Week 13: More on DID

JJ Chen

April 10, 2015

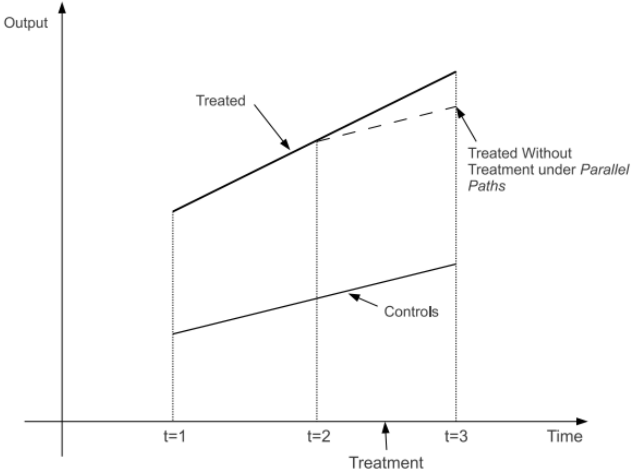
Today's Plan

- Parallel trend and parallel growth assumptions
  - Mora and Reggio (2014)

Motivation

- The ID assumption for DID design is the common trend
  - Parallel path for outcome levels in post-treatment periods
- In many DID papers, authors include group-specific linear time trends and call it
  - a robustness check, or
  - a relaxation of the parallel trend assumption
- But once we include group-specific time trends, is the ID assumption still the common trend?
- Today we try to be more specific about the ID assumption under group-specific time trends

Motivation: Why Group-Specific Trends



Review: Common Trend Assumption

- Two groups of units: $\text{T}_i = 0$, $\text{T}_i = 1$
- Two periods: $\text{P}_t = 0$, $\text{P}_t = 1$
- Treatment group receives treatment in post period: $\text{TP}_{it} = \text{T}_i \times \text{P}_t$
- Two potential outcomes: Y_{0it}, Y_{1it}
  - Observed outcomes: $\text{Y}_{it} = \text{Y}_{0it} \times (1 - \text{TP}_{it}) + \text{Y}_{1it} \times \text{TP}_{it}$
- ATET: $\delta^{ATET} = \text{E}(\text{Y}_{1it} - \text{Y}_{0it} \mid \text{TP}_{it} = 1)$
  - $\implies \delta^{ATET} = \text{E}(\text{Y}_{1i1} - \text{Y}_{0i1} \mid \text{T} = 1)$

Regressions Gives Conditional Mean Comparison

- $\text{Y}_{it} = \alpha + \beta\text{T}_i + \gamma\text{P}_t + \delta\text{TP}_{it} + e_{it}$
  - $\text{E}(\text{Y}_{i1} \mid \text{T} = 1) = \alpha + \beta + \gamma + \delta$
  - $\text{E}(\text{Y}_{i0} \mid \text{T} = 1) = \alpha + \beta + \delta$
  - $\text{E}(\text{Y}_{i1} \mid \text{T} = 0) = \alpha + \gamma$
  - $\text{E}(\text{Y}_{i0} \mid \text{T} = 0) = \alpha$
- $\delta^{DID} = \text{E}(\text{Y}_{i1} - \text{Y}_{i0} \mid \text{T} = 1) - \text{E}(\text{Y}_{i1} - \text{Y}_{i0} \mid \text{T} = 0)$

Potential Outcome DID

- ATET: $\delta^{ATET} = \mathrm{E}(\mathrm{Y}_{1i1} - \mathrm{Y}_{0i1} \mid \mathrm{T} = 1)$
- DID gives ATET?

$$
\begin{aligned}
\delta^{DID} &= \mathrm{E}(\mathrm{Y}_{i1} - \mathrm{Y}_{i0} \mid \mathrm{T} = 1) - \mathrm{E}(\mathrm{Y}_{i1} - \mathrm{Y}_{i0} \mid \mathrm{T} = 0) \\
&= \mathrm{E}(\mathrm{Y}_{1i1} - \mathrm{Y}_{0i0} \mid \mathrm{T} = 1) - \mathrm{E}(\mathrm{Y}_{0i1} - \mathrm{Y}_{0i0} \mid \mathrm{T} = 0) \\
&= \mathrm{E}(\mathrm{Y}_{1i1} - \mathrm{Y}_{0i0} \mid \mathrm{T} = 1) - \mathrm{E}(\mathrm{Y}_{0i1} - \mathrm{Y}_{0i0} \mid \mathrm{T} = 1) \\
&= \mathrm{E}(\mathrm{Y}_{1i1} - \mathrm{Y}_{0i1} \mid \mathrm{T} = 1) \\
&= \delta^{ATET}
\end{aligned}
$$

- ID assumption for ATET:
  $\mathrm{E}(\mathrm{Y}_{0i1} - \mathrm{Y}_{0i0} \mid \mathrm{T} = 1) = \mathrm{E}(\mathrm{Y}_{0i1} - \mathrm{Y}_{0i0} \mid \mathrm{T} = 0)$
  - Or: $\mathrm{Y}_{0i1} - \mathrm{Y}_{0i0} \perp\!\!\!\perp \mathrm{T}$
  - Can be extended to selection-on-observable: $\mathrm{Y}_{0i1} - \mathrm{Y}_{0i0} \perp\!\!\!\perp \mathrm{T} \mid X$

Multiple Periods

- Now suppose we have three periods: $t = 1, 2, 3$
  - Assume two pre-pretreatment periods and one post-treatment period
  - $\mathrm{P}_t = 1$ if $t = 3$; $\mathrm{P}_t = 0$ if $t = 1, 2$
- ATET: $\delta^{ATET} = \mathrm{E}(\mathrm{Y}_{1i3} - \mathrm{Y}_{0i3} \,|\mathrm{TP}_{it} = 1)$
- Regression that allows group-specific linear time trend
  - $\mathrm{Y}_{it} = \alpha + \beta \mathrm{T}_i + \gamma \mathrm{P}_t + \delta \mathrm{TP}_{it} + \theta(\mathrm{T}_i \times t) + e_{it}$
  - A simple version of eqn (5.2.7) in MHE
- What are the conditional means?

Multiple Periods

- $\mathrm{Y}_{it} = \alpha + \beta \mathrm{T}_i + \gamma \mathrm{P}_t + \delta \mathrm{TP}_{it} + \theta(\mathrm{T}_i \times t) + e_{it}$
  - $\mathrm{E}(\mathrm{Y}_{i3} \,|\mathrm{T} = 1) = \alpha + \beta + \gamma + \delta + 3\theta$
  - $\mathrm{E}(\mathrm{Y}_{i2} \,|\mathrm{T} = 1) = \alpha + \beta + 2\theta$
  - $\mathrm{E}(\mathrm{Y}_{i1} \,|\mathrm{T} = 1) = \alpha + \beta + \theta$
  - $\mathrm{E}(\mathrm{Y}_{i3} \,|\mathrm{T} = 0) = \alpha + \gamma$
  - $\mathrm{E}(\mathrm{Y}_{i2} \,|\mathrm{T} = 0) = \alpha$
  - $\mathrm{E}(\mathrm{Y}_{i1} \,|\mathrm{T} = 0) = \alpha$
-

$$\begin{aligned}\delta^{DID} =& \mathrm{E}\left((\mathrm{Y}_{i3} - \mathrm{Y}_{i2}) - (\mathrm{Y}_{i2} - \mathrm{Y}_{i1}) \,|\mathrm{T} = 1\right) - \\ & \mathrm{E}\left((\mathrm{Y}_{i3} - \mathrm{Y}_{i2}) - (\mathrm{Y}_{i2} - \mathrm{Y}_{i1}) \,|\mathrm{T} = 0\right) \\ =& \mathrm{E}(\Delta \mathrm{Y}_{i3} - \Delta \mathrm{Y}_{i2} \,|\mathrm{T} = 1) - \mathrm{E}(\Delta \mathrm{Y}_{i3} - \Delta \mathrm{Y}_{i2} \,|\mathrm{T} = 0)\end{aligned}$$

ID Assumption: Pallele Growth

- ATET: $\delta^{ATET} = \mathrm{E}(\mathrm{Y}_{1i3} - \mathrm{Y}_{0i3} \mid \mathrm{T}_{it} = 1)$
-

$$
\begin{aligned}
\delta^{DID} &= \mathrm{E}(\Delta \mathrm{Y}_{i3} - \Delta \mathrm{Y}_{i2} \mid \mathrm{T} = 1) - \mathrm{E}(\Delta \mathrm{Y}_{i3} - \Delta \mathrm{Y}_{i2} \mid \mathrm{T} = 0) \\
&= \mathrm{E}(\Delta \mathrm{Y}_{1i3} - \Delta \mathrm{Y}_{0i2} \mid \mathrm{T} = 1) - \mathrm{E}(\Delta \mathrm{Y}_{0i3} - \Delta \mathrm{Y}_{0i2} \mid \mathrm{T} = 0) \\
&= \mathrm{E}(\Delta \mathrm{Y}_{1i3} - \Delta \mathrm{Y}_{0i2} \mid \mathrm{T} = 1) - \mathrm{E}(\Delta \mathrm{Y}_{0i3} - \Delta \mathrm{Y}_{0i2} \mid \mathrm{T} = 1) \\
&= \mathrm{E}(\Delta \mathrm{Y}_{1i3} - \Delta \mathrm{Y}_{0i3} \mid \mathrm{T} = 1) \\
&= \mathrm{E}\left((\mathrm{Y}_{1i3} - \mathrm{Y}_{0i2}) - (\mathrm{Y}_{0i3} - \mathrm{Y}_{0i2}) \mid \mathrm{T} = 1\right) \\
&= \mathrm{E}(\mathrm{Y}_{1i3} - \mathrm{Y}_{0i3} \mid \mathrm{T} = 1) = \delta^{ATET}
\end{aligned}
$$

- ID Assumption:
  $\mathrm{E}(\Delta \mathrm{Y}_{0i3} - \Delta \mathrm{Y}_{0i2} \mid \mathrm{T} = 1) = \mathrm{E}(\Delta \mathrm{Y}_{0i3} - \Delta \mathrm{Y}_{0i2} \mid \mathrm{T} = 0)$
  - Compare to parallel trend
    $\mathrm{E}(\mathrm{Y}_{0i3} - \mathrm{Y}_{0i2} \mid \mathrm{T} = 1) = \mathrm{E}(\mathrm{Y}_{0i3} - \mathrm{Y}_{0i2} \mid \mathrm{T} = 0)$

Parellel Trend and Parellel Growth is Non-nested

Table 1: *Parallel*-(1) *vs.* *Parallel*-(2)

	$Y^0_{t^*-1}$	$Y^0_{t^*}$	$Y^0_{t^*+1}$	$\Delta Y^0_{t^*+1}$	$\Delta\left(\Delta Y^0_{t^*+1}\right)$
Case (a)					
$D=1$	1	2	3	1	0
$D=0$	0	0	1	1	1
Case (b)					
$D=1$	1	2	4	2	1
$D=0$	0	0	1	1	1
Case (c)					
$D=1$	1	2	4	2	1
$D=0$	0	1	3	2	1

Note: Expected outcomes in the absence of treatment conditional on treatment and X. Case (a) illustrates a situation whereby Parallel-(1) is satisfied while Parallel-(2) is violated. Case (b) illustrates the opposite situation. In Case (c), both Parallel assumptions are satisfied.

Parellel Trend and Parellel Growth

- Mora and Reggio (2014) formalize the above ideas and characterize a family of identification assumptions: *Parallel-(q)*
  - Parallel trend:*Parallel-(1)*; Parallel growth: *Parallel-(2)*
- Notations:
  - Let t^* be the last pre-treatment period. Assume the treatment finishes before the first post-treatment period, $t^* + 1$
  - Denote the s-period difference operator as $\Delta_s \equiv (1 - L^s)$, and the $(q-1)$-th difference operator as $\Delta^{q-1} \equiv (1-L)^{q-1}$
- *Parallel-(q)* is defined for a given pre-treatment period $q \leq t^*$, and for a given post-treatment periods $s, 1 \leq s \leq J - t^*$, $\mathrm{E}(\Delta_s \Delta^{q-1} \mathrm{Y}^0_{i,t^*+s} \mid \mathrm{X}_i, \mathrm{D}_i = 1) = \mathrm{E}(\Delta_s \Delta^{q-1} \mathrm{Y}^0_{i,t^*+s} \mid \mathrm{X}_i, \mathrm{D}_i = 0)$
  - The number of pre-treatment periods sets the max. number of assumptions that can be assumed

More on DID Regressions

- Extensions of Standard Models when there are multiple periods:
  - $\mathrm{Y}_{it} = \alpha + \beta \mathrm{T}_i + \sum_{\tau=2}^{J} \gamma_\tau \mathrm{I}_t^\tau + \delta \mathrm{TP}_{it} + e_{it}$
  - $\mathrm{Y}_{it} = \alpha + \beta \mathrm{T}_i + \sum_{\tau=2}^{J} \gamma_\tau \mathrm{I}_t^\tau + \delta \mathrm{TP}_{it} + \theta(\mathrm{TP}_{it} \times t) + e_{it}$
  - $\mathrm{Y}_{it} = \alpha + \beta \mathrm{T}_i + \sum_{\tau=2}^{J} \gamma_\tau \mathrm{I}_t^\tau + \sum_{s=1}^{J-t^*} \delta_s(\mathrm{TP}_{it} \times \mathrm{I}_t^{t^*+s}) + e_{it}$
- For all the three regressions, pre-treatment dynamics are identical for controls and treated
  - In the presence of group-specific trends, these models are inappropriate

More on DID Regressions

- Incorporating group-specific time trends:
  - $$\mathrm{Y}_{it} = \alpha + \beta \mathrm{T}_i + \sum_{\tau=2}^{J} \gamma_\tau \mathrm{I}_t^\tau + \sum_{s=1}^{J-t^*} \delta_s (\mathrm{T}_i \times \mathrm{I}_t^{t^*+s}) + e_{it}$$
  - $$\mathrm{Y}_{it} = \alpha + \beta \mathrm{T}_i + \sum_{\tau=2}^{J} \gamma_\tau \mathrm{I}_t^\tau + \sum_{s=1}^{J-t^*} \delta_s (\mathrm{T}_i \times \mathrm{I}_t^{t^*+s}) + \sum_{r=1}^{R} \delta_r (\mathrm{T}_i \times t^r) + e_{it}$$
- Fully flexible model (multiple periods, two groups):
  - $$\mathrm{Y}_{it} = \alpha + \beta \mathrm{T}_i + \sum_{\tau=2}^{J} \gamma_\tau \mathrm{I}_t^\tau + \sum_{\tau=2}^{J} \delta_\tau (\mathrm{T}_i \times \mathrm{I}_t^\tau) + e_{it}$$

A Brief Review of Current Practive

Table B1: *List of Selected Papers*

Author	Year	Journal	Title	No. Pre	No. Post
Aaronson and Mazumder	2011	JPE	The impact of Rosenwald Schools on Black achievement	2	2
Abramitzky, Delavande, and Vasconcelos	2011	AEJ:AE	Marrying Up: The Role of Sex Ratio in Assortative Matching	6	11
Currie and Walker	2011	AEJ:AE	Traffic Congestion and Infant Health: Evidence from E-ZPass	300	168
De Jong, Lindeboom, and Van der Klaauw	2011	JEEA	Screening disability insurance applications	2	1
Jayachandran, Lleras-Muney, and Smith	2010	AEJ:AE	Modern Medicine and the Twentieth Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs	12	7
Furman and Stern	2011	AER	Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research	14	18
Kotchen and Grant	2011	REStat	Does Daylight Saving Time Save Energy? Evidence from a Natural Experiment in Indiana	2	1
Moser and Voena	2012	AER	Compulsory Licensing: Evidence from the Trading with the Enemy Act	43	22
Redding, Sturm, and Wolf	2011	REStat	History and industry location: Evidence from German airports	12	40

Note: Papers are listed by the alphabetical order obtained from the author's name. The papers selected satisfy the following conditions: (a) There is an application of DID; (b) the sample includes more than one period before treatment; (c) data are publicly available; and (d) the paper is published in the period 2009:2012 in one of the following 10 Economics journals: AEJ:AE, AER, JAppEcon, JEcon, JEEA, JLabEc, JPE, QJE, REStat, and REStud. "No. Pre." refers to the number of pre-treatment periods and "No. Post." refers to the number of post-treatment periods.

A Brief Review of Current Practive

Table B2: *Fully flexible model results and reported results from selected papers*

	Reported	Fully Flexible Model						
	Estimated	$q=1$		$q=2$		Equiv.	Common	Linear
Article	Effect	Effect	Dynamics	Effect	Dynamics	Test	Trends	Trend
Aaronson and Mazumder	0.072***	0.039***	3.337	0.053***	6.488	1.420		
(2011)	(0.007)	(0.012)	[0.068]	(0.017)	[0.011]	[0.234]		
Abramitzky, Delavande,	-0.020**	0.036	22.651	0.106	23.428	-0.069	13.41	13.287
and Vasconcelos (2011) - 1	(0.010)	(0.039)	[0.012]	(0.073)	[0.009]	[0.118]	[0.020]	[0.010]
Abramitzky, Delavande,	-0.010***	0.008	15.983	0.010	16.205	-0.003	4.339	2.633
and Vasconcelos (2011) - 2	(0.004)	(0.016)	[0.100]	(0.030)	[0.094]	[0.870]	[0.502]	[0.621]
Abramitzky, Delavande,	-0.017***	0.003	26.989	0.031	28.664	-0.028	11.27	10.76
and Vasconcelos (2011) - 3	(0.005)	(0.013)	[0.003]	(0.022)	[0.001]	[0.042]	[0.046]	[0.029]
Currie and Walker (2011) - 1	-0.208***	-0.506***	13.748	-0.386	13.796	-0.121	652.85	323.79
	(0.028)	(0.198)	[0.132]	(0.395)	[0.131]	[0.600]	[0.000]	[0.000]
Currie and Walker (2011) - 2	-0.090***	-0.582***	33.123	-1.071***	30.811	0.489	173.23	172.45
	(0.024)	(0.198)	[0.000]	(0.353)	[0.000]	[0.013]	[0.000]	[0.000]
Currie and Walker (2011) - 3	-0.065***	0.029	13.304	0.136	15.950	-0.107	351.47	282.06
	(0.017)	(0.101)	[0.149]	(0.128)	[0.068]	[0.079]	[0.000]	[0.000]
Currie and Walker (2011) - 4	-0.181***	-0.191*	25.404	-0.380*	27.992	0.189	581.98	316.51
	(0.023)	(0.108)	[0.003]	(0.204)	[0.001]	[0.100]	[0.000]	[0.000]
Currie and Walker (2011) - 5	0.018	-0.421	20.420	-0.592	14.565	0.171	268.15	246.61
	(0.038)	(0.374)	[0.016]	(0.736)	[0.104]	[0.714]	[0.000]	[0.000]
Furman and Stern (2011)	0.535***	0.471***	1.605	0.666	1.562	0.262	69.26	69.86
	(0.142)	(0.123)	[0.071]	(0.417)	[0.083]	[0.610]	[0.000]	[0.000]
Kotchen and Grant (2011)-1	0.009***	0.006*		-0.002		7.28		
	(0.003)	(0.003)		(0.005)		[0.007]		
Kotchen and Grant (2011)-2	-0.003	-0.006**		-0.013***		3.97		
	(0.003)	(0.003)		(0.005)		[0.0471]		
Moser and Voena (2012)	0.151***	0.075	4.606	0.006	3.995	2.362	6.84	2.89
	(0.036)	(0.046)	[0.000]	(0.081)	[0.000]	[0.124]	[0.000]	[0.000]

References I

Mora, Ricardo, and Iliana Reggio. 2014. "Treatment Effect Identification Using Alternative Parallel Assumptions."