Generalization of the Difference-in-Differences Model for Applied Research

Guido Cataife, Brandon Hesgrove, Rekha Varghese

IMPAQ International

IMPAQ INTERNATIONAL

Evaluating & Enhancing Public Programs & Policy
www.impaqint.com | @impaqint | (443) 259-5500

MOTIVATION

Difference-in-differences (DD) models are a very popular method for policy evaluation

However, the identifying assumption (parallel trends assumption) is rarely satisfied

- Can address using matching/weighting
- Alternatively, can control for this observed difference between groups

We discuss the properties of the generalized DD model (GDD), which adds linear time trends and relaxes the stringent parallel trends assumption, and provide insights relevant for the applied researcher.

LITERATURE

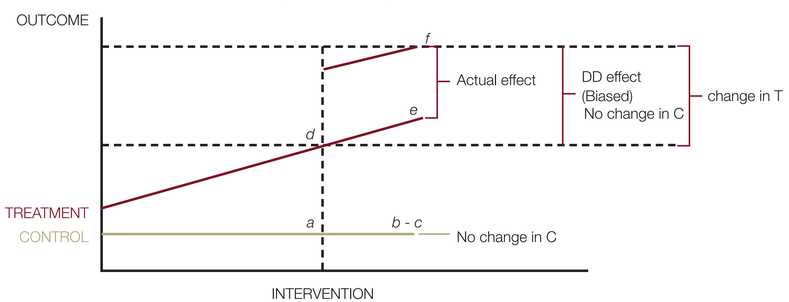
Similar models presented decades ago in policy evaluation literature

- Simonton (1977) presented a model based on Donald Campbell's comparative interrupted time series (CITS) work
- CITS remains popular in policy evaluation literature, but has been mostly ignored by economists

Lee (2016) introduced a similar model and called it GDD. Mora and Reggio (2017) discuss a broad class of models that include GDD as a particular case. The model specification we use is straightforward to estimate in health economics applications. We describe how it works, how to interpret its results, what the main properties are, and show how it performs under different data generating processes.

METHODS

- DD model controls for external trends using the change observed over time in the control group as the counterfactual for what would have occurred to the treatment group in the absence of the treatment
- DD produces correct results *only if* the treatment and control groups would have had the same trends in the post-period in the absence of the treatment. Usually referred to as *Parallel Trends Assumption*
- Parallel Trends Assumption rarely holds



We can account for the fact that baseline trends are different

Instead of assuming parallel trends, assume that any observed difference in trends between T and C would have persisted throughout post period

- We call this the "Generalized Difference-in-Differences Model". The estimator is identical to DD when the trends are parallel in baseline
- Estimator can identify effect under certain violations of DD assumption

$$Y = \boldsymbol{\beta_0} + \beta_1 treat * post * time + \boldsymbol{\beta_2 treat * post} + \beta_3 treat * time + \beta_4 post * time + \boldsymbol{\beta_5 treat} + \beta_6 time + \boldsymbol{\beta_7 post} + \boldsymbol{\varepsilon}$$

The DD model only contained the bolded terms. The newly added terms allow identification of the four trends: Trends of T and C in the baseline and post periods.

GDD average treatment effect on the treated (ATT) estimate is

$$\beta_2 + \left(\frac{1 + 2 + \cdots + N}{N}\right)\beta_1$$

Where N is the number of post-periods (normalize last period of baseline = 0)
Equivalently, average all post-period treatment effect estimates

We simulate data under multiple scenarios to compare the estimates produced by DD and GDD models.

Our data generating processes (DGP)

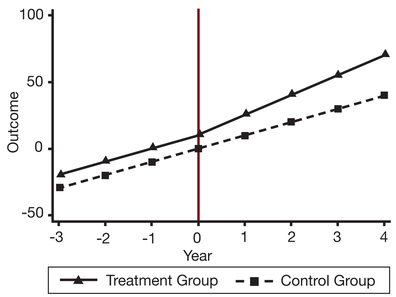
- Group-period level observations
- 4 pre-periods and 4 post-periods
- No noise (*e* always 0)

Scenarios

- Either baseline parallel trends or non-parallel trends that differ up to a linear term
- We use different slope and level shocks, for both periods and groups

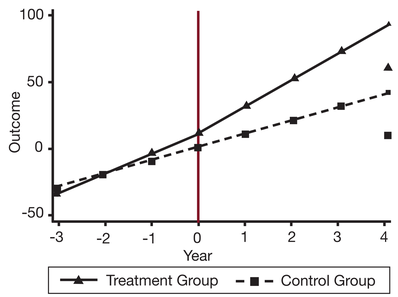
Use regression analysis to estimate DD and GDD

RESULTS



DGP 1

- Y = 10*year + 5*treat*post*year+10*treat
- DD Estimate = 12.5
- GDD Estimate = 12.5



DGP 2

- Y= 10*year + 5*year*treat + 5*treat*post*year + 10*treat
- DD Estimate = 32.5
- GDD Estimate = 12.5

Variables	(1) DGP3_DD	(2) DGP3_GDD	(3) DGP4_DD	(4) DGP4_GDD
Year*Treat*Post		0.09 (0.09)		0.09 (0.09)
Post*Year		4.94*** (0.06)		-0.06 (0.06)
Treat*Post	-0.04 (0.89)	-0.09 (0.21)	4.96*** (0.10)	4.91*** (0.21)
Year*Treat		-0.04 (0.06)		-0.04 (0.06)
Treat	10.06*** (0.63)	10.00*** (0.12)	5.06*** (0.07)	5.00*** (0.12)
Post	32.51*** (0.63)	0.00 (0.15)	5.01*** (0.07)	5.00*** (0.15)
Year		5.04*** (0.04)		0.04 (0.04)
Constant	-7.51*** (0.45)	0.05 (0.08)	-0.01 (0.05)	0.05 (0.08)
Observations	1,600	1,600	1,600	1,600
R-squared	0.78	1.00	0.97	.97
P_Value_Est	0.97	0.66	0.00	0.00
Estimate	-.004	0.12	4.96***	5.12***
Std Error	(0.89)	(0.27)	(0.10)	(0.27)

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

DGP 3: 10*treat + 5*year + 5*year*post
DGP 4: 5*treat + 5*post + 5*treat*post

- Standard errors larger with DD than GDD when DGP has linear time trends
- Standard errors smaller with DD than GDD when DGP has only level shocks

CONSIDERATIONS

GDD relies on the assumption that baseline trend of C group would persist

If GDD assumption is violated

- GDD will not identify the effect
- DD will not identify the effect either
  - Unless by chance the counterfactual trends turn parallel in the post period (and the DD parallel trend assumption is satisfied)
- In such a case GDD might not produce lower bias than DD
  - Should be cautious with a short baseline and volatile trends

CONCLUSIONS AND RECOMMENDATIONS

GDD is a useful tool for policy evaluation

- Relaxes the identifying assumption of DD at little cost
- It has low data requirements
- Provides the same treatment effect estimate as DD with baseline parallel trends
- Identifies estimates when baseline trends differ linearly

Applied researchers should consider GDD in situations where DD would be used

Lee, MJ. (2016). "Generalized Difference In Differences With Panel Data and Least Squares Estimator." Sociological Methods & Research 45(1): 134-157.
Simonton, D. (1977). "Cross-Sectional Time-Series Experiments: Some Suggested Statistical Analyses." Psychological Bulletin 84(3): 489-502.
Mora, R., and Reggio, I. (2017). "Alternative Diff-In-Diffs Estimators With Several Pretreatment Periods." Econometric Reviews 1-22.