Frontiers in Difference-in-Differences

Relaxing Parallel Trends

Brantly Callaway

brantly.callaway@uga.edu

University of Georgia

October 17, 2023

Covariates in the Parallel Trends Assumption

Including covariates in the parallel trends assumption can often make DID identification strategies more plausible:

Example

- Minimum wage example: path of teen employment may depend on a state's population / population growth / region of the country
- Job displacement example: path of earnings may depend on year's of education / race / occupation

However, there are a number of new issues that can arise in this setting…

Outline

1. Limitations of TWFE Regression
2. Identification with Two Periods
3. Alternative Estimation Strategies
4. Multiple Periods
5. Minimum Wage Application
6. Dealing with “Bad” Controls

Notation

Start with the case with only two time periods

Only need a little bit of new notation here:

- X_{t^*} and X_{t^*-1} – time-varying covariates
- Z – time-invariant covariates

Covariates in the Parallel Trends Assumption

Conditional Parallel Trends Assumption

$$\mathbb{E}[\Delta Y_{t^*}(0)|X_{t^*}, X_{t^*-1}, Z, D = 1] = \mathbb{E}[\Delta Y_{t^*}(0)|X_{t^*}, X_{t^*-1}, Z, D = 0]$$

In words: Parallel trends holds conditional on having the same covariates (X_t, X_{t-1}, Z).

Minimum wage example (e.g.) Parallel trends conditional on counties have the same population (like X_t) and being in the same region of the country (like Z)

Limitations of TWFE Regressions

Limitations of TWFE Regressions

In this setting, it is common to run the following TWFE regression:

$Y_{it} = \theta_t + \eta_i + \alpha D_{it} + X_{it}'\beta + e_{it}$

However, there are a number of issues:

Issue 1: Issues related to multiple periods and variation in treatment timing still arise

Issue 2: Hard to allow parallel trends to depend on time-invariant covariates

Issue 3: Hard to allow for covariates that could be affected by the treatment

Limitations of TWFE Regressions

In this setting, it is common to run the following TWFE regression:

$$Y_{it} = \theta_t + \eta_i + \alpha D_{it} + X_{it}'\beta + e_{it}$$

However, there are a number of issues:

Issue 4: Linearity results in mixing identification and estimation…e.g., with 2 periods

$$\Delta Y_{it} = \Delta\theta_t + \alpha D_{it} + \Delta X_{it}'\beta + \Delta e_{it}$$

$\implies$ differencing out unit fixed effects can have implications about what researcher controls for

- This doesn't matter if model for untreated potential outcomes is truly linear
- However, if we think of linear model as an approximation, this may have meaningful implications.

Limitations of TWFE Regressions

Even if none of the previous 4 issues apply, α will still be equal to a weighted average of underlying (conditional-on-covariates) treatment effect parameters.

- The weights can be negative, and suffer from "weight reversal" (similar to the issue discussed in Sloczynski (2020) in the context of cross-sectional regressions with covaraites)
- In other words, weights α is a weighted average of $ATT(X)$ where (relative to a baseline of weighting based on the distribution of X for the treated group), the weights put larger weight on $ATT(X)$ for values of the covariates that are *uncommon* for the treated group relative to the untreated group and smaller weight on $ATT(X)$ for values of the covariates that are *common* for the treated group relative to the untreated group

See Caetano and Callaway (2023) for more details

Identification under Conditional Parallel Trends

Under conditional parallel trends, we have that

$\text{ATT} = \mathbb{E}[\Delta Y_{t^*}|D = 1] - \mathbb{E}[\Delta Y_{t^*}(0)|D = 1]$

Identification under Conditional Parallel Trends

Under conditional parallel trends, we have that

$$
\begin{aligned}
ATT &= \mathbb{E}[\Delta Y_{t^*} | D = 1] - \mathbb{E}[\Delta Y_{t^*}(0) | D = 1] \\
&= \mathbb{E}[\Delta Y_{t^*} | D = 1] - \mathbb{E}\Big[\mathbb{E}[\Delta Y_{t^*}(0) | X, D = 1] \Big| D = 1\Big]
\end{aligned}
$$

Identification under Conditional Parallel Trends

Under conditional parallel trends, we have that

$$\begin{aligned} ATT &= \mathbb{E}[\Delta Y_{t^*}|D=1] - \mathbb{E}[\Delta Y_{t^*}(0)|D=1] \\ &= \mathbb{E}[\Delta Y_{t^*}|D=1] - \mathbb{E}\Big[\mathbb{E}[\Delta Y_{t^*}(0)|X, D=1]\Big|D=1\Big] \\ &= \mathbb{E}[\Delta Y_{t^*}|D=1] - \mathbb{E}\Big[\underbrace{\mathbb{E}[\Delta Y_{t^*}(0)|X, D=0]}_{=:m_0(X)}\Big|D=1\Big] \end{aligned}$$

Intuition: (i) Compare path of outcomes for treated group to (conditional on covariates) path of outcomes for untreated group, (ii) adjust for differences in the distribution of covariates between groups.

This argument also require an overlap condition

- For all possible values of the covariates, $p(x) := P(D=1|X=x) < 1$.
- In words: for all treated units, we can find untreated units that have the same characteristics

There are several possible ways to turn this identification result into an estimation strategy →

Regression Adjustment

The challenging term to deal with in the previous expression for ATT is

$$\mathbb{E}\left[\underbrace{\mathbb{E}[\Delta Y_{t^*}(0)|X, D=0]}_{=:m_0(X)}\middle|D=1\right]$$

The most direct way to proceed is by proposing a model for $m_0(X)$.

This expression suggests a regression adjustment estimator. For example, if we assume that $m_0(X) = X'\text{beta}_0$, then we have that

$$ATT = \mathbb{E}[\Delta Y_{t^*}|D=1] - \mathbb{E}[X'|D=1]\beta_0$$

and we can estimate the ATT by

- Step 1: Estimate β_0 using untreated group
- Step 2: Combine this estimate with estimates of $\mathbb{E}[\Delta Y_t|D=1]$ and $\mathbb{E}[X'|D=1]$ (using treated group) to estimate ATT

Covariate Balancing

Alternatively, if we could choose balancing weights $\nu_0(X)$ such that the distribution of X was the same in the untreated group as it is in the treated group after applying the balancing weights, then we would have that (from the second term above)

$$\begin{aligned}\mathbb{E}\Big[\mathbb{E}[\Delta Y_{t^*}(0)|X, D=0]\Big|D=1\Big] &= \mathbb{E}\Big[\nu_0(X)\mathbb{E}[\Delta Y_{t^*}(0)|X, D=0]\Big|D=0\Big] \\ &= \mathbb{E}[\nu_0(X)\Delta Y_{t^*}(0)|D=0]\end{aligned}$$

where the first equality is due to balancing weights and the second by the law of iterated expectations.

The most common way to re-weight is based on the propensity score, you can show:

$$\nu_0(x) = \frac{p(x)(1-p)}{(1-p(x))p}$$

where $p(x) = P(D=1|X=x)$ and $p = P(D=1)$.

This is the approach suggested in Abadie (2005). In practice, you need to estimate the propensity score. The most common choices are probit or logit.

Doubly Robust

Alternatively, you can show

$$ATT = \mathbb{E}\left[\left(\frac{D}{p} - \frac{p(X)(1-D)}{(1-p(X))p}\right)(\Delta Y_t - \mathbb{E}[\Delta Y_t | X, D=0])\right]$$

This requires estimating both $p(X)$ and $\mathbb{E}[\Delta Y_{t^*} | X, D=0]$.

Big advantage:

- This expression for ATT is *doubly robust*. This means that, it will deliver consistent estimates of ATT if either the model for $p(X)$ or for $\mathbb{E}[\Delta Y_{t^*} | X, D=0]$.
- In my experience, doubly robust estimators perform much better than either the regression or propensity score weighting estimators
- This also provides a connection to estimating ATT under conditional parallel trends using machine learning for $p(X)$ and $\mathbb{E}[\Delta Y_{t^*} | X, D=0]$ (see: Chang (2020) and Callaway, Drukker, Liu, and Sant'Anna (2023))

More Details

Multiple Time Periods and Variation in Treatment Timing

Conditional Parallel Trends with Multiple Periods

For all groups $g \in \bar{\square}$ (all groups except the never-treated group) and for all time periods $t = 2, \ldots, \square$,

$$\mathbb{E}[\Delta Y_t(0) | \mathbf{X}, Z, G = g] = \mathbb{E}[\Delta Y_t(0) | \mathbf{X}, Z, U = 1]$$

where $\mathbf{X} := (X_1, X_2, \ldots, X_{\square})$.

Under this assumption, using similar arguments to the ones above, one can show that

$$ATT(g,t) = \mathbb{E}\left[\left(\frac{\mathbf{1}\{G = g\}}{p_g} - \frac{p_g(\mathbf{X}, Z)U}{(1 - p_g(\mathbf{X}, Z))p_g}\right)\left(Y_t - Y_{g-1} - m^0_{gt}(\mathbf{X}, Z)\right)\right]$$

where $p_g(\mathbf{X}, Z) := P(G = g | \mathbf{X}, Z, \mathbf{1}\{G = g\} + U = 1)$ and $m^0_{gt}(\mathbf{X}, Z) := \mathbb{E}[Y_t - Y_{g-1} | \mathbf{X}, Z, U = 1]$.

Practical Considerations

Because $\mathbf{X}$ contains X_t for all time periods, terms like $m^0_{gt}(\mathbf{X}, Z)$ can be quite high-dimensional (and hard to estimate) in many applications. In many cases, it may be reasonable to replace with lower dimensional function $\mathbf{X}$:

- $\bar{X}_i$ — the average of X_{it} across time periods
- X_{it}, X_{ig-1} — the covariates in the current period and base period (this is possible in the `pte` package currently and may be added to `did` soon).
- X_{ig-1} — the covariates in the base period (this is the default in `did`)
- $(X_{it} - X_{ig-1})$ — the change in covariates over time

Empirical Example

Back to Minimum Wage Example

We'll allow for path of outcomes to depend on region of the country

```
# run TWFE regression
twfe_x <- fixest::feols(lemp ~ post | id + region^year,
                        data=data2)
modelsummary(twfe_x, gof_omit=".*")
```

	(1)
post	0.001
	(0.008)

Relative to previous results, this is much smaller and statistically insignificant and is similar to the result in Dube et al. (2010).

Callaway and Sant'Anna (2021) Results

```r
# callaway and sant'anna including covariates
cs_x <- att_gt(yname="lemp",
               tname="year",
               idname="id",
               gname="G",
               xformla=~region,
               control_group="nevertreated",
               base_period="universal",
               data=data2)
cs_x_res <- aggte(cs_x, type="group")
summary(cs_x_res)
cs_x_dyn <- aggte(cs_x, type="dynamic")
ggdid(cs_x_dyn)
```

Callaway and Sant'Anna (2021) Results

```
Call:
aggte(MP = cs_x, type = "group")

Reference: Callaway, Brantly and Pedro H.C. Sant'Anna.  "Difference-in-Differences with Multiple Time
Periods." Journal of Econometrics, Vol. 225, No. 2, pp. 200-230, 2021.
<https://doi.org/10.1016/j.jeconom.2020.12.001>, <https://arxiv.org/abs/1803.09015>


Overall summary of ATT's based on group/cohort aggregation:
     ATT    Std. Error     [ 95%  Conf. Int.]
 -0.0273         0.008    -0.0429     -0.0116 *


Group Effects:
 Group Estimate Std. Error [95% Simult.  Conf. Band]
  2004  -0.0436     0.0199       -0.0862     -0.0011 *
  2006  -0.0199     0.0077       -0.0364     -0.0034 *
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Callaway and Sant’Anna (2021) Results

Comments

Even more than in the previous case, the results in this case are notably different depending on the estimation strategy.

Add Covariates

```r
# run TWFE regression
twfe_x <- fixest::feols(lemp ~ post + lpop + lavg_pay | id + region^year,
                        data=data2)
modelsummary(twfe_x, gof_omit=".*")
```

	(1)
post	−0.022
	(0.008)
lpop	1.057
	(0.137)
lavg_pay	0.074
	(0.079)

Regression Adjustment

```
# callaway and sant'anna including covariates
cs_x <- att_gt(yname="lemp",
               tname="year",
               idname="id",
               gname="G",
               xformla=~region + lpop + lavg_pay,
               control_group="nevertreated",
               base_period="universal",
               est_method="reg",
               data=data2)
cs_x_res <- aggte(cs_x, type="group")
summary(cs_x_res)
cs_x_dyn <- aggte(cs_x, type="dynamic")
ggdid(cs_x_dyn)
```

Regression Adjustment

```
Call:
aggte(MP = cs_x, type = "group")

Reference: Callaway, Brantly and Pedro H.C. Sant'Anna.  "Difference-in-Differences with Multiple Time
Periods." Journal of Econometrics, Vol. 225, No. 2, pp. 200-230, 2021.
<https://doi.org/10.1016/j.jeconom.2020.12.001>, <https://arxiv.org/abs/1803.09015>


Overall summary of ATT's based on group/cohort aggregation:
     ATT    Std. Error     [ 95%  Conf. Int.]
 -0.0321        0.0081     -0.048     -0.0162 *


Group Effects:
 Group Estimate Std. Error [95% Simult.  Conf. Band]
  2004  -0.0596     0.0195       -0.1016     -0.0175 *
  2006  -0.0197     0.0081       -0.0372     -0.0022 *
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Outcome Regression
```

Regression Adjustment

Inverse Propensity Score Weighting

```r
# callaway and sant'anna including covariates
cs_x <- att_gt(yname="lemp",
               tname="year",
               idname="id",
               gname="G",
               xformla=~region + lpop + lavg_pay,
               control_group="nevertreated",
               base_period="universal",
               est_method="ipw",
               data=data2)
cs_x_res <- aggte(cs_x, type="group")
summary(cs_x_res)
cs_x_dyn <- aggte(cs_x, type="dynamic")
ggdid(cs_x_dyn)
```

Inverse Propensity Score Weighting

```
Call:
aggte(MP = cs_x, type = "group")

Reference: Callaway, Brantly and Pedro H.C. Sant'Anna.  "Difference-in-Differences with Multiple Time
Periods." Journal of Econometrics, Vol. 225, No. 2, pp. 200-230, 2021.
<https://doi.org/10.1016/j.jeconom.2020.12.001>, <https://arxiv.org/abs/1803.09015>


Overall summary of ATT's based on group/cohort aggregation:
     ATT    Std. Error     [ 95%  Conf. Int.]
 -0.0313        0.0081    -0.0472     -0.0154 *


Group Effects:
 Group Estimate Std. Error [95% Simult.  Conf. Band]
  2004  -0.0514     0.0191       -0.0911     -0.0117 *
  2006  -0.0222     0.0079       -0.0387     -0.0057 *
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Inverse Probability Weighting
```

Inverse Propensity Score Weighting

Doubly Robust

```
# callaway and sant'anna including covariates
cs_x <- att_gt(yname="lemp",
               tname="year",
               idname="id",
               gname="G",
               xformla=~region + lpop + lavg_pay,
               control_group="nevertreated",
               base_period="universal",
               est_method="dr",
               data=data2)
cs_x_res <- aggte(cs_x, type="group")
summary(cs_x_res)
cs_x_dyn <- aggte(cs_x, type="dynamic")
ggdid(cs_x_dyn)
```

Doubly Robust

```
Call:
aggte(MP = cs_x, type = "group")

Reference: Callaway, Brantly and Pedro H.C. Sant'Anna.  "Difference-in-Differences with Multiple Time
Periods." Journal of Econometrics, Vol. 225, No. 2, pp. 200-230, 2021.
<https://doi.org/10.1016/j.jeconom.2020.12.001>, <https://arxiv.org/abs/1803.09015>


Overall summary of ATT's based on group/cohort aggregation:
     ATT    Std. Error     [ 95%  Conf. Int.]
 -0.0317         0.008    -0.0474     -0.0159 *


Group Effects:
 Group Estimate Std. Error [95% Simult.  Conf. Band]
  2004  -0.0509     0.0187       -0.0918     -0.0099 *
  2006  -0.0230     0.0069       -0.0381     -0.0080 *
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Doubly Robust

Change base period

```
# callaway and sant'anna including covariates
cs_x <- att_gt(yname="lemp",
               tname="year",
               idname="id",
               gname="G",
               xformla=~region + lpop + lavg_pay,
               control_group="nevertreated",
               base_period="varying",
               est_method="dr",
               data=data2)
cs_x_res <- aggte(cs_x, type="group")
summary(cs_x_res)
cs_x_dyn <- aggte(cs_x, type="dynamic")
ggdid(cs_x_dyn)
```

Change base period

```
Call:
aggte(MP = cs_x, type = "group")

Reference: Callaway, Brantly and Pedro H.C. Sant'Anna.  "Difference-in-Differences with Multiple Time
Periods." Journal of Econometrics, Vol. 225, No. 2, pp. 200-230, 2021.
<https://doi.org/10.1016/j.jeconom.2020.12.001>, <https://arxiv.org/abs/1803.09015>


Overall summary of ATT's based on group/cohort aggregation:
     ATT    Std. Error     [ 95%  Conf. Int.]
 -0.0317        0.0077    -0.0469     -0.0165 *


Group Effects:
 Group Estimate Std. Error [95% Simult.  Conf. Band]
  2004  -0.0509      0.018       -0.0897     -0.0120 *
  2006  -0.0230      0.007       -0.0382     -0.0079 *
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Change base period

Change comparison group

```
# callaway and sant'anna including covariates
cs_x <- att_gt(yname="lemp",
               tname="year",
               idname="id",
               gname="G",
               xformla=~region + lpop + lavg_pay,
               control_group="notyettreated",
               base_period="universal",
               est_method="dr",
               data=data2)
cs_x_res <- aggte(cs_x, type="group")
summary(cs_x_res)
cs_x_dyn <- aggte(cs_x, type="dynamic")
ggdid(cs_x_dyn)
```

Change comparison group

```
Call:
aggte(MP = cs_x, type = "group")

Reference: Callaway, Brantly and Pedro H.C. Sant'Anna.  "Difference-in-Differences with Multiple Time
Periods." Journal of Econometrics, Vol. 225, No. 2, pp. 200-230, 2021.
<https://doi.org/10.1016/j.jeconom.2020.12.001>, <https://arxiv.org/abs/1803.09015>


Overall summary of ATT's based on group/cohort aggregation:
     ATT    Std. Error     [ 95%  Conf. Int.]
 -0.0312         0.008    -0.0468     -0.0155 *


Group Effects:
 Group Estimate Std. Error [95% Simult.  Conf. Band]
  2004  -0.0493     0.0189       -0.0917     -0.0069 *
  2006  -0.0230     0.0072       -0.0392     -0.0069 *
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Not Yet Treated,  Anticipation Periods:  0
Estimation Method:  Doubly Robust
```

Change comparison group

Allow for anticipation

```
# callaway and sant'anna including covariates
cs_x <- att_gt(yname="lemp",
               tname="year",
               idname="id",
               gname="G",
               xformla=~region + lpop + lavg_pay,
               control_group="nevertreated",
               base_period="universal",
               est_method="dr",
               anticipation=1,
               data=data2)
cs_x_res <- aggte(cs_x, type="group")
summary(cs_x_res)
cs_x_dyn <- aggte(cs_x, type="dynamic")
ggdid(cs_x_dyn)
```

Allow for anticipation

```
Call:
aggte(MP = cs_x, type = "group")

Reference: Callaway, Brantly and Pedro H.C. Sant'Anna.  "Difference-in-Differences with Multiple Time
Periods." Journal of Econometrics, Vol. 225, No. 2, pp. 200-230, 2021.
<https://doi.org/10.1016/j.jeconom.2020.12.001>, <https://arxiv.org/abs/1803.09015>


Overall summary of ATT's based on group/cohort aggregation:
    ATT    Std. Error     [ 95%  Conf. Int.]
 -4e-04        0.0095    -0.0191      0.0182


Group Effects:
 Group Estimate Std. Error [95% Pointwise  Conf. Band]
  2006   -4e-04     0.0091         -0.0183      0.0174
---
Signif. codes: `*' confidence band does not cover 0

Control Group:  Never Treated,  Anticipation Periods:  1
Estimation Method:  Doubly Robust
```

Allow for anticipation

Covariates Affected by the Treatment

Motivation

So far, our discussion has been for the case where the time-varying covariates evolve exogenously.

- Many (probably most) covariates fit into this category: in the minimum wage example, a county's population probably fits here.

But in other cases, there can exist covariates that we would like to include in the parallel trends assumption that could be affected by the treatment (this type of covariate is often referred to as a bad control.

- Example: Job displacement
  - → Path of earnings (in the absence of job displacement) likely depends on occupation, industry, and/or union status, which could all be affected by job displacement

The traditional approach in empirical work is to completely drop covariates that could have been affected by the treatment.

- Not clear if this is the right idea though…

Additional Notation

To wrap our heads around this, let's go back to the case with two time periods.

Define treated and untreated potential covariates: $X_{it}(1)$ and $X_{it}(0)$. Notice that in the "textbook" two period setting, we observe

$$X_{it^*} = D_i X_{it^*}(1) + (1 - D_i) X_{it^*}(0) \qquad \text{and} \qquad X_{it^*-1} = X_{it^*-1}(0)$$

If the covariates are literally not affected by the treatment at all, then we can write the version of conditional parallel trends that we have been using above in terms of potential outcomes:

Conditional Parallel Trends using Untreated Potential Covariates

$$\mathbb{E}[\Delta Y_{t^*}(0) | X_{t^*}(0), X_{t^*-1}(0), Z, D = 1] = \mathbb{E}[\Delta Y_{t^*}(0) | X_{t^*}(0), X_{t^*-1}(0), Z, D = 0]$$

Alternative Identification Strategies

One idea is to just ignore that the covariates may have been affected by the treatment:

Alternative Conditional Parallel Trends 1

$$\mathbb{E}[\Delta Y_{t^*}(0)|X_{t^*}, X_{t^*-1}(0), Z, D = 1] = \mathbb{E}[\Delta Y_{t^*}(0)|X_{t^*}, X_{t^*-1}(0), Z, D = 0]$$

The limitations of this approach are well known (even discussed in MHE), and this is not typically the approach taken in empirical work

Alternative Identification Strategies

It is more common in empirical work to drop $X_t(0)$ entirely from the parallel trends assumption

Alternative Conditional Parallel Trends 2

$$\mathbb{E}[\Delta Y_{t^*}(0)|Z, D = 1] = \mathbb{E}[\Delta Y_{t^*}(0)|Z, D = 0]$$

In my view, this is not attractive either though. If we believe this assumption, then we have basically solved the bad control problem by assuming that it does not exist.

Alternative Identification Strategies

Perhaps a better alternative identifying assumption is the following one

Alternative Conditional Parallel Trends 3

$$\mathbb{E}[\Delta Y_{t^*}(0)|X_{t^*-1}(0), Z, D=1] = \mathbb{E}[\Delta Y_{t^*}(0)|X_{t^*-1}(0), Z, D=0]$$

This says that:

- Conditional parallel trends holds after conditioning on pre-treatment time-varying covariates that could have been affected by treatment
- It is not the same identifying assumption as the one that we started with, but we are at least allowing for the bad control to show up in the identifying assumption
- Since $X_{t^*-1}(0)$ is observed for all units, we can immediately operationalize this assumption use our arguments from earlier (i.e., the ones without bad controls)

- In practice, you can just include the bad control among other covariates in `did`

Identification under CPT

Going back to the original conditional parallel trends assumption (that include both $X_{t^*}(0)$ and $X_{t^*-1}(0)$)...Using the same sort of arguments as for regression adjustment earlier, it follows that

$$ATT = \mathbb{E}[\Delta Y_{t^*}|D=1] - \mathbb{E}\Big[\mathbb{E}[\Delta Y_{t^*}(0)|X_{t^*}(0), X_{t^*-1}(0), Z, D=0]\Big|D=1\Big]$$

Here:

- The inside conditional expectation is identified — we see untreated potential outcomes and covariates for the untreated group
- But the outside expectation is infeasible $\implies$ we are stuck

Dealing with $X_{t^*}(0)$

Covariate Unconfoundedness Assumption

$X_{t^*}(0) \perp\!\!\!\perp D | X_{t^*-1}(0), Z$

Intuition: For the treated group, the time-varying covariate would have evolved in the same way over time as it actually did for the untreated group, conditional on X_{t^*-1} and Z.

- Notice that this assumption only concerns untreated potential covariates $\implies$ it allows for X_{t^*} to be affected by the treatment
- Making an assumption like this indicates that $X_{t^*}(0)$ is playing a dual role: (i) start by treating it as if it's an outcome, (ii) have it continue to play a role as a covariate

Under this assumption, can show that we can recover the ATT:

$$ATT = \mathbb{E}[\Delta Y_{t^*} | D = 1] - \mathbb{E}\left[\mathbb{E}[\Delta Y_{t^*} | X_{t^*-1}, Z, D = 0] \middle| D = 1\right]$$

Additional Discussion

In some cases, it may make sense to condition on other additional variables (e.g., the lagged outcome Y_{t^*-1}) in the covariate unconfoundedness assumption. In this case, it is still possible to identify ATT, but it is more complicated

It could also be possible to use alternative identifying assumptions besides covariate unconfoundedness — at a high-level, we somehow need to recover the distribution of $X_{t^*}(0)$

- e.g., Brown, Butts, and Westerlund (2023)

See Caetano et al. (2023) for more details about bad controls.

Appendix

Understanding Double Robustness

To understand double robustness, we can rewrite the expression for ATT as

$$\text{ATT} = \mathbb{E}\left[\frac{D}{p}\left(\Delta Y_{t^*} - m_0(X)\right)\right] - \mathbb{E}\left[\frac{p(X)(1-D)}{(1-p(X))p}\left(\Delta Y_{t^*} - m_0(X)\right)\right]$$

The first term is exactly the same as what comes from regression adjustment

- If we correctly specify a model for $m_0(X)$, it will be equal to ATT.
- If $m_0(X)$ not correctly specified, then, by itself, this term will be biased for ATT

The second term can be thought of as a de-biasing term

- If $m_0(X)$ is correctly specified, it is equal to 0
- If $p(X)$ is correctly specified, it reduces to $\mathbb{E}[\Delta Y_t(0)|D=1] - \mathbb{E}[m_0(X)|D=1]$ which both delivers counterfactual untreated potential outcomes and removes the (possibly misspecified) second term from the first equation

Back