10) Differences-in-Differences

Vitor Kamada

August 2018

First-Differenced Estimator

$$y_{it} = \beta_0 + \delta_0 d2_t + \beta_1 x_{it} + \alpha_i + u_{it}$$

$$y_{i2} = (\beta_0 + \delta_0) + \beta_1 x_{i2} + \alpha_i + u_{i2} \quad [\text{t=2}]$$

$$y_{i1} = \beta_0 + \beta_1 x_{i1} + \alpha_i + u_{i1} \quad [\text{t=1}]$$

$$y_{i2} - y_{i1} = \delta_0 + \beta_1 (x_{i2} - x_{i1}) + u_{i2} - u_{i1}$$

$$\Delta y_i = \delta_0 + \beta_1 \Delta x_i + \Delta u_i$$

Difference-in-Differences Estimator

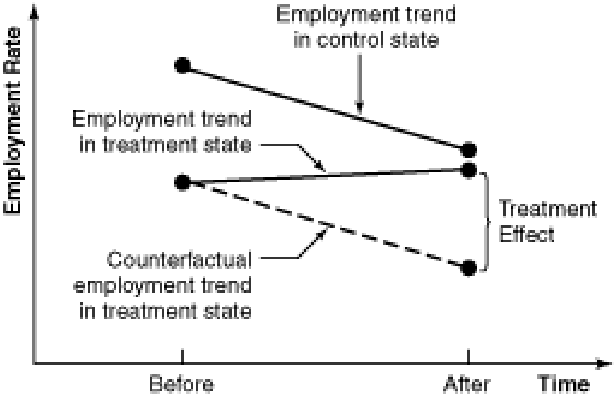
$$y = \beta_0 + \delta_0 d2 + \beta_1 dT + \delta_1 d2 \cdot dT + u$$

$$\hat{\delta}_1 = (\bar{y}_{2,T} - \bar{y}_{2,C}) - (\bar{y}_{1,T} - \bar{y}_{1,C})$$

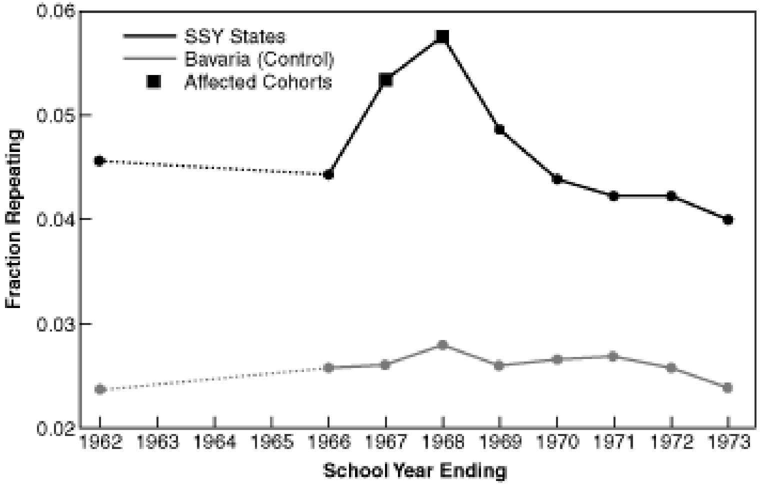
$$\hat{\delta}_1 = (\bar{y}_{2,T} - \bar{y}_{1,T}) - (\bar{y}_{2,C} - \bar{y}_{1,C})$$

	Before	After	After-Before
Control	β_0	$\beta_0 + \delta_0$	δ_0
Treatment	$\beta_0 + \beta_1$	$\beta_0 + \delta_0 + \beta_1 + \delta_1$	$\delta_0 + \delta_1$
Treatment-Control	β_1	$\beta_1 + \delta_1$	δ_1

Angrist & Pischke (2009)



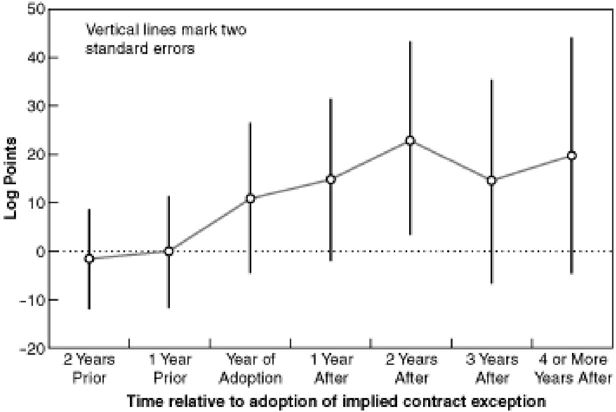
The Effect of School Term Length on Student Performance



Pischke (2007)

$$y_{ist} = \gamma_s + \lambda_t + \sum_{\tau=0}^{m} \delta_{-\tau} D_{s,t-\tau} + \sum_{\tau=1}^{q} \delta_{+\tau} D_{s,t+\tau} + X'_{ist}\beta + \epsilon_{ist}$$

The Impact of Implied-Contract Exceptions on the Use of Temporary Workers



(Autor, 2003)

State-Specific Time Trends

$$y_{ist} = \gamma_{0s} + \gamma_{1s}t + \lambda_t$$

$$+\delta_\tau D_{st} + X'_{ist}\beta + \epsilon_{ist}$$

The Effects of Labor Regulation on Productivity

All models include state and year effects.
Robust standard errors clustered at the state level.

	(1)	(2)	(3)	(4)
Labor regulation (lagged)	−.186	−.185	−.104	.0002
	(.064)	(.051)	(.039)	(.020)
Log development expenditure per capita		.240	.184	.241
		(.128)	(.119)	(.106)
Log installed electricity capacity per capita		.089	.082	.023
		(.061)	(.054)	(.033)
Log state population		.720	0.310	−1.419
		(.96)	(1.192)	(2.326)
Congress majority			−.0009	.020
			(.01)	(.010)
Hard left majority			−.050	−.007
			(.017)	(.009)
Janata majority			.008	−.020
			(.026)	(.033)
Regional majority			.006	.026
			(.009)	(.023)
State-specific trends	No	No	No	Yes
Adjusted R^2	.93	.93	.94	.95

Besley and Burgess (2004)

Fixed Effects versus Lagged Dependent Variables

$$E[y_{0it}|\alpha_i, X_{it}, D_{it}] = E[y_{0it}|\alpha_i, X_{it}] \qquad (1)$$

$$E[y_{0it}|y_{i,t-h}, X_{it}, D_{it}] = E[y_{0it}|y_{i,t-h}, X_{it}] \qquad (2)$$

Fixed Effects and Lagged Dependent Variables

$$E[y_{0it}|\alpha_i, y_{i,t-h}, X_{it}, D_{it}] = E[y_{0it}|\alpha_i, y_{i,t-h}, X_{it}]$$

$$Y_{it} = \alpha_i + \theta Y_{i,t-h} + \lambda_t + \delta D_{it} + X_{it}'\beta + \epsilon_{it}$$

$$\Delta Y_{it} = \theta \Delta Y_{i,t-1} + \Delta\lambda_t + \delta \Delta D_{it} + \Delta X_{it}'\beta + \Delta\epsilon_{it}$$

Both $\Delta Y_{i,t-1}$ and $\Delta\epsilon_{it}$ are a function of $\epsilon_{i,t-1}$

True Model: Fixed Effects

$$Y_{it} = \alpha_i + \delta D_{it} + \epsilon_{it}$$

$$Y_{i,t-1} = \alpha_i + \epsilon_{i,t-1}$$

$$Y_{it} = Y_{i,t-1} + \delta D_{it} + \epsilon_{it} - \epsilon_{i,t-1}$$

$$\frac{Cov(Y_{it}, \tilde{D}_{it})}{V(\tilde{D}_{it})} = \delta - \frac{Cov(\epsilon_{i,t-1}, \tilde{D}_{it})}{V(\tilde{D}_{it})} = \delta + \frac{\gamma \sigma_\epsilon^2}{V(\tilde{D}_{it})}$$

where $\tilde{D}_{it} = D_{it} - \gamma Y_{i,t-1}$ is the residual from a regression of D_{it} on $Y_{i,t-1}$

True Model: Lagged Dependent Variables

$$Y_{it} = \alpha + \theta Y_{i,t-1} + \delta D_{it} + \epsilon_{it}$$

$$Y_{it} - Y_{i,t-1} = \alpha + (\theta - 1) Y_{i,t-1} + \delta D_{it} + \epsilon_{i,t}$$

$$\frac{Cov(Y_{it} - Y_{i,t-1}, D_{it})}{V(D_{it})} = \delta + (\theta - 1)\left[\frac{Cov(Y_{i,t-1}, D_{it})}{V(D_{it})}\right]$$