Difference in Difference Models

Mae McDonnell

AOM 2019

Difference in difference models

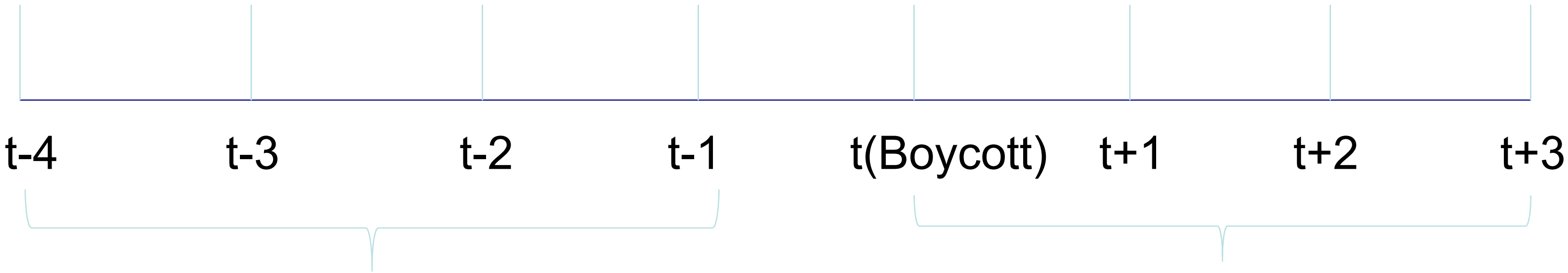
- Maybe the most popular identification strategy in applied work today
- Attempts to mimic random assignment with treatment and “comparison” sample

Basic problem

- You want to estimate the effect of some treatment (ie, "intervention")
- Intervention occurs at time period t_1
- Have data from pre- and post- intervention
  - Note, can be two-period panel or longitudinal panel with fixed effects
- But how do we know that observed difference is due to the intervention and not secular (ie, population-level) trends?

H{dp sdh#urp #P fG rqqhdd# #F ree#Irukfrp lqj,

How do boycotts affect board turnover?



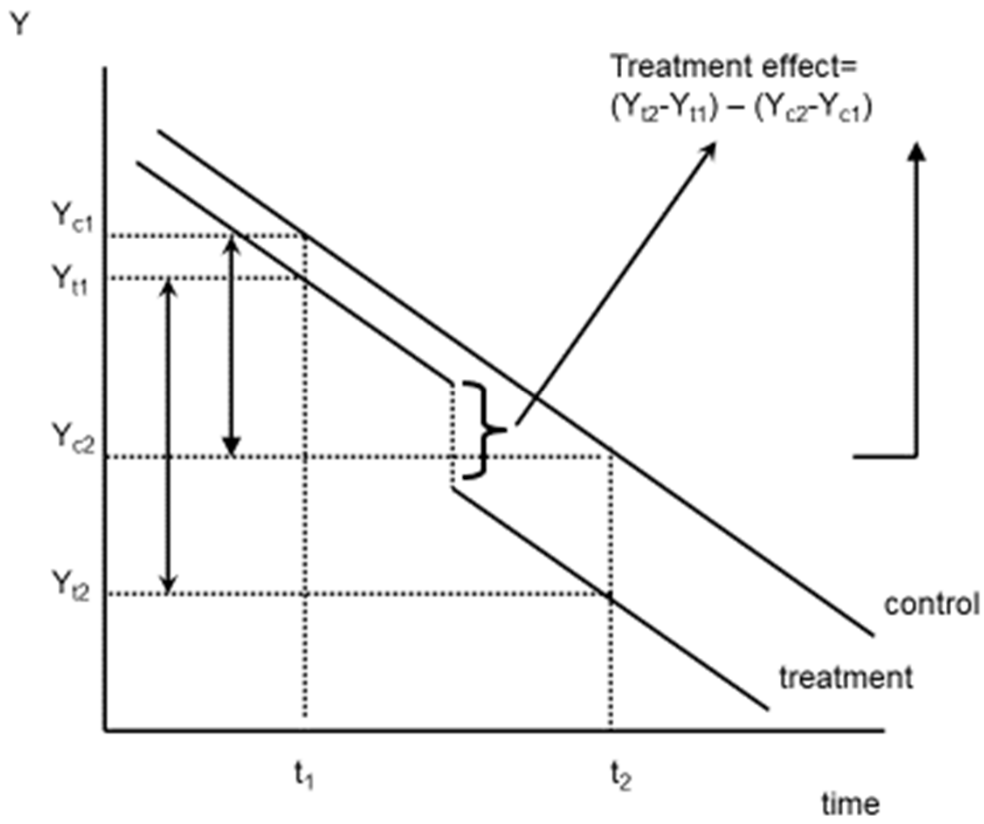
Pre-treatment Turnover (Baseline): 22.9%

Post-treatment Turnover: 29.7%

Boycotts are associated with a 30% increase over the baseline turnover rate (p < .05)

Difference in difference models

- Use observed trend in an untreated control group to establish what would have occurred in the absence of the intervention



Key Assumption

- “But for” assumption:
  - Control group identifies the trend in outcomes that would have happened in the absence of the treatment
- For this reason it is necessary to ensure that the control group does not differ on any pre-treatment observables that might explain variation in post-treatment trends

Defining a Control Group

- Sometimes it is obvious
  - (i.e., if a program is rolled out in a random selection of counties within a state, the unselected counties make a natural control group)
- If treatment is not random, you have to account for idiosyncratic selection into treatment when constructing control group
  - Propensity score matching
  - Synthetic controls (in cases where no viable control group is available).

VARIABLES	DV: Boycott
Boycotts against Firm in Prior 3 Years	0.505***
	(0.143)
Boycotts against Industry Peers in Prior 3 Years	0.163*
	(0.067)
Lagged Logged Employees	0.892***
	(0.092)
Policy Liberalism of HQ State Legislature	0.295**
	(0.105)
KLD Reputation Strengths	0.009
	(0.043)
KLD Reputation Weaknesses	0.097*
	(0.044)
Average Annual Turnover in Prior Three Years	-1.005
	(2.090)
Fixed Effects for Year	YES
Constant	-10.572
	(1.28)
Number of Observations	18054
--	
Robust standard errors in parentheses	
*** p<0.001, ** p<0.01, * p<0.05	

- From model, predicted log-odds of being boycotted for each firm-year
- Used nearest-neighbor propensity score matching to match boycotted firms with control firms that had a similar risk of being boycotted
- Can maximize efficiency but reduce bias with a caliper restriction
  - typically .25 s.d. of the log-odds

Validation Checks

- Ensure that the control group does not differ on any pre-treatment observables that might explain variation in post-treatment trends

Assessing the Quality of the Matched Sample in McDonnell & Cobb (Forthcoming)

	Group A Entire Sample			Group C -- 1:2 Nearest Neighbor Within-Year within Caliper		
	N (Treatment) = 120 N (Control) = 17941			N (Treatment) = 111 N (Control) = 217		
	Mean, Treatment	Mean, Control	t-test	Mean, Treatment	Mean, Control	t-test
Average Annual Turnover in Prior 3 Yrs	0.088	0.080	0.23	0.088	0.088	0.99
Boycotts of Firm in Prior 5 Years	0.529	0.023	**0.00**	0.450	0.309	0.11
Boycotts of Industry in Prior 5 Years	0.756	0.225	**0.00**	0.667	0.737	0.74
Logged Employees	4.498	2.107	**0.00**	4.307	4.261	0.75
State Policy Liberalism	0.842	0.596	0.06	0.889	1.102	0.17
KLD Reputation Strengths	4.202	1.639	**0.00**	4.369	4.548	0.74
KLD Reputation Weaknesses	4.824	1.709	**0.00**	4.676	4.281	0.35
Logged Assets	10.000	7.892	**0.00**	9.947	9.668	0.18
ROA	0.045	0.035	0.23	0.046	0.057	0.23

Validation Checks

- Verify the parallel trends assumption

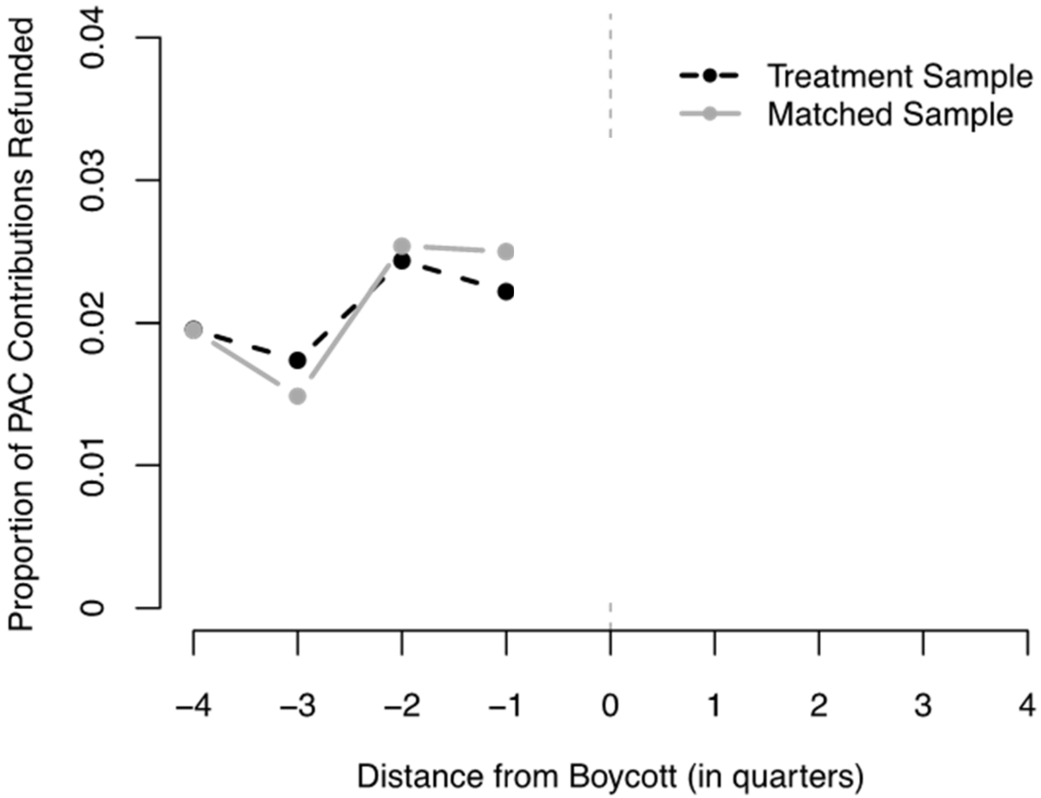


*Figure from McDonnell & Werner, 2016

Simplest (Two-period) Model

- Data varies by
  - observation (i)
  - time (t)
  - Outcome is Y_{it}
- Only two periods; structure as a panel
- Intervention will occur in a group of observations (e.g. states, firms, people, etc.)

- Three key variables
  - T_{it} =1 if obs i is eventually treated
  - A_{it} =1 in the post-treatment period
  - $T_{it}A_{it}$ -- interaction term produces the **difference estimator** (i.e., treatment effect)
- $Y_{it} = \beta_0 + \beta_1 T_{it} + \beta_2 A_{it} + \beta_3 T_{it}A_{it} + \varepsilon_{it}$

McDonnell & Cobb (Forthcoming):
Difference-in-difference Models Predicting Board Turnover after Social Movement Boycotts

	Full RiskMetrics Sample	1:2 Propensity-Score-Matched Model
VARIABLES	3-Year Turnover	3-Year Turnover
Treatment: Boycott	-0.013	-0.917
	(0.091)	(0.716)
Period Indicator	0.015	0.026
	(0.010)	(0.058)
Difference Estimator	0.328**	0.360**
	(0.152)	(0.201)
Firm-Boycott-level Fixed Effects	NO	YES
Constant	-1.193***	-1.370***
	(0.007)	(0.261)
Observations	37309	656
Robust standard errors in parentheses		
*** p<0.001, ** p<0.01, * p<0.05		

Best practice to also demonstate the effects graphically

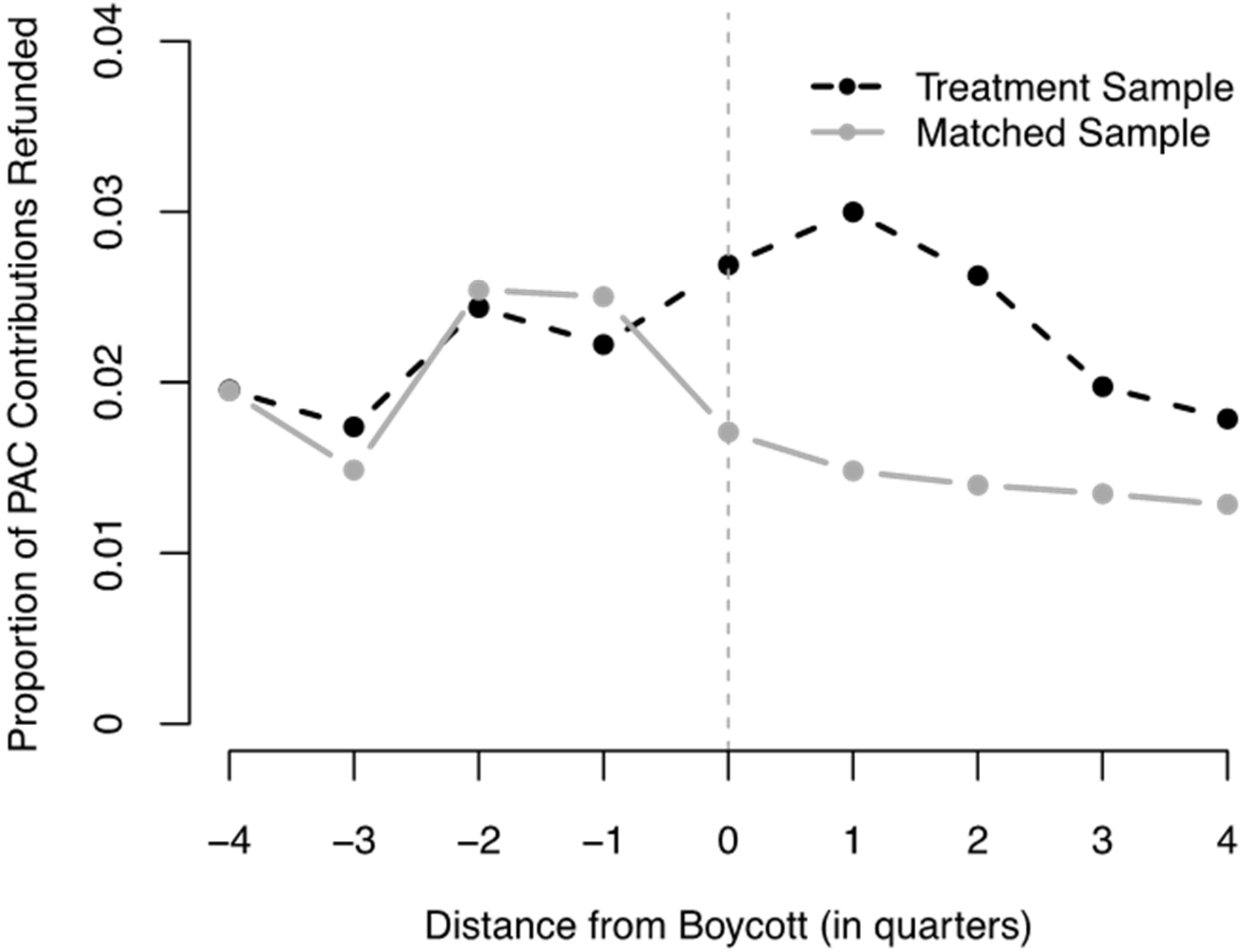


*Figure from McDonnell & Werner, 2016

Robustness Checks

- Placebo in time and placebo in treatment
  - example from Marie & Zolitz (2015): “High” Achievers
  - Treatment: local policy that only allowed Dutch, Belgian and German nationalities to purchase marijuana locally

Table 7: Placebo in Policy Timing and Treated Group

	Placebo Specification			
	Policy 1 Year Earlier		**Belgians Treated Group**	
	Std. Grade	**Passed**	**Std. Grade**	**Passed**
Placebo Policy Effect	-0.0129 (0.030)	-0.0004 (0.013)	0.0103 (0.048)	0.0284 (0.022)
All Controls and FEs	Yes	Yes	Yes	Yes
Observations	34,325	34,325	48,762	48,762

Note: The controls and Fes included in all specifications are as in the last column of Table 3 (i.e., age in months, number of courses enrolled in, teaching period dummies, and a cubic in time trend, course specific Fes, and student specific Fes). Robust standard errors clustered at the nationality level are reported in parenthesis. * and ** indicate significance at the 5 and 1 percent level, respectively.