1	1989	1990	1991	1992
Pretreat. year 2	1988	1989	1990	1991
Nobs	21,832	21,832	21,832	21,832
R^2	0.92	0.92	0.92	0.92

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2 displays the results for the canonical $DD_{[1]}/Parallel_{[1]}$. Year $t^* = 1993$ is the most immediate year before the treatment, and $t^*+s = 2003$ the moment the treatment is evaluated. Results confirm what was already visible in Figure 1. Compare to the rest of Belgium, the income handicap grew larger between 1993 and 2003 (−329.8€). But placebo $DD_{[1]}$ point at a rising income handicap prior to treatment. Thus $Parallel_{[1]}$ does no hold.

Thus, it is necessary to go beyond $Parallel_{[1]}$ to capture EU-Objective 1's true impact. Interestingly, as we possess many pretreatment periods, we can implement both the traditional trend-corrected DD method and $DD_{[2]}/Parallel_{[2]}$. Results are reported in Table 3. The last two columns correspond to year 2003 (t^*+10). As anticipated, the two estimators deliver treatment effect estimates that significantly diverge. Whereas the traditional linear trend-corrected method concludes to a negative impact (i.e. the income handicap rose by −238.5€), our preferred fully flexible $DD_{[2]}/Parallel_{[2]}$ method displays a gain of 916.2€. This illustrates the striking differences induced by a method that only uses pretreatment observations to account for trend differences, and also lifts the constraint of outcome as one-in-a-time trend shift.

Finally, we assess the legitimacy of $Parallel_{[2]}$ by estimating placebo $DD_{[2]}$ using 3 pretreatment years. Results (Table 4) are all supportive of $DD_{[2]} = 0$, suggesting that $Parallel_{[2]}$ was a realistic description of the relative dynamics of Hainaut's income per head prior to EU-Objective 1, and that the positive $DD_{[2]}$ values in Table 3 properly identify the programme's causal impact.

Table 3. Linear trend-corrected DD vs $DD_{[2]}$.

	2000		2003	
	Corrected $DD_{[1]}$	$DD_{[2]}$	Corrected $DD_{[1]}$	$DD_{[2]}$
DD	76.55	255.56***	−238.52***	961.19***
Prob DD = 0	0.154	0.000	0.001	0.000
Post-treat. year	1997	1997	2003	2003
Pretreat. year 1	1993	1993	1993	1993
Pretreat. year 2	1988	1988	1988	1988
Nobs	5,890	21,832	9,421	21,832
R^2	0.96	0.92	0.95	0.92

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Disclosure statement

No potential conflict of interest was reported by the author.

References

Angrist, J. D., and S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. London, UK: Princeton University Press.

Autor, D. 2003. "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing." *Journal of Labor Economics* 21 (1): 1–42. doi:10.1086/344122.

Besley, T., and R. Burgess. 2004. "Can Labor Regulation Hinder Economic Performance? Evidence from India." *The Quarterly Journal of Economics* 119 (1): 91–134. Oxford University Press. doi:10.1162/003355304772839533.

Friedberg, L. 1998. "Did Unilateral Divorce Raise Divorce Rates? Evidence from Panel Data." *American Economic Review* 88 (3): 608–627.

Mora, R., and I. Reggio. 2012. "Treatment Effect Identification Using Alternative Parallel Assumptions". *UC3M Working papers- Economics*, Universidad Carlos III de Madrid. Departamento de Economía.

Vandenberghe, V. 2018. "Treatment-Effect Identification Without Parallel Paths, *Economics: The Open-Access, Open-Assessment E-Journal*, Kiel Institute for the World Economy (IfW), 12 (2018-9): pp. 1–19.

Wolfers, J. 2006. "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results." *American Economic Review* 96 (5): 1802–1820. doi:10.1257/aer.96.5.1802.