渐进DID模型的理论计量进展

宗庆庆
上海财经大学公共经济与管理学院
zong.qingqing@mail.shufe.edu.cn

2023年11月

Difference in Difference

Basic Setup

- $t \in \{0,1\}$: t=0 if before the treatment time; t=1 if after
- $r \in \{0,1\}$: r=0 if in control group; r=1 if in treated group
- $\tau = t \cdot r$ is the treatment. So τ=1 only if t=1 and r=1.
- Like in experiments, only treated group after the treatment is treated.
- $Y_{\tau t}$ is the potential outcome in time t when treatment status is τ.
- So, $E(Y_{11} - Y_{01}|r = 1)$ is the interested ATT.
- There're four observed volumes (two groups in two times):

➢ $E(Y_{11}|r = 1)$: treatment group + after the treatment
➢ $E(Y_{00}|r = 1)$: treatment group + before the treatment
➢ $E(Y_{01}|r = 0)$: control group + after the treatment
➢ $E(Y_{00}|r = 0)$: control group + before the treatment

Identification

- We use the control group to infer the time effect of treated group. Then we can figure out the treatment effects.

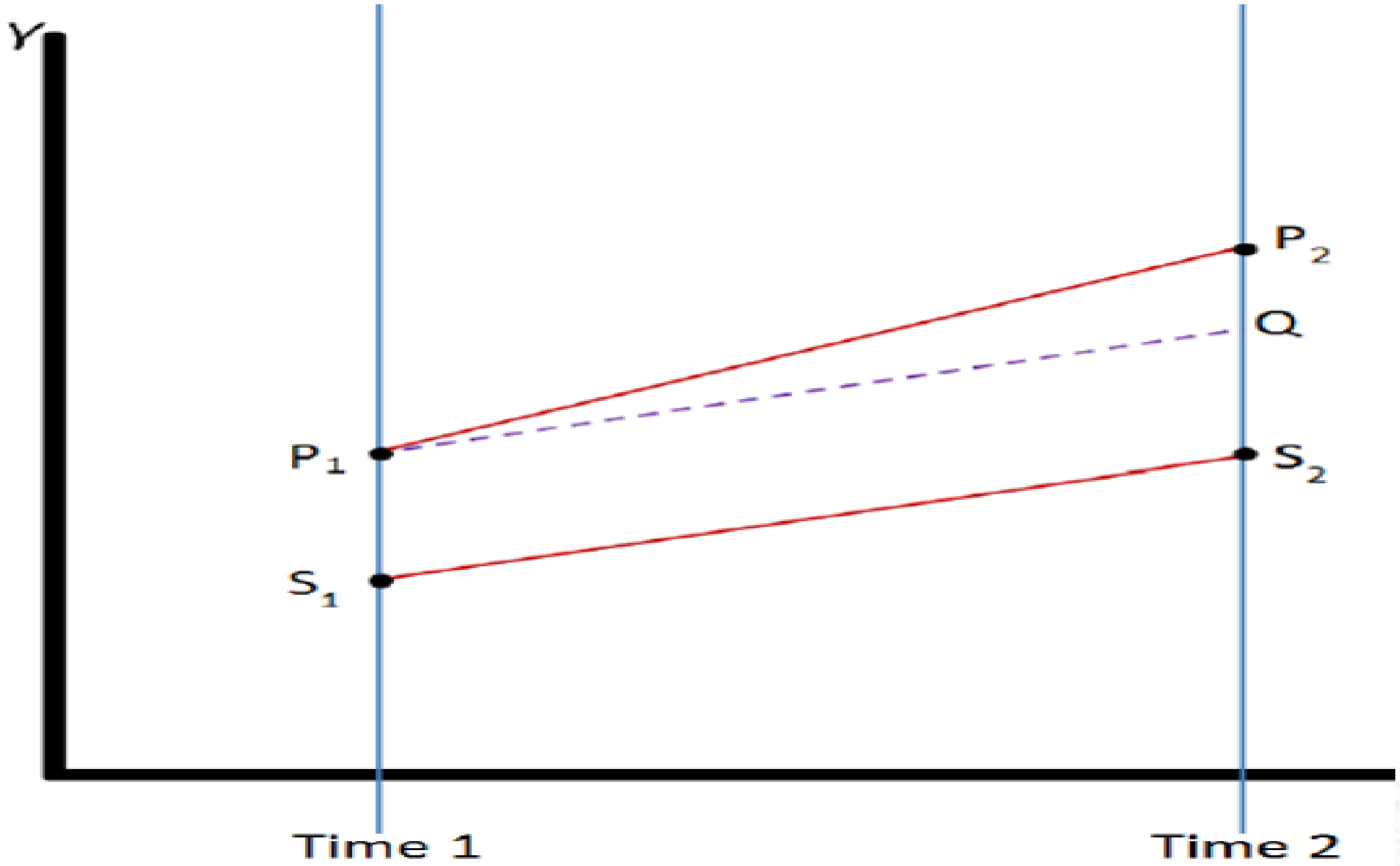
$$\begin{aligned}\text{DID} &= [E(Y|r=1,t=1)-E(Y|r=1,t=0)]-[E(Y|r=0,t=1)-E(Y|r=0,t=0)] \\ &= E(Y_{11}-Y_{00}|r=1)-E(Y_{01}-Y_{00}|r=0) \\ &= E(Y_{11}-Y_{01}|r=1)+[E(Y_{01}-Y_{00}|r=1)-E(Y_{01}-Y_{00}|r=0)]\end{aligned}$$

- DID is the interested ATT ($E(Y_{11}-Y_{01}|r=1)$) if the **Parallel Trend Assumption** holds

$$E(Y_{01}-Y_{00}|r=1)=E(Y_{01}-Y_{00}|r=0)$$

- The assumption means the treatment group shares the same time trend with control group, or the time trend in the absence of the intervention are the same in both groups

Parallel Trend Assumption



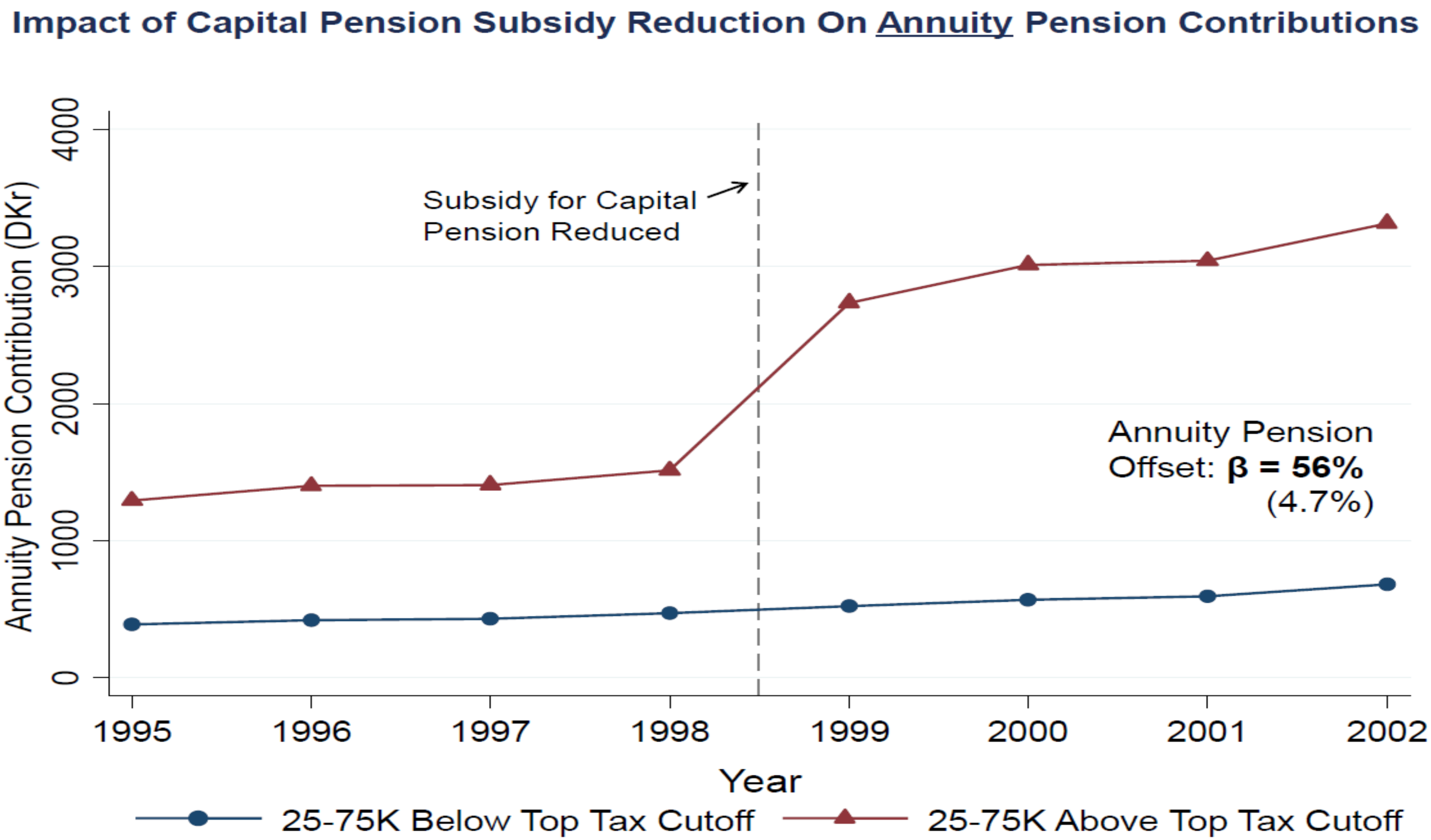
Estimation

- $\widehat{DID} = \left[\bar{Y}_{t_1,treated} - \bar{Y}_{t_0,treated}\right] - [\bar{Y}_{t_1,control} - \bar{Y}_{t_0,control}]$
- Linear form: $y = \beta_0 + \beta_t t + \beta_r r + \delta\tau + u$; by OLS $\Rightarrow \hat{\delta}$
- $\hat{\delta} = \widehat{DID}$. OLS or fixed effects can be used to estimate DID.

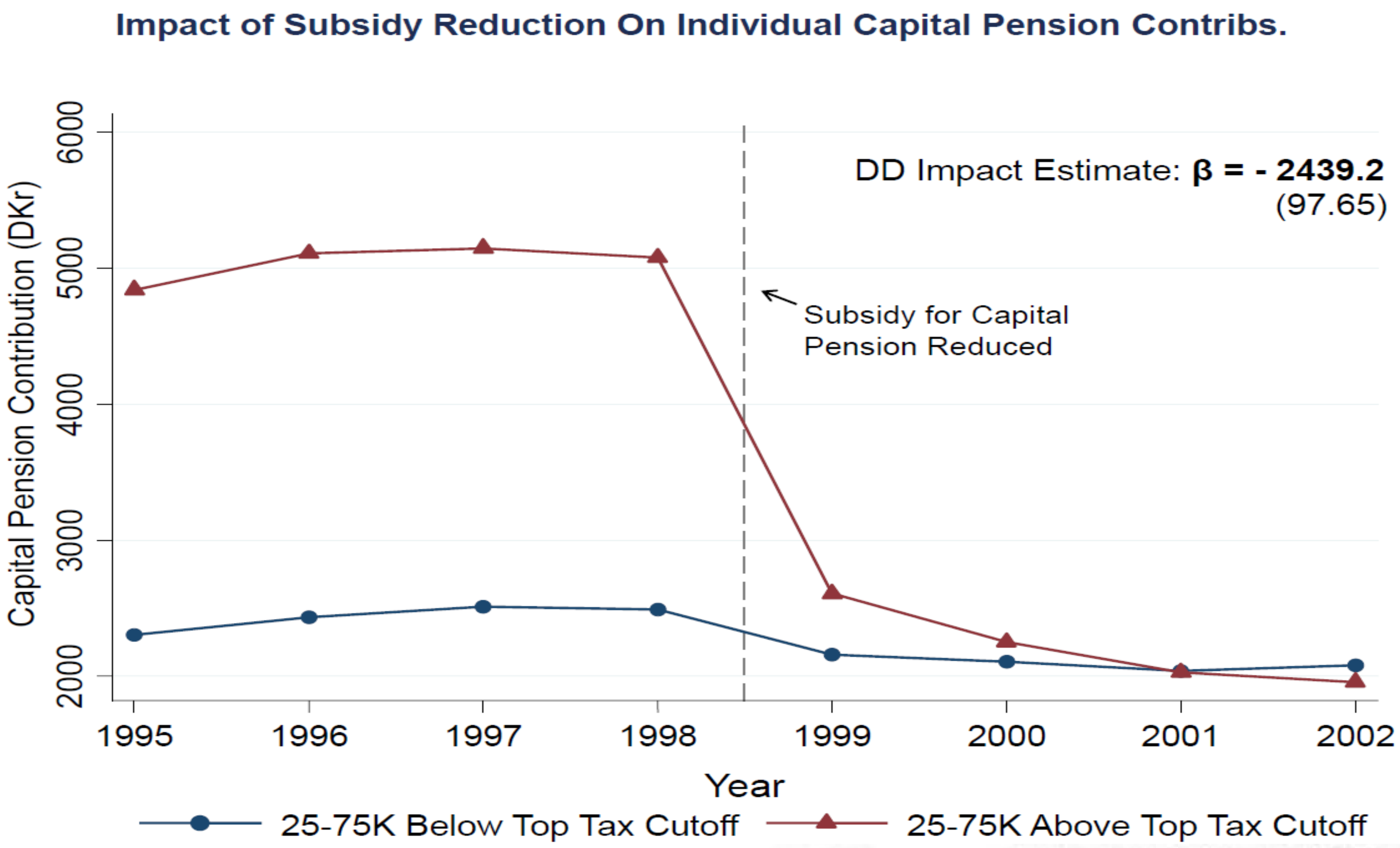
Why δ is DID Estimator?

	Before (t=0)	After (t=1)	Difference
Treatment (r=1)	$\beta_0+\beta_r$	$\beta_0+\beta_r+\beta_t+\delta$	$\Delta Y_t=\beta_t+\delta$
Control (r=0)	β_0	$\beta_0+\beta_t$	$\Delta Y_c=\beta_t$
Difference	β_1	$\beta_1+\delta$	$\Delta\Delta Y=\delta$

Test Parallel Trend



Chetty et al.(2014,QJE)



How Much Should We Trust DID?

Statistical Inference of DID

- Bertrand Mariance,Esther Duflo,Sendhil Mullainathan,2004,How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics*, 119(1),249-275.
- Because of **serial correlation**, DD estimation as it is commonly performed grossly under-states the standard errors around the estimated intervention effect.

① First, DD estimation usually relies on fairly long time periods

② Second, the most commonly used dependent variables in DD estimation are typically highly positively serially correlated

③ Third, the treatment variable changes itself very little within a state over time

- These three factors reinforce each other to create potentially large mis-measurement in the standard errors coming from the OLS estimation.

Survey of DID Paper

- Data comes from a survey of all articles in six journals between 1990 and 2000: *American Economic Review*; *Industrial Labor Relations Review*; *Journal of Labor Economics*; *Journal of Political Economy*; *Journal of Public Economics*; and *Quarterly Journal of Economics*.

- They define an article as "Difference-in-Difference" if it: (1) examines the effect of a specific interventions and (2) uses units unaffected by the intervention as a control group.

- Their survey of DD papers, which we discuss below, finds an average of **16.5** periods.

Survey of DID Paper (cont'd)

Number of DD papers	92	
Number with more than 2 periods of data	69	
Number which collapse data into before-after	4	
Number with potential serial correlation problem	65	
Number with some serial correlation correction	5	
GLS	4	
Arbitrary variance-covariance matrix	1	
Distribution of time span for papers with more than 2 periods	Average	16.5
	Percentile	Value
	1%	3
	5%	3
	10%	4
	25%	5.75
	50%	11
	75%	21.5
	90%	36
	95%	51
	99%	83

Over-rejection in DID Estimation

- A sample of women's wages from the Current Population Survey (CPS).
- 1979-1999, all women between the ages 25 and 50
- The sample contains nearly 900,000 observations, approximately 540,000 report strictly positive weekly earnings
- Dependent variable: log(weekly earnings)

① Draw a year at random from a uniform distribution between 1985 and 1995 (ensure having enough observations prior and post-intervention)

② Select exactly half the states (25) at random and designate them as "affected" by the law

③ Estimate DD 200 times where the control variables contain education ,age , state dummies and year dummies

Over-rejection in DID Estimation (cont'd)

A. CPS DATA

Data	$\hat{\rho}_1, \hat{\rho}_2, \hat{\rho}_3$	Modifications	Rejection rate No effect	Rejection rate 2% effect
1) CPS micro, log wage			.675 (.027)	.855 (.020)
2) CPS micro, log wage		Cluster at state-year level	.44 (.029)	.74 (.025)
3) CPS agg, log wage	.509, .440, .332		.435 (.029)	.72 (.026)
4) CPS agg, log wage	.509, .440, .332	Sampling w/replacement	.49 (.025)	.663 (.024)
5) CPS agg, log wage	.509, .440, .332	Serially uncorrelated laws	.05 (.011)	.988 (.006)
6) CPS agg, employment	.470, .418, .367		.46 (.025)	.88 (.016)
7) CPS agg, hours worked	.151, .114, .063		.265 (.022)	.280 (.022)
8) CPS agg, changes in log wage	−.046, .032, .002		0	.978 (.007)

Over-rejection in DID Estimation (cont'd)

- The stylized exercise above focused on data with 50 states
- and 21 time periods.
- Many DD papers use fewer states and several DD papers use fewer time periods.
- They examined how the rejection rate varies with these two important parameters

① Varying the number of states does not change the extent of the over-rejection

② Over-rejection falls as the time span gets shorter, but it does so at a rather slow rate

Over-rejection in DID Estimation (cont'd)

TABLE III
VARYING N AND T

Data	N	T	Rejection rate: No effect	Rejection rate: 2% effect
A. CPS DATA				
1) CPS aggregate	50	21	.49 (.025)	.663 (.024)
2) CPS aggregate	20	21	.39 (.024)	.54 (.025)
3) CPS aggregate	10	21	.443 (.025)	.510 (.025)
4) CPS aggregate	6	21	.383 (.025)	.433 (.025)
5) CPS aggregate	50	11	.20 (.020)	.638 (.024)
6) CPS aggregate	50	7	.15 (.017)	.635 (.024)
7) CPS aggregate	50	5	.078 (.013)	.5 (.025)
8) CPS aggregate	50	3	.048 (.011)	.363 (.024)
9) CPS aggregate	50	2	.055 (.011)	.28 (.022)
B. MONTE CARLO SIMULATIONS WITH SAMPLING FROM AR(1) DISTRIBUTION				
10) AR(1), $\rho = .8$	50	21	.35 (.028)	.638 (.028)
11) AR(1), $\rho = .8$	20	21	.35 (.028)	.538 (.029)
12) AR(1), $\rho = .8$	10	21	.3975 (.028)	.505 (.029)
13) AR(1), $\rho = .8$	6	21	.393 (.028)	.5 (.029)
14) AR(1), $\rho = .8$	50	11	.335 (.027)	.588 (.028)
15) AR(1), $\rho = .8$	50	5	.175 (.022)	.5525 (.029)
16) AR(1), $\rho = .8$	50	3	.09 (.017)	.435 (.029)
17) AR(1), $\rho = .8$	50	50	.4975 (.029)	.855 (.020)

Solutions

- Parametric Methods: misspecification??

- Block Bootstrap: complicated!

- Ignoring Time Series Information

Parametric Methods

Data	Technique	Estimated $\hat{\rho}_1$	Rejection rate: No effect	Rejection rate: 2% Effect
A. CPS DATA				
1) CPS aggregate	OLS		.49 (.025)	.663 (.024)
2) CPS aggregate	Standard AR(1) correction	.381	.24 (.021)	.66 (.024)
3) CPS aggregate	AR(1) correction imposing $\rho = .8$		.18 (.019)	.363 (.024)
B. OTHER DATA GENERATING PROCESSES				
4) AR(1), $\rho = .8$	OLS		.373 (.028)	.765 (.024)
5) AR(1), $\rho = .8$	Standard AR(1) correction	.622	.205 (.023)	.715 (.026)
6) AR(1), $\rho = .8$	AR(1) correction imposing $\rho = .8$		.06 (.023)	.323 (.027)
7) AR(2), $\rho_1 = .55$ $\rho_2 = .35$	Standard AR(1) correction	.444	.305 (.027)	.625 (.028)
8) AR(1) + white noise, $\rho = .95$, noise/signal = .13	Standard AR(1) correction	.301	.385 (.028)	.4 (.028)

Block Bootstrap

- For each placebo intervention , compute the absolute *t*-statistic

- Construct a bootstrap sample by drawing with replacement 50 matrices (y_s,v_s)

- Run OLS on this sample, obtain an estimate beta_r_hat, t_r = abs(beta_r_hat - beta_hat)/se(beta_r_hat)

- The difference between this distribution and the sampling distribution of t becomes small as N goes to infinity,even in the presence of arbitrary autocorrelation

Block Bootstrap(cont'd)

TABLE V
BLOCK BOOTSTRAP

Data	Technique	N	Rejection rate: No effect	Rejection rate: 2% effect
A. CPS DATA				
1) CPS aggregate	OLS	50	.43 (.025)	.735 (.022)
2) CPS aggregate	Block bootstrap	50	.065 (.013)	.26 (.022)
3) CPS aggregate	OLS	20	.385 (.022)	.595 (.025)
4) CPS aggregate	Block bootstrap	20	.13 (.017)	.19 (.020)
5) CPS aggregate	OLS	10	.385 (.024)	.48 (.024)
6) CPS aggregate	Block bootstrap	10	.225 (.021)	.25 (.022)
7) CPS aggregate	OLS	6	.48 (.025)	.435 (.025)
8) CPS aggregate	Block bootstrap	6	.435 (.022)	.375 (.025)
B. AR(1) DISTRIBUTION				
9) AR(1), $\rho = .8$	OLS	50	.44 (.035)	.70 (.032)
10) AR(1), $\rho = .8$	Block bootstrap	50	.05 (.015)	.25 (.031)

Ignoring Time Series Information

IGNORING TIME SERIES DATA

Data	Technique	N	Rejection rate: No effect	Rejection rate: 2% effect
A. CPS DATA				
1) CPS agg	OLS	50	.49 (.025)	.663 (.024)
2) CPS agg	Simple aggregation	50	.053 (.011)	.163 (.018)
3) CPS agg	Residual aggregation	50	.058 (.011)	.173 (.019)
4) CPS agg, staggered laws	Residual aggregation	50	.048 (.011)	.363 (.024)
5) CPS agg	OLS	20	.39 (.025)	.54 (.025)
6) CPS agg	Simple aggregation	20	.050 (.011)	.088 (.014)
7) CPS agg	Residual aggregation	20	.06 (.011)	.183 (.019)
8) CPS agg, staggered laws	Residual aggregation	20	.048 (.011)	.130 (.017)
9) CPS agg	OLS	10	.443 (.025)	.51 (.025)
10) CPS agg	Simple aggregation	10	.053 (.011)	.065 (.012)
11) CPS agg	Residual aggregation	10	.093 (.014)	.178 (.019)
12) CPS agg, staggered laws	Residual aggregation	10	.088 (.014)	.128 (.017)
13) CPS agg	OLS	6	.383 (.024)	.433 (.024)
14) CPS agg	Simple aggregation	6	.068 (.013)	.07 (.013)
15) CPS agg	Residual aggregation	6	.11 (.016)	.123 (.016)
16) CPS agg, staggered laws	Residual aggregation	6	.09 (.014)	.138 (.017)
B. AR(1) DISTRIBUTION				
17) AR(1), $\rho = .8$	Simple aggregation	50	.050 (.013)	.243 (.025)
18) AR(1), $\rho = .8$	Residual aggregation	50	.045 (.012)	.235 (.024)
19) AR(1), $\rho = .8$, staggered laws	Residual aggregation	50	.075 (.015)	.355 (.028)

Staggered DID

Basic Setup

- The canonical difference-in-differences (DID) model contains two time periods, “pre” and “post”, and two groups, “treatment” and “control”.
- Most DID applications, however, exploit variation across groups of units that receive treatment at different times.
- Staggered DID adopts a two-way fixed effects specification:

$$y_{it} = \alpha_i + \lambda_t + \beta D_{it} + u_{it}$$

Use of DID in Finance and Accounting: 2000-2019

	(1) DiD	(2) Staggered DiD	(3) Staggered DiD / DiD (%)
Journal of Finance	54	29	53.70%
Journal of Financial Economics	162	79	48.77%
Review of Financial Studies	139	66	47.48%
Review of Finance	28	12	42.86%
Journal of Financial and Quantitative Analysis	56	32	57.14%
Finance	**439**	**218**	**49.66%**
Journal of Accounting Research	52	21	40.38%
Journal of Accounting and Economics	63	34	53.97%
The Accounting Review	108	52	48.15%
Review of Accounting Studies	46	24	52.17%
Contemporary Accounting Research	43	17	39.53%
Accounting	**312**	**148**	**47.44%**
Finance and Accounting	**751**	**366**	**48.74%**

Recent Advances

- Goodman-Bacon,A.2021,Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*
- Baker, Andrew C., David F. Larcker, and Charles C.Y. Wang. "How Much Should We Trust Staggered Difference-In-Differences Estimates?" 2021,Working Paper

Main Conclusions

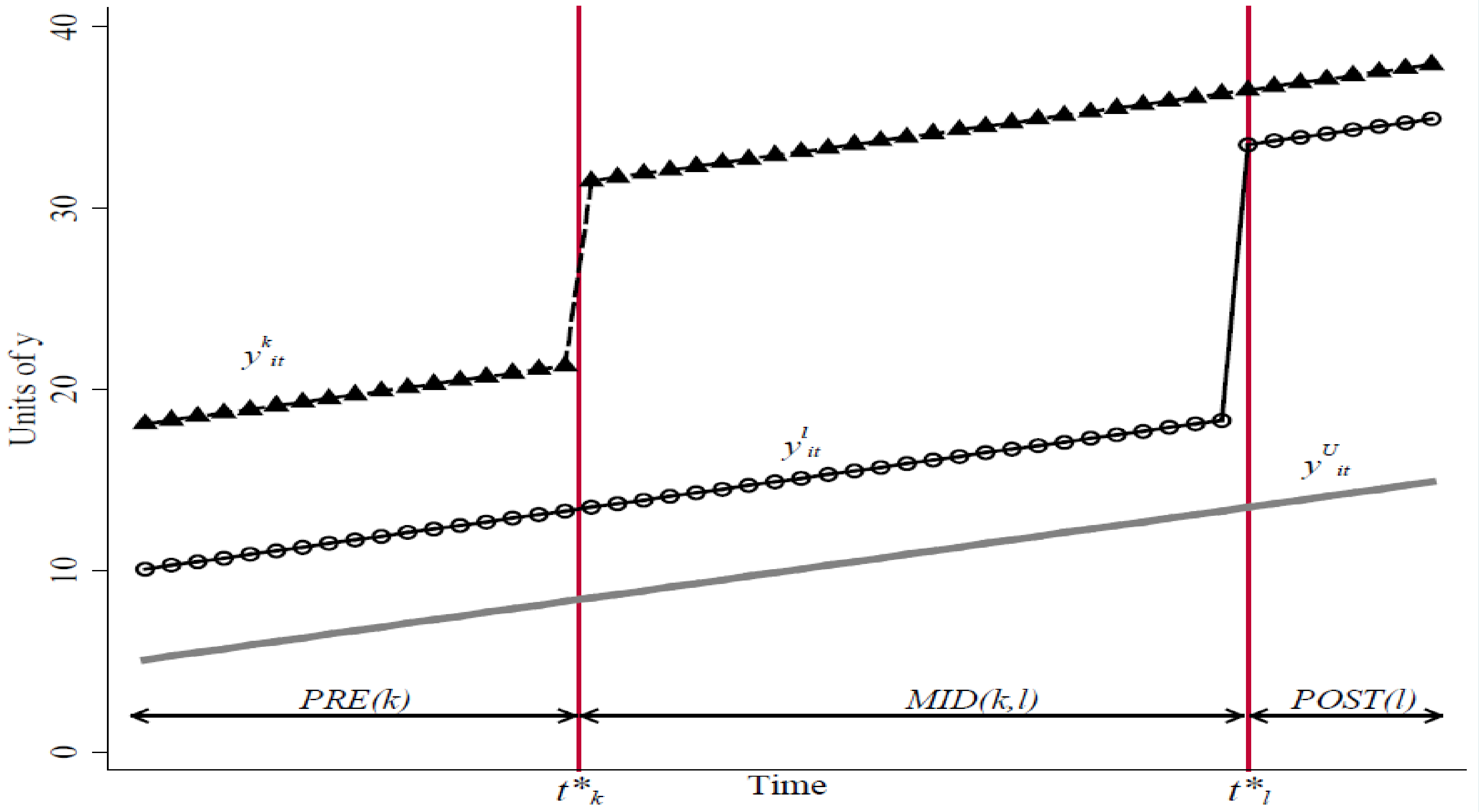
- Recent advances in econometric theory show that such designs are likely to be biased in the presence of treatment effect heterogeneity

- Goodman-Bacon(2021) derives an expression for this general DID estimator according to the ***DD Decomposition Theorem***, and shows that it is a weighted average of all possible two-group/ two-period DID estimators in the data

- Baker et al.(2021) apply recently proposed methods to a set of prior published results and find that the reported effects in prior research become indistinguishable from zero in many cases

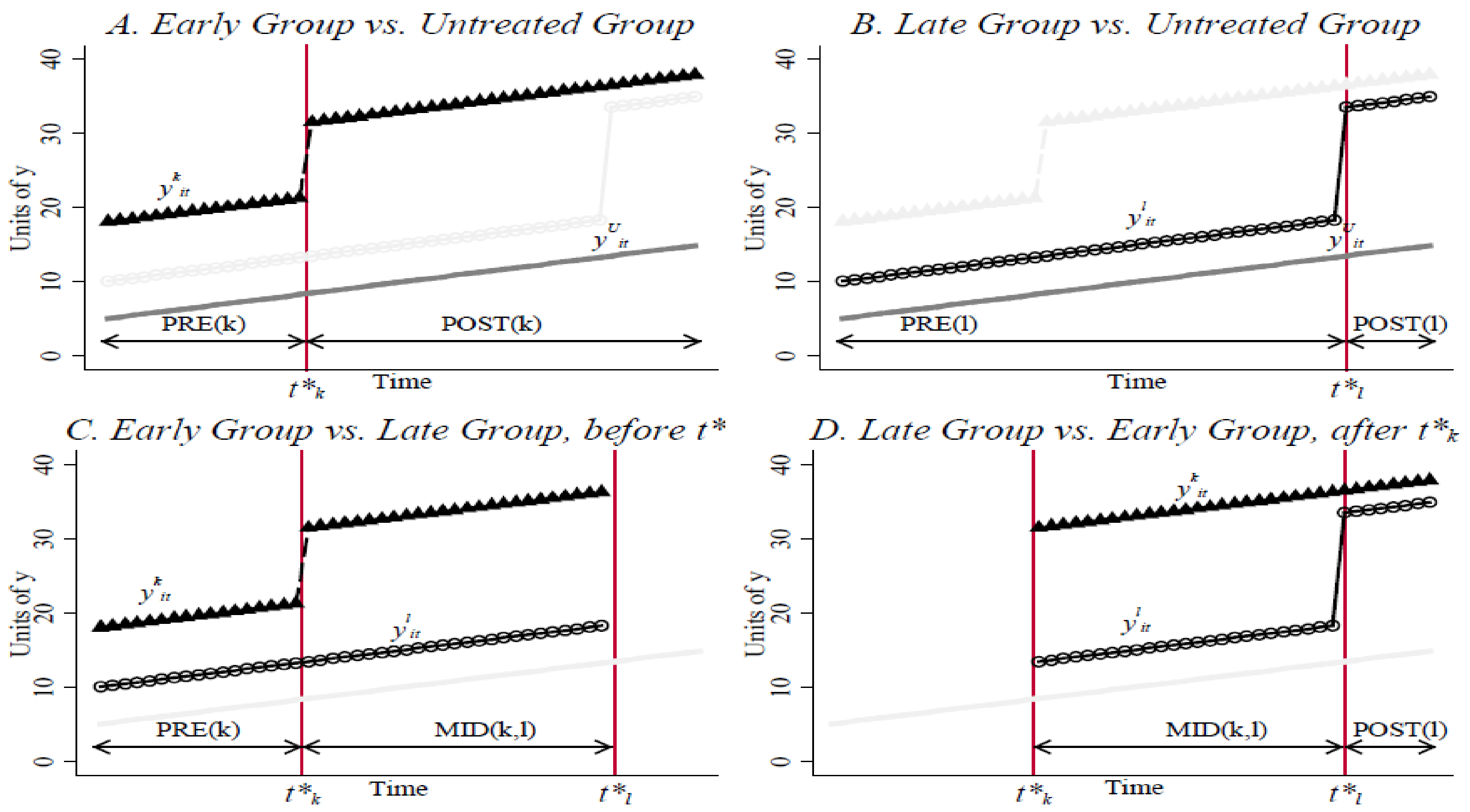
Three Groups

Figure 1. Difference-in-Differences with Variation in Treatment Timing: Three Groups



Decomposition: Graph

Figure 2. The Four Simple (2x2) Difference-in-Differences Estimates from the Three Group Case



Decomposition

$$\widehat{\beta}^{DD} = \sum_{k \neq U} s_{kU} \widehat{\beta}^{2x2}_{kU} + \sum_{k \neq U} \sum_{\ell > k} s_{k\ell} \left[\mu_{k\ell} \widehat{\beta}^{2x2,k}_{k\ell} + (1 - \mu_{k\ell}) \widehat{\beta}^{2x2,\ell}_{k\ell} \right]$$

$$\widehat{\beta}^{2x2}_{kU} \equiv \left(\overline{y}_k^{POST(k)} - \overline{y}_k^{PRE(k)} \right) - \left(\overline{y}_U^{POST(j)} - \overline{y}_U^{PRE(j)} \right)$$

$$\widehat{\beta}^{2x2,k}_{k\ell} \equiv \left(\overline{y}_k^{MID(k,\ell)} - \overline{y}_k^{PRE(k)} \right) - \left(\overline{y}_\ell^{MID(k,\ell)} - \overline{y}_\ell^{PRE(k)} \right)$$

$$\widehat{\beta}^{2x2,\ell}_{k\ell} \equiv \left(\overline{y}_\ell^{POST(\ell)} - \overline{y}_\ell^{MID(k,\ell)} \right) - \left(\overline{y}_k^{POST(\ell)} - \overline{y}_k^{MID(k,\ell)} \right)$$

Weights

$$s_{kU} = \frac{(n_k + n_U)^2 \overbrace{n_{kU}(1 - n_{kU})\overline{D}_k(1 - \overline{D}_k)}^{\hat{V}^D_{kU}}}{\hat{V}^D},$$

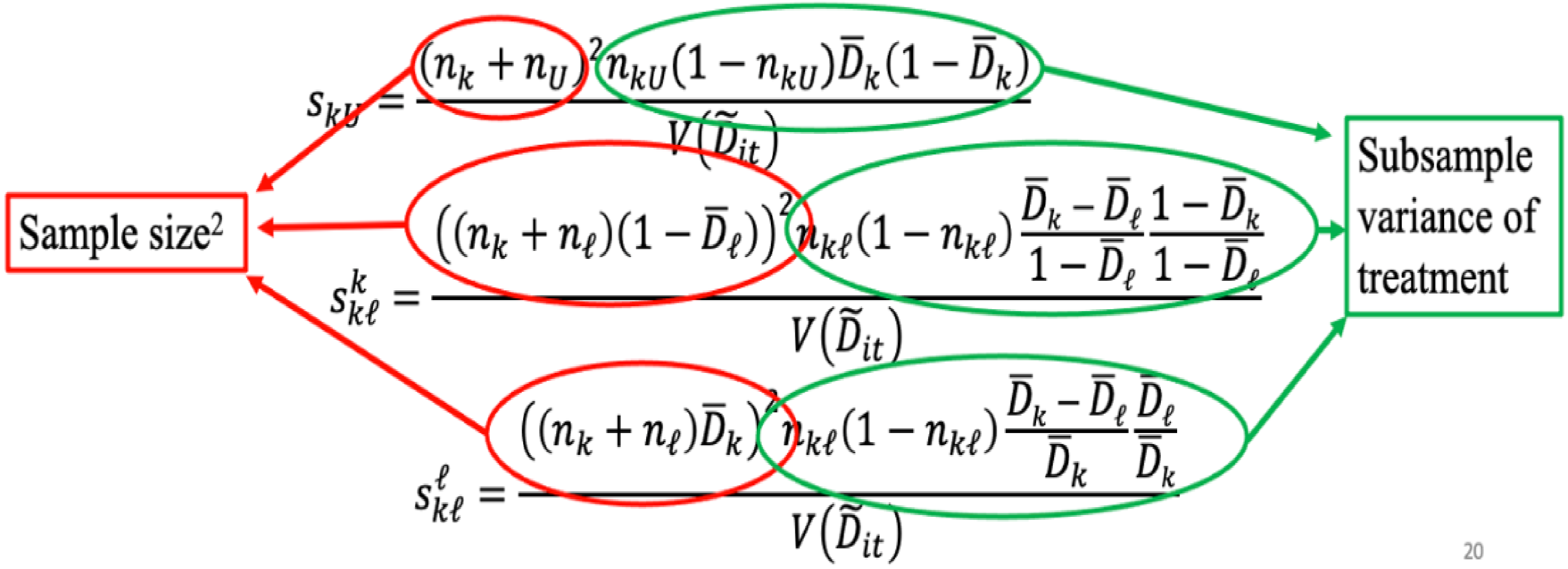
$$s^k_{k\ell} = \frac{\left((n_k + n_\ell)\left(1 - \overline{D}_\ell\right)\right)^2 \overbrace{n_{k\ell}\left(1 - n_{k\ell}\right) \frac{\overline{D}_k - \overline{D}_\ell}{1 - \overline{D}_\ell} \frac{1 - \overline{D}_k}{1 - \overline{D}_\ell}}^{\hat{V}^{D,k}_{k\ell}}}{\hat{V}^D},$$

$$s^\ell_{k\ell} = \frac{\left((n_k + n_\ell)\,\overline{D}_k\right)^2 \overbrace{n_{k\ell}\left(1 - n_{k\ell}\right) \frac{\overline{D}_\ell}{\overline{D}_k} \frac{\overline{D}_k - \overline{D}_\ell}{\overline{D}_k}}^{\hat{V}^{D,\ell}_{k\ell}}}{\hat{V}^D}.$$

$$\textit{and}\ \sum_{k \neq U} s_{kU} + \sum_{k \neq U} \sum_{\ell > k} \left[s^k_{k\ell} + s^\ell_{k\ell}\right] = 1.$$

Weights (cont’d)

- The weights come both from group sizes and the treatment variance in each pair



$$s_{kU} = \frac{(n_k + n_U)^2 n_{kU}(1 - n_{kU})\bar{D}_k(1 - \bar{D}_k)}{V(\tilde{D}_{it})}$$

$$s_{k\ell}^{k} = \frac{\left((n_k + n_\ell)(1 - \bar{D}_\ell)\right)^2 n_{k\ell}(1 - n_{k\ell}) \frac{\bar{D}_k - \bar{D}_\ell}{1 - \bar{D}_\ell} \frac{1 - \bar{D}_k}{1 - \bar{D}_\ell}}{V(\tilde{D}_{it})}$$

$$s_{k\ell}^{\ell} = \frac{\left((n_k + n_\ell)\bar{D}_k\right)^2 n_{k\ell}(1 - n_{k\ell}) \frac{\bar{D}_k - \bar{D}_\ell}{\bar{D}_k} \frac{\bar{D}_\ell}{\bar{D}_k}}{V(\tilde{D}_{it})}$$

Is the estimated parameter ATT?

$$ATT_k(W) \equiv E[Y_{it}^1 - Y_{it}^0 | k, t \in W]$$

$$\Delta Y_k^h(W_1, W_0) \equiv E[Y_{it}^h | k, W_1] - E[Y_{it}^h | k, W_0], \quad h = 0,1$$

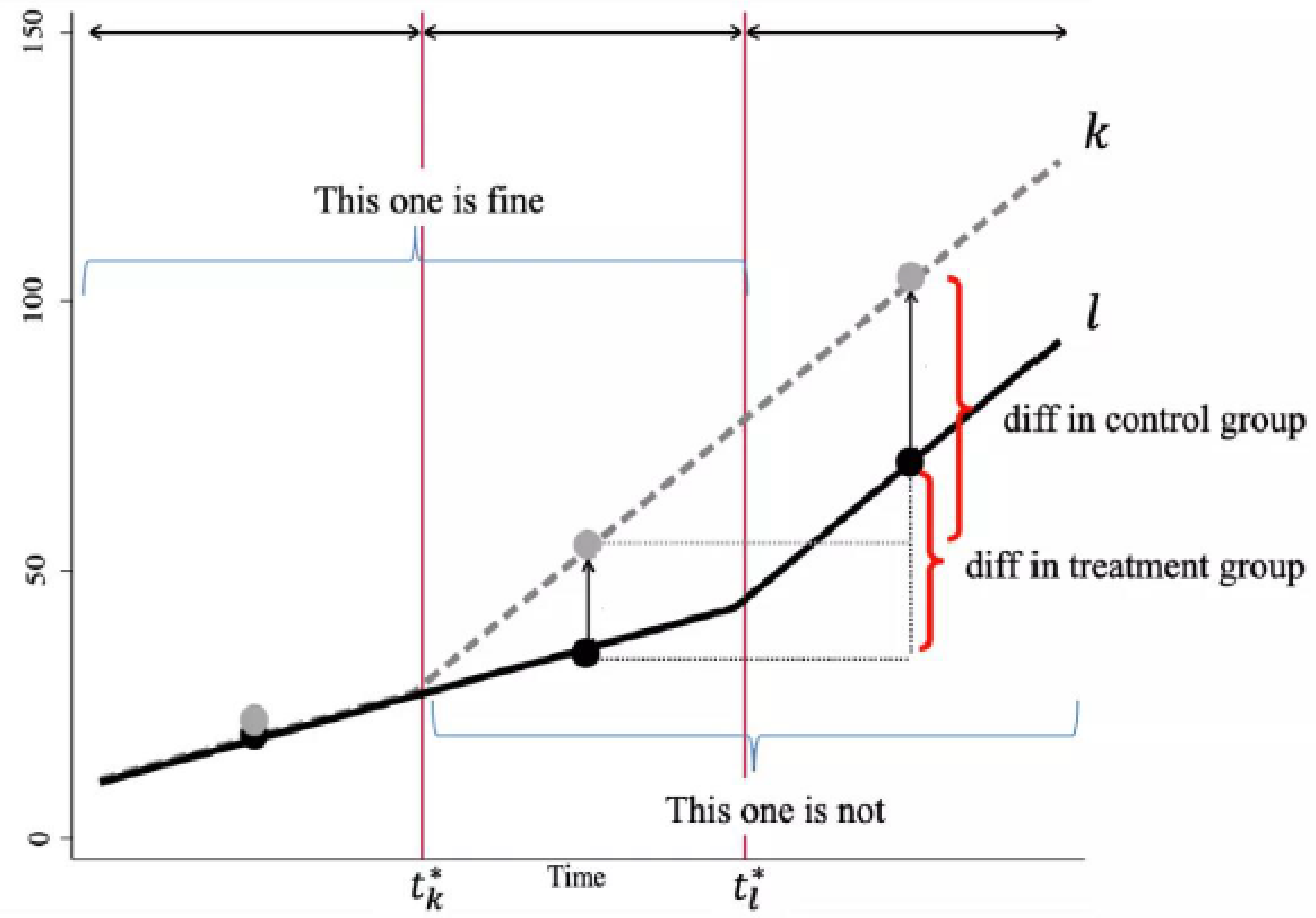
$$\beta_{kU}^{2x2} = ATT_k(POST(k)) + \Delta Y_k^0(POST(k), PRE(k)) - \Delta Y_U^0(POST(k), PRE(k)) \quad (11a)$$

$$\beta_{k\ell}^{2x2,k} = ATT_k(MID(k,\ell)) + \Delta Y_k^0(MID(k,\ell), PRE(k)) - \Delta Y_\ell^0(MID(k,\ell), PRE(k)) \quad (11b)$$

$$\beta_{k\ell}^{2x2,\ell} = ATT_\ell(POST(\ell)) + \Delta Y_\ell^0(POST(\ell), MID(k,\ell)) - \Delta Y_k^0(POST(\ell), MID(k,\ell))$$

$$- [ATT_k(POST(\ell)) - ATT_k(MID(k,\ell))] \quad (11c)$$

The Bias



150
100
50
0
This one is fine
k
l
diff in control group
diff in treatment group
This one is not
t_k^*
Time
t_l^*

The Meaning of Beta

$$plim\ \hat{\beta}^{DD} = VWATT + VWCT - \Delta ATT$$

(1) Variance-weighted ATT

$$VWATT \equiv \sum_{k \neq U} \sigma_{kU}\, ATT_k(POST(k)) + \sum_{k \neq U} \sum_{\ell > k} [\sigma_{k\ell}^{k} ATT_k(MID(k,\ell)) + \sigma_{k\ell}^{\ell} ATT_l(POST(k))]$$

(2) Variance-weighted common trends

$$VWCT \equiv \sum_{k \neq U} \sigma_{kU} [\Delta Y_k^0(POST(k), PRE(k)) - \Delta Y_U^0(POST(k), PRE(k))]$$
$$+ \sum_{k \neq U} \sum_{\ell > k} [\sigma_{k\ell}^{k} \{\Delta Y_k^0(MID(k,l), PRE(k)) - \Delta Y_l^0(MID(k,l), PRE(k))\}$$
$$+ \sigma_{k\ell}^{l} \{\Delta Y_l^0(POST(l), MID(k,l)) - \Delta Y_k^0(POST(l), MID(k,l))\}] \approx \sum_{k} \Delta Y_k^0 [w_k^T - w_k^C]$$

(3) Weighted sum of the change in treatment effects

$$\Delta ATT \equiv \sum_{k \neq U} \sum_{\ell > k} \sigma_{k\ell}^{\ell} [ATT_k(POST(\ell)) - ATT_k(MID(k,\ell))]$$

总结

- 多时点DID的估计本质上是多个传统2×2 DID估计的加权平均，其权重是子样本规模、处理组与控制组相对规模、以及子样本方差的函数：子样本规模越大，相互比较的两个组子样本规模越近，处于中期（如两次政策实施时点之间）的处理组会被赋予更高的权重
- 已经被处理的观测还能作为“控制组”，即便他们本身并不是控制组
- 多时点DID方法极易产生研究偏误，尤其是在政策渐进实施过程中处理效应变化的情况下（Heterogeneous treatment effect）
- 只有满足平行趋势及零时变处理效应，系数才能被解读为平均处理效应（政策带来的因果效应）！！！
- Stata 命令：help bacondecomp

No-fault divorce reforms and female suicide

Table 1. The No-Fault Divorce Rollout: Treatment Times, Group Sizes, and Treatment Shares

No-Fault Divorce Year (t_k^*)	Number of States	Share of States (n_k)	Treatment Share ($\overline{D}_k$)
Non-Reform States	5	0.10	.
Pre-1964 Reform States	8	0.16	.
1969	2	0.04	0.85
1970	2	0.04	0.82
1971	7	0.14	0.79
1972	3	0.06	0.76
1973	10	0.20	0.73
1974	3	0.06	0.70
1975	2	0.04	0.67
1976	1	0.02	0.64
1977	3	0.06	0.61
1980	1	0.02	0.52
1984	1	0.02	0.39
1985	1	0.02	0.36

Notes: The table lists the dates of no-fault divorce reforms from Stevenson and Wolfers (2006), the number and share of states that adopt in each year, and the share of periods each treatment timing group spends treated in the estimation sample from 1964-1996.

bacondecomp

```
. ssc install bacondecomp
checking bacondecomp consistency and verifying not already installed...
installing into c:\ado\plus\...
installation complete.

. help bacondecomp

.  use http://pped.org/bacon_example.dta

. help bacondecomp

. xtreg asmrs post pcinc asmrh cases i.year, fe robust

Fixed-effects (within) regression               Number of obs     =      1,617
Group variable: stfips                          Number of groups  =         49

R-sq:                                           Obs per group:
     within  = 0.3655                                         min =         33
     between = 0.0032                                         avg =       33.0
     overall = 0.1375                                         max =         33

                                                F(36,48)          =      19.40
corr(u_i, Xb)  = -0.1421                        Prob > F          =     0.0000

                                 (Std. Err. adjusted for 49 clusters in stfips)
------------------------------------------------------------------------------
             |               Robust
       asmrs |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        post |  -2.515964   2.283101    -1.10   0.276    -7.106446    2.074519
       pcinc |  -.0011182   .0003614    -3.09   0.003    -.0018448   -.0003916
       asmrh |   1.184598   .5765446     2.05   0.045     .0253772    2.343819
       cases |  -178.5179   137.9739    -1.29   0.202     -455.933    98.89717
```

Variables

Filter variables here

Name
stfips
year
_nfd
post
asmrs
pcinc
asmrh
cases
weight
copop

Properties

Variables

Name	
Label	
Type	
Format	
Value label	
Notes	

Data

bacondecomp (cont'd)

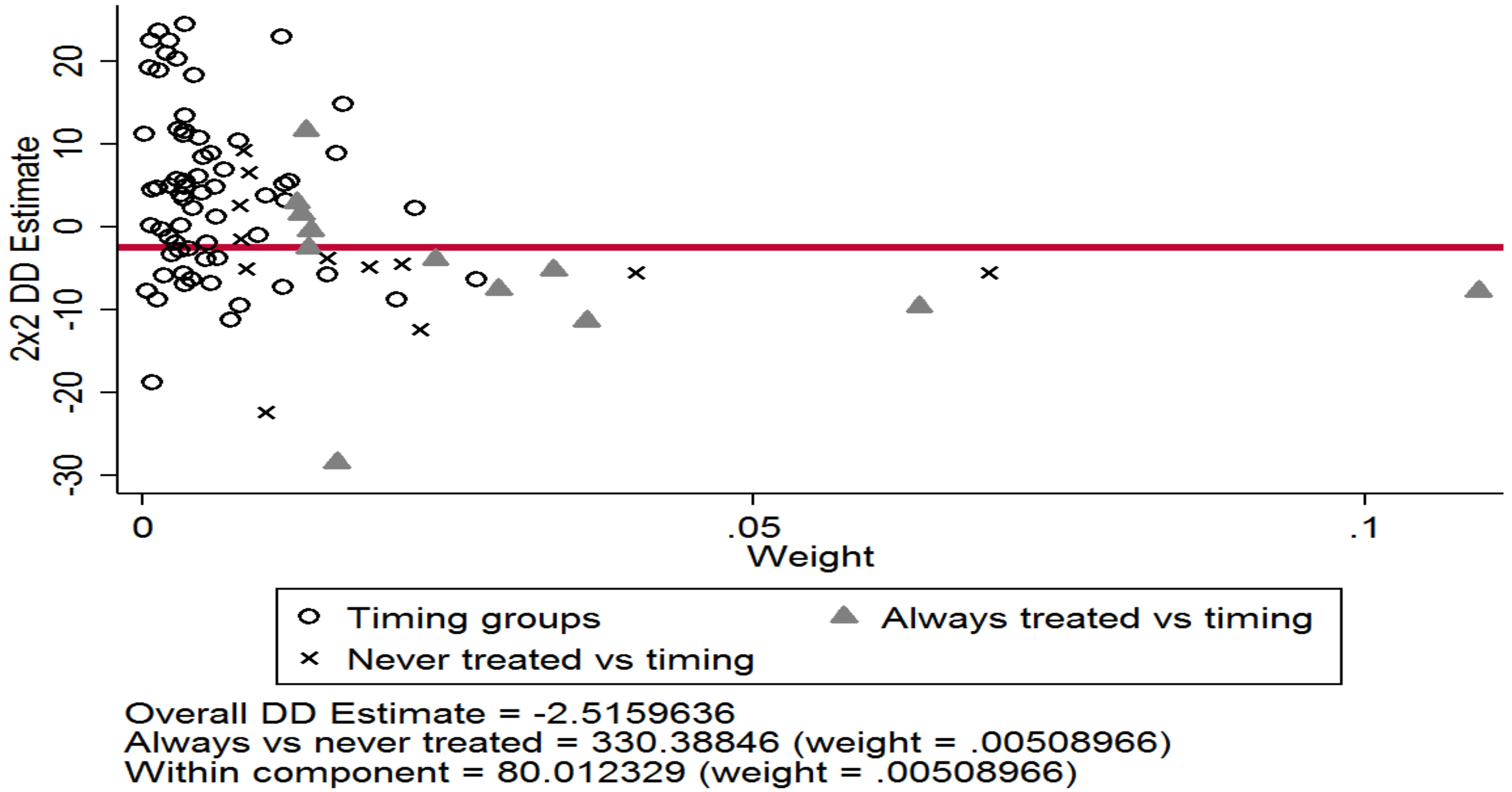
```
. bacondecomp asmrs post pcinc asmrh cases, stub(Bacon_) robust
Computing decomposition across 14 timing groups
including an always-treated group and a never-treated group
```

asmrs	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
post	-2.515964	2.283101	-1.10	0.270	-6.99076	1.958833

Bacon Decomposition

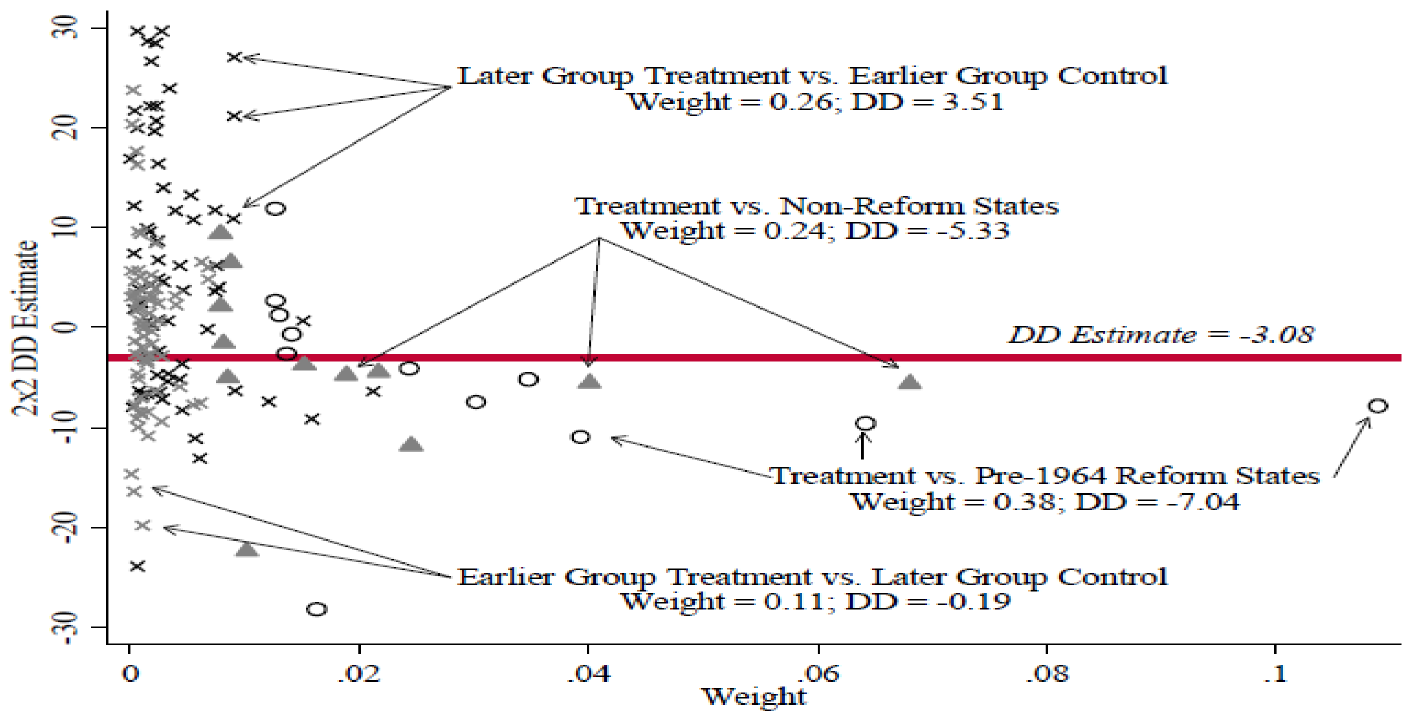
	Beta	TotalWeight
Timing_groups	2.602167327	.3776606651
Always_v_timing	-7.02043576	.3783086972
Never_v_timing	-5.256988806	.2389229013
Always_v_never	330.3884583	.0000180736
Within	80.0123291	.0050896627

bacondecomp (cont'd)



DID decomposition

Figure 6. Difference-in-Differences Decomposition for Unilateral Divorce and Female Suicide



Notes: Notes: The figure plots each 2x2 DD components from the decomposition theorem against their weight for the unilateral divorce analysis. The open circles are terms in which one timing group acts as the treatment group and the pre-1964 reform states act as the control group. The closed triangles are terms in which one timing group acts as the treatment group and the non-reform states act as the control group. The x's are the timing-only terms. The figure notes the average DD estimate and total weight on each type of comparison. The two-way fixed effects estimate, -3.08, equals the average of the y-axis values weighted by their x-axis value.

Thanks!