Revisiting the Effects of Cigarette Taxes on Smoking Outcomes

Vinish Shrestha

TOPS Presentation.

2023-03-10

Disclosures

1. No funding was obtained for this work by the author.
2. No tobacco-related funding has been acquired by the author in the past 10 years.

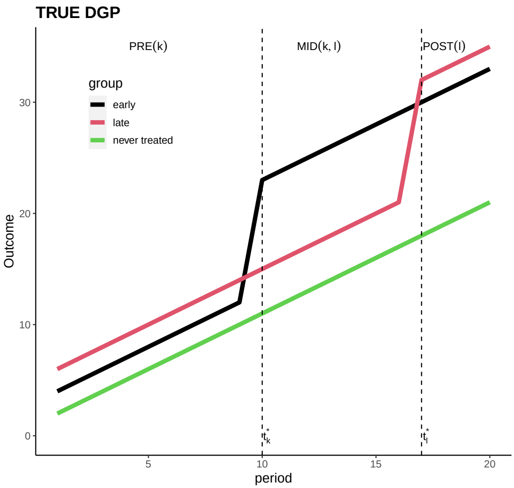
Note: I am happy to share codes for replication. The codes will eventually be posted on my personal website.

TOC

Section 1

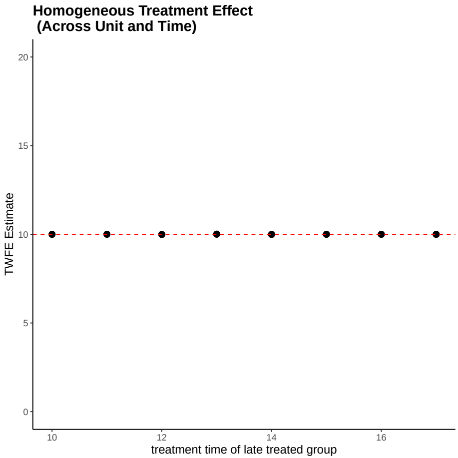
Illustrations of 3 cases

Case 1 (Homogeneous Treatment Effects across Units and Time)

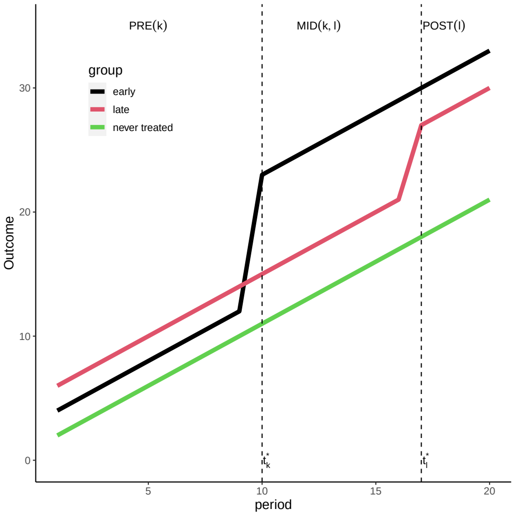


Simulation Results (varying treatment time of later unit)

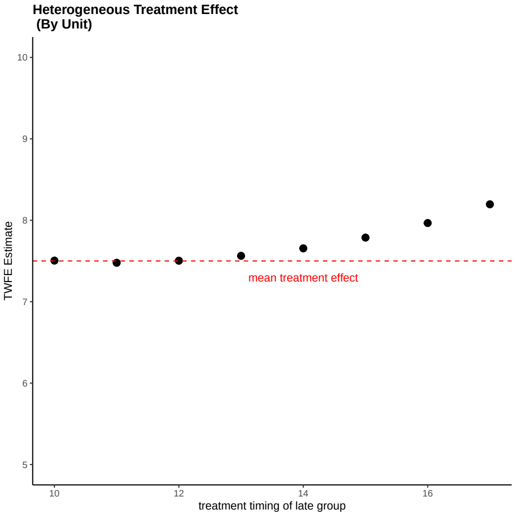
TWFE: $Y_{it} = \beta D_{it} + \eta_i + \theta_t + \epsilon_{it}$



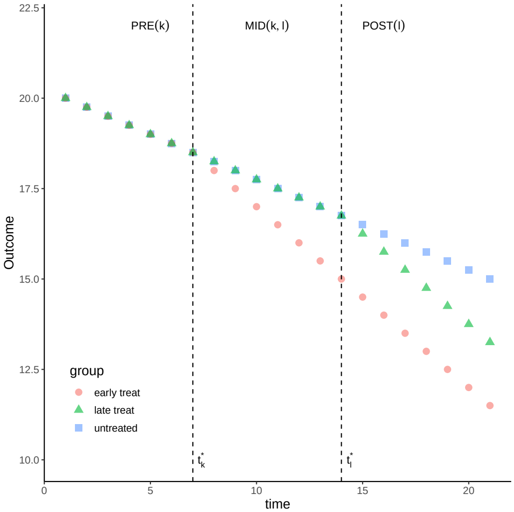
Case 2 (Heterogeneous Treatment Effects by Units)



Simulation Results (varying treatment time of later unit)



Case 3 (Heterogeneous Treatment Effects by Time)



Some Realizations

1. Homogeneous treatment effects
   - TWFE works fine
2. Heterogeneous treatment effects across unit
   - TWFE can be incorrect
   - depends on treatment timing
3. Heterogeneous treatment effects over time
   - early treated units acting as control for later treated units
   - "bad comparison"
   - *negative weighting problem*

- In cases 2 and 3 $TW\hat{F}E \neq A\hat{T}T$ (average treatment effect on the treated estimate)

Section 2

Motivation and Main Findings

Motivation

- Cigarette taxes widely used as a policy instrument
  - reduce smoking and increase revenue
- Research heavily rely on TWFE specifications (Review DeCicca, Kenkel, and Lovenheim (2020))
  - "... an important issue for the analysis of cigarette taxes that has not been sufficiently explored by researchers"

TWFE specification

$$smoking_{st} = \alpha + \beta \times tax_{st} + \theta_t + \eta_s + \epsilon_{st}$$

- Continuous measure of cigarette taxes (prices)
  - within unit (state) variation in cigarette taxes (prices) over time
  - multiple-treatment and multiple-control group framework (staggered framework)

Recent advancements in staggered DiD literature

- Highlights TWFE concerns (De Chaisemartin and d'Haultfoeuille (2020), Goodman-Bacon (2021), Callaway and Sant'Anna (2021), Sun and Abraham (2021), Callaway (2022))
- One main issue
  - *negative weighting problem*
  - if ATT varies with the length of exposure to treatment, then early treated group forms a "bad comparison group" for later treated units
- Particularly dire
  - if a significant number of units are eventually treated

Note: Between 2004-2010 38 states increased cigarette taxes at least once.

Study's Purpose

- Revisit the literature of cigarette taxes and smoking outcomes
- How different are the TWFE estimates from $A\hat{T}E$?
  - TWFE versus $A\hat{T}T$ from Callaway and Sant'Anna (2021) (CS estimator)
  - TWFE versus i) canonical event-study, ii) interaction-weighted estimator (Sun and Abraham (2021)), iii) event-study-type estimates (Callaway and Sant'Anna (2021))

1. Balanced panel data Behavioral Risk Factor Survelliance System Selected Metropolitan/Micropolitan Area Risk Trends (BRFSS SMART)
2. Two periods: i) 2004-2010; and ii) 2015-2020
3. TWFE specification:
   - $smoking_{st} = \alpha + \beta \times tax_{st} + \theta_t + \eta_s + \epsilon_{st}$
   - $tax_{st} \in \{0,\ 1\}$ (binary treatment)

Main Findings

- Different approaches demonstrate effectiveness of tax incidence in reducing smoking-related outcomes

1. |TWFE estimate| $<$ $|\hat{ATT}|$ from CS estimator
   - 2004-2010 period: TWFE estimate is about 65% of the overall $\hat{ATT}$ from CS
2. Decomposition of TWFE following Goodman-Bacon (2021) shows huge weight (32%) is placed on cases that use *later treated units* in comparison to *early treated units* in 2004-2010 sample
   - Not too bad in 2015-2020 sample (4.7%)
3. Canonical event study, SA approach, and CS event-study type estimates all show gradual but effects increasing in magnitude over time
4. $|\hat{ATT}_{2015-2020}|$ only 63% of $|\hat{ATT}_{2004-2010}|$

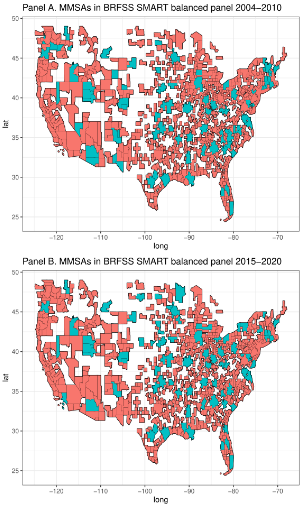
Section 3

Data

BRFSS SMART

- Behavioral Risk Factor Survelliance System (BRFSS) Selected Metropolitan/Micropolitan Area Risk Trends (SMART)
    - years 2004-2010 and 2015-2020
- Smart project initiated to produce local areas defined as Metropolitan/Micropolitan *(MMSAs) == locality of interest*
- Each MMSAs include at least 500 individuals
- The number of MMSAs vary by year
    - 134 in 2004, while 198 in 2010 (entry and exit)
- Focus on the status of *current smoker* as the outcome variable
- Create a balanced panel of the percent of current smokers collapsed at the MMSA-year level

MMSA map (balanced panel)



- green MMSAs are covered in the BRFSS SMART balanced panel
- at least 1 MMSA for 46 states; more than 2 MMSAs in many states
- 108 and 95 MMSAs in balanced panel 2004-2010 and 2015-2020
- states not represented: Alaska, Hawai, North Dakota, Rhode Island

Change (Increase) in cigarette taxes as treatment

- Tax Burden of Tobacco for years 1970-2019 (prepared by Orzechowski and Walker)
- Binary variable to represent tax change within state
  - treatment assignment
  - “tax change year” takes a value *1* and MMSAs within the state retain this value

- A handful of states with multiple tax increases
  - PA in July 2004 and November 2009
  - both fall within 2004-2010 survey year
  - use the first one to denote the treatment assignment

Table 1. States with tax changes by year

year	states	count of MMSAs	average tax increase
2004	AL, HW, MI, NJ, PA, RI, VA	108	0.26
2005	AK, CO, KY, ME, MN, MT, NC, NH, OH, OK, WA	108	0.49
2006	AZ, IA, VT	108	0.67
2007	CT, DE, IN, SD, TN, TX	108	0.75
2008	DC, MA, MD, NY, WI	108	0.97
2009	AR, FL, MS	108	0.74
2010	NM, SC, UT	108	0.75
2015	DC, KS, LA, NV, OH, RI, VT	95	0.53
2016	AL, CT, PA, WV	95	0.51
2017	CA	95	2
2018	DE, KY, OK	95	0.75
2019	IL, NM	95	0.78
2020	VA	95	0.3

Other variables

- *Tobacco Control Variable:* The percentage of a state's population under a bar ban
  - American Nonsmokers' Rights Foundation (ANRF)

Pre-treatment variables (posttreatment bias Rosenbaum (1984))

- Locality specific unemployment rate for 2000 and 2010
  - Merged Outgoing Rotation Group Earnings Data (2000 and 2010)
- CPS tobbaco supplement
  - Anti-smoking sentiment measure 1998-1999
    - in spirit of DeCicca et al. (2008)
    - collapsed at the locality level
  - Change in the proportion of current smokers between 1998-1999 and 2001-2002

Section 4

Method: TWFE

Method 1 (TWFE: explanation borrowed from Roth et al. (2022))

$$Y_{it} = \beta D_{it} + \theta_t + \eta_i + \epsilon_{it}, ...i)$$

$$\text{Also, } Y_{it}(g) = Y_{it}(0) + \tau_{it}(g), ...ii)$$

Using Frisch-Lovell Theorem:

$$\hat{\beta} = \sum_i \sum_t \frac{(D_{it} - \hat{D}_{it})(Y_{it})}{(D_{it} - \hat{D}_{it})^2}, ...iii)$$

$$\text{where, } \hat{D}_{it} = \bar{D}_i + \bar{D}_t - \bar{D}$$

- weight is proportional to $(D_{it} - \hat{D}_{it})$
- For *early treated units*: $\bar{D}_i \approx 1$
- If eventually almost all units are treated then $\bar{D}_t \approx 1$ towards the end period
- So, towards the end period: $\hat{D}_{it} > 1$ as $\bar{D} < 1$
- Numerator $(D_{it} - \hat{D}_{it})$ negative even if $D_{it} = 1$
    - puts negative weight on $\tau_{it}(g)$

##

$(D_{it} - \hat{D}_{it})$ for units treated in 2005 and 2006

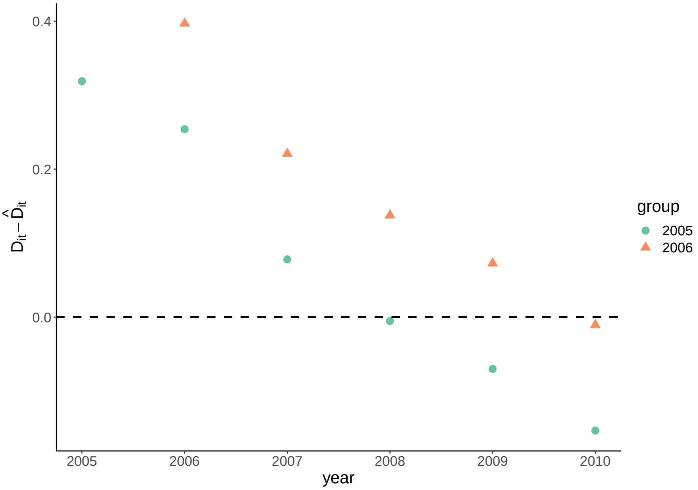


Table 2. TWFE estimate decomposed following Goodman-Bacon (2021)

	2004-2010		2015-2020	
type	weight	avg.estimate	weight	avg.estimate
Earlier vs Later Treated	0.218	-0.898	0.029	-0.048
Later vs Always Treated	0.177	-0.296	0.130	-0.471
Later vs Earlier Treated	0.316	-0.233	0.047	0.350
Treated vs Untreated	0.290	-0.834	0.790	-0.549

Note: Summary of Goodman Bacon decomposition of TWFE estimate as all possible 2 times 2 DiD estimates summarized by groups in column 1.

Method: static and dynamic TWFE

1. TWFE (Static)

$$Y_{ist} = \alpha + \beta D_{ist} + \eta_i + \theta_t + \epsilon_{it} \tag{1}$$

2. TWFE canonical event study (Dynamic)

$$Y_{ist} = \alpha + \underbrace{\sum_{k=-K}^{L}}_{k \neq \{E, -1\}} \gamma_k D_{ist}^k + \eta_i + \theta_t + \epsilon_{it} \tag{2}$$

- $1(t - g_i = k) = D_{st}^k$; relative time indicator away from policy year g_i
- omitted category include E and year before the treatment

Section 5

Method: Alternatives to TWFE

Group time ATT

period first treated	units
2005	
2006	S1, S2
2007	S3, S4, S5
2008	

- say, S0 is never treated
- define group g as units first treated in period g

Group(g) time(t) ATT

- $ATT_{g=2006,t=2006}$; $ATT_{g=2006,t=2007}$; $ATT_{g=2006,t=2008}$
- $ATT_{g=2007,t=2007}$; $ATT_{g=2007,t=2008}$

Callaway and Sant'Anna Estimator (Callaway and Sant'Anna (2021))

- Identify group-time ATT

$$ATT(g,\ t) = E(Y_t(g) - Y_t(0)|G_g = 1) \tag{3}$$

Under a) unconditional parallel trend assumption b) no-anticipation

$$A\hat{T}T(g,\ t = t^*) = [\bar{Y}_{t^*}(g) - \bar{Y}_{pretreat}(g)] - [\bar{Y}_{t^*}(C) - \bar{Y}_{pretreat}(C)]$$

$$A\hat{T}T(g,\ t) = \underbrace{\frac{\sum_i (Y_{i,t}.1(G_i = g) - Y_{i,g-1}.1(G_i = g)}{\sum_i 1(G_i = g)}}_{group\ g\ before\ \&\ after} - \underbrace{\frac{\sum_i (Y_{i,t}.1(G_i = C) - Y_{i,g-1}.1(G_i = C)}{\sum_i 1(G_i = C)}}_{group\ C\ before\ \&\ after} \quad ($$

- C can include i) never treated; or ii) not-yet-treated (until t) show results

CS Doubly Robust Estimator

- parallel trend satisfied conditional upon pretreatment covariates

$$\hat{ATT}(g,\ t) = \frac{1}{N}\sum_{i}\left[\left(\frac{1.(G_i = 1)}{\frac{\sum_i 1.(G_i=g)}{N}} - \frac{\frac{\hat{p}_g(X)1.(G_i=C)}{1-\hat{p}_g(X).1(G_i=C)}}{\frac{1}{N}\sum_i \frac{\hat{p}_g(X)1.(G_i=C)}{1-\hat{p}_g(X).1(G_i=C)}}\right)(Y_{i,t} - Y_{i,g-1} - \hat{m}_{g,t}(X))\right] \quad (5)$$

Combines 1) IPW (Abadie (2005)) 2) Outcome Regression (Heckman, Ichimura, and Todd (1997))

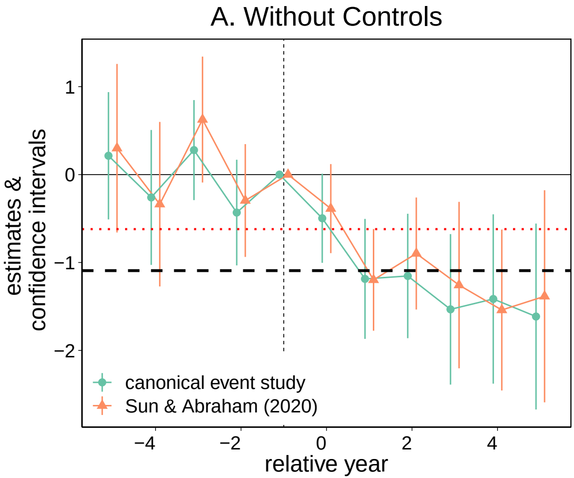
- These $\hat{ATT}(g\ ,t)$ are then aggregated to form *i) event study type estimates* and *ii) point estimate* $\hat{ATT}$

Section 6

Results (using parsimonious specification)

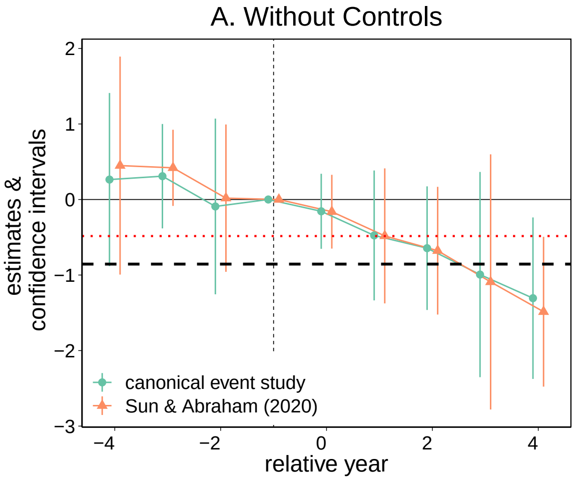
R1. TWFE and Event Study Estimates (2004-2010 Sample)

Note: i) red dot = TWFE static estimate, ii) green = Canonical event study estimates, iii) orange = SA event study estimates, iv) black dash = average of estimates from canonical event study estimates



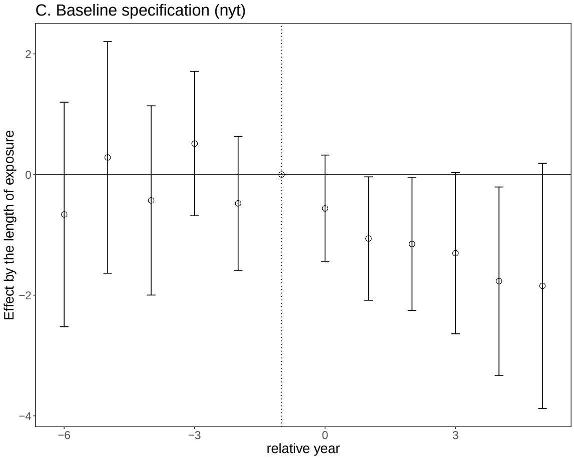
R2. TWFE and Event Study Estimates (2015-2020 Sample)

Note: i) red dot = TWFE static estimate, ii) green = Canonical event study estimates, iii) orange = SA event study estimates, iv) black dash = average of estimates from canonical event study estimates



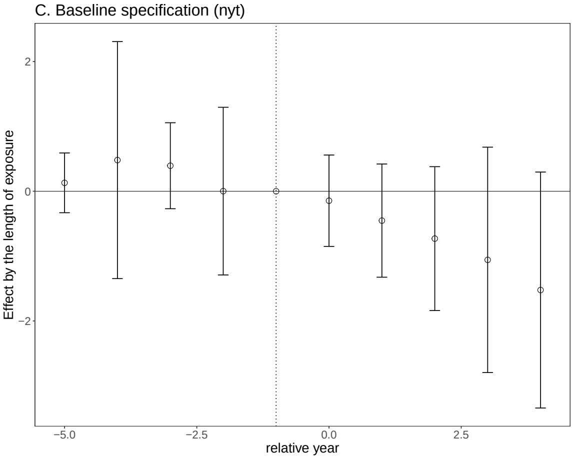
R2. CS Event-Study-Type Estimates (2004-2010 Sample)

Note: The analysis use not-yet-treated units (nyt) as the comparison.



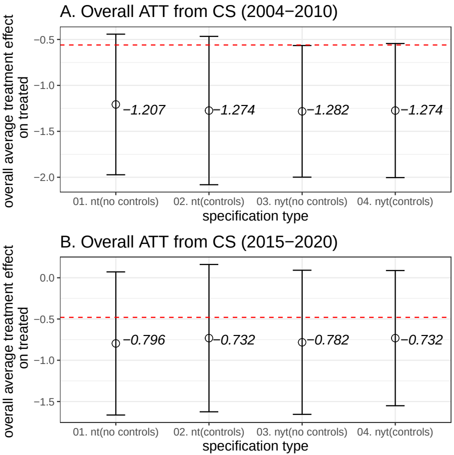
R2. CS Event-Study-Type Estimates (2015-2020 Sample)

Note: The analysis use not-yet-treated units (nyt) as the comparison.



R3. TWFE and $\hat{ATT}$ from Callaway and Sant'Anna (2021)

Note: The red dashed line is the TWFE estimate. The $\hat{ATT}$ are obtained from aggregating the group time ATT estimates.



Section 7

Conclusion

Some concluding remarks

- Cigarette tax are an effective means of reducing smoking prevalence
  - *consistent with earlier studies*
- However, TWFE estimates tend to be biased downwards in magnitude
  - particularly in a sample when the treatment is of multiple time-multiple group and the majority of units are eventually treated
- Canonical event study estimates capture heterogeneity by time
  - estimates are similar to CS-type event study and SA-type event study

Using point estimates of ATT that respects treatment heterogeneity can increase the magnitude of the elasticity estimates (until now the elasticity estimates are mainly based on TWFE estimates)

References I

Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *The Review of Economic Studies* 72 (1): 1–19.

Callaway, Brantly. 2022. "Difference-in-Differences for Policy Evaluation." *arXiv Preprint arXiv:2203.15646*.

Callaway, Brantly, and Pedro HC Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225 (2): 200–230.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110 (9): 2964–96.

DeCicca, Philip, Donald S Kenkel, and Michael F Lovenheim. 2020. "The Economics of Tobacco Regulation: A Comprehensive Review."

DeCicca, Philip, Donald Kenkel, Alan Mathios, Yoon-Jeong Shin, and Jae-Young Lim. 2008. "Youth Smoking, Cigarette Prices, and Anti-Smoking Sentiment." *Health Economics* 17 (6): 733–49.

References II

Goodman-Bacon, Andrew. 2021. "Difference-in-Differences with Variation in Treatment Timing." *Journal of Econometrics* 225 (2): 254–77.

Heckman, James J, Hidehiko Ichimura, and Petra E Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *The Review of Economic Studies* 64 (4): 605–54.

Rosenbaum, Paul R. 1984. "The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment." *Journal of the Royal Statistical Society: Series A (General)* 147 (5): 656–66.

Roth, Jonathan, Pedro HC Sant'Anna, Alyssa Bilinski, and John Poe. 2022. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *arXiv Preprint arXiv:2201.01194*.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." *Journal of Econometrics* 225 (2): 175–99.