Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects

Liyang Sun\(^1\)    Sarah Abraham\(^2\)
Econometric Society Winter Meeting

\(^1\)Department of Economics, MIT
\(^2\)Cornerstone Research
Introduction
• Researchers often estimate dynamic treatment effects by the estimates for coefficients $\mu_\ell$ in a (dynamic) two-way FE specification that resembles the following:

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_\ell \mu_\ell D_{i,t}^\ell + \nu_{i,t}$$

where $D_{i,t}^\ell$ is an indicator for $\ell$ periods relative to $i$’s initial treatment ($\ell = 0$ is the period of initial treatment).
Researchers often estimate dynamic treatment effects by the estimates for coefficients $\mu_\ell$ in a (dynamic) two-way FE specification that resembles the following:

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_\ell \mu_\ell D_{i,t}^\ell + \nu_{i,t}$$

where $D_{i,t}^\ell$ is an indicator for $\ell$ periods relative to $i$’s initial treatment ($\ell = 0$ is the period of initial treatment).

Goal of our project: characterize $\mu_\ell$ under heterogenous treatment effect in event studies

- absorbing treatment
- variation in treatment timing
Examples of event studies

- Deryugina (2017): the non-transient fiscal cost for a county that has been hit by a hurricane, where the initial treatment is the first hurricane experienced in a county
Examples of event studies

• Deryugina (2017): the non-transient fiscal cost for a county that has been hit by a hurricane, where the initial treatment is the first hurricane experienced in a county

• Dobkin et al. (2018): the dynamic effects of a hospitalization, where the initial treatment is the initial hospital admission
• Intuition: compare earlier cohort with later cohort to estimate the dynamic effect
• Report estimates for relative period coefficients $\mu_\ell$ as estimates for dynamic effects
• Potentially problematic when there are multiple cohorts:
  • Decompose $\mu_\ell$ in terms of cohort-specific effects
  • Demonstrate potential contamination from $\ell' \neq \ell$ due to heterogeneity
Overview

- Intuition: compare earlier cohort with later cohort to estimate the dynamic effect
- Report estimates for relative period coefficients $\mu_\ell$ as estimates for dynamic effects
- Potentially problematic when there are multiple cohorts:
  - Decompose $\mu_\ell$ in terms of cohort-specific effects
  - Demonstrate potential contamination from $\ell' \neq \ell$ due to heterogeneity
- Event studies can show up under different names
  - “staggered adoption” (Athey and Imbens, 2018)
  - “stepped wedge design” (Ellenberg, JAMA 2018)
For today’s talk

- Literature review
- Cast event studies in a potential outcomes framework
- Decomposition
- Alternative methods
- Empirical illustration
Literature review
• Active literature on the causal interpretations of two-way fixed effects regressions (Athey and Imbens, 2018; Borusyak and Jaravel, 2017; Callaway and Sant’Anna, 2020; de Chaisemartin and D’Haultfoeuille, 2020; Goodman-Bacon, 2018)

• Mostly focus on “static” specifications:

\[ Y_{i,t} = \alpha_i + \lambda_t + \mu D_{i,t} + \nu_{i,t} \]

• Decompose \( \mu \) into non-convex combination of heterogeneous treatment effects

• We focus on “dynamic” specifications
• de Chaisemartin and D’Haultfœuille (2020) propose diagnostic tools and alternative estimators
• Callaway and Sant’Anna (2020) allows for conditioning on time-varying covariates
• We propose simple regression-based alternative estimation strategy for dynamic treatment effects
Potential outcome framework
Potential outcome framework: 1/3

- A random sample of $N$ units observed over $T + 1$ time periods with i.i.d. observations $\{Y_{i,t}, D_{i,t}\}_{t=0}^T$
- Treatment status $D_{i,t} \in \{0, 1\}$
- Treatment is actually a sequence $\{D_{i,s}\}_{s=0}^T$.
  - In event studies, can be identified with the scalar $E_i = \min \{t : D_{i,t} = 1\}$, the period of initial treatment
  - Let $E_i = \infty$ for those never treated
  - **Treatment cohort**: a group $\{i : E_i = e\}$ of unit first treated at the same time
- Results hold with or without a never treated cohort (control cohort).
• Potential outcome $Y_{i,t}^e$ is the outcome in response to treatment that first starts in period $e$
• “Baseline outcome” $Y_{i,t}^\infty$ is the potential outcome if never treated
• Observed outcome is therefore

$$Y_{i,t} = Y_{i,t}^E = Y_{i,t}^\infty + \sum_{0\leq e \leq T} (Y_{i,t}^e - Y_{i,t}^\infty) \cdot 1\{E_i = e\}$$

• Unit-level treatment effect $Y_{i,t} - Y_{i,t}^\infty$
• “Building blocks” for causal interpretation are cohort-specific average treatment effect on the treated (CATT) \( \ell \) periods from initial treatment:

\[
CATT_{e,\ell} = E[Y_{i,e+\ell} - Y_{i,e+\ell}^\infty \mid E_i = e]
\]

• This object coincides with the “group-time average treatment effect” studied by Callaway and Sant’Anna (2020)

• Next use potential outcome notations to articulate identifying assumptions underlying the dynamic specification
Assumption 1.

*(Parallel trends in baseline outcome.*) For all \( s \neq t \), the
\[ E[Y_{i,t} - Y_{i,s} | E_i = e] \] is the same for all \( e \in supp(E_i) \).

- Potential violation: Ashenfelter’s dip
Assumption 2.

(No anticipation.) There is no treatment effect in pre-treatment periods i.e. $E[Y_{i,e+\ell}^e - Y_{i,e+\ell}^\infty | E_i = e] = 0$ for all $e \in \text{supp}(E_i)$ and all $\ell < 0$.

- Potential violation: Hendren (2017) shows that knowledge of future job loss leads to decreases in consumption due to anticipation
- Similar to Malani and Reif (2015) and Botosaru and Gutierrez (2018)
Assumption 3.

(Treatment effect homogeneity.) For each relative period $\ell$, $CATT_{e,\ell}$ does not depend on cohort $e$ and is equal to $ATT_{\ell}$.

Potential violations:

- Effects vary with covariates
- Selection into treatment timing based on effects
- Calendar time-varying effects (e.g. macroeconomic conditions could govern the effects on labor market outcomes)
Decompose the dynamic specification
Dynamic specification

\[ Y_{i,t} = \alpha_i + \lambda_t + \sum_{\ell=-K}^{-2} \mu_\ell D_{i,t}^\ell + \sum_{\ell=0}^{L} \mu_\ell D_{i,t}^\ell + \nu_{i,t} \]

- \( \mu_\ell \) denotes the population regression coefficient, i.e. the probability limit of the associated OLS estimator \( \hat{\mu}_\ell \)
- Included relative periods collected in \( g^{incl} = \{-K, \ldots, 0, \ldots, L\} \)
- Excluded relative periods collected in \( g^{excl} = \{-T, \ldots, -K - 1, -1, L + 1, \ldots, T\} \)
- Excluding some relative periods can be necessary due to multi-collinearity (Borusyak and Jaravel, 2017)
Proposition 1.

Under parallel trends, we can write $\mu_\ell$ as a linear combination of $CATT_{e,\ell}$ as well as $CATT_{e,\ell'}$ from other relative periods $\ell' \neq \ell$,

$$\mu_\ell = \sum_{e} \omega_{e,\ell} CATT_{e,\ell} + \sum_{\ell' \neq \ell, \ell' \in g^{incl}} \sum_{e} \omega_{e,\ell'} CATT_{e,\ell'} + \sum_{\ell' \in g^{excl}} \sum_{e} \omega_{e,\ell'} CATT_{e,\ell'}$$

1. For own relative period: weights sum to one
2. For other relative periods included in the specification: weights sum to zero for each $\ell' \neq \ell$
3. For relative periods excluded from the specification:
   $$\sum_{\ell' \in g^{excl}} \sum_{e} \omega_{e,\ell'} = -1$$

If the weights $\omega_{e,\ell'}$ are non-zero, then effects from $\ell' \neq \ell$ can potentially contaminate the interpretation of $\mu_\ell$
Consider a saturated regression:

\[ Y_{i,t} = \sum_{e} \alpha_{e} \cdot 1\{E_i = e\} + \sum_{s} \lambda_{s} \cdot 1\{t = s\} \]

\[ + \sum_{\ell \in g^{incl}} \sum_{e} \gamma_{e,\ell} \cdot \left( D_{i,t}^{\ell} \cdot 1\{E_i = e\} \right) \]

\[ + \sum_{\ell' \in g^{excl}} \sum_{e} \gamma_{e,\ell'} \cdot \left( D_{i,t}^{\ell'} \cdot 1\{E_i = e\} \right) + \epsilon_{i,t} \]
Consider a saturated regression:

\[
Y_{i,t} = \sum_{e} \alpha_e \cdot 1\{E_i = e\} + \sum_{s} \lambda_s \cdot 1\{t = s\} \\
\quad + \sum_{\ell \in g^{incl}} \sum_{e} \gamma_{e,\ell} \cdot \left( D_{i,t}^\ell \cdot 1\{E_i = e\} \right) \\
\quad + \sum_{\ell' \in g^{excl}} \sum_{e} \gamma_{e,\ell'} \cdot \left( D_{i,t}^{\ell'} \cdot 1\{E_i = e\} \right) + \epsilon_{i,t}
\]

Under parallel trends, we know the population coefficients:

\[
Y_{i,t} = \sum_{e} E[Y_{i,0}^\infty | E_i = e] \cdot 1\{E_i = e\} + \sum_{s} E[Y_{i,s}^\infty - Y_{i,0}^\infty] \cdot 1\{t = s\} \\
\quad + \sum_{\ell \in g^{incl}} \sum_{e} CATT_{e,\ell} \cdot \left( D_{i,t}^\ell \cdot 1\{E_i = e\} \right) \\
\quad + \sum_{\ell' \in g^{excl}} \sum_{e} CATT_{e,\ell'} \cdot \left( D_{i,t}^{\ell'} \cdot 1\{E_i = e\} \right) + \epsilon_{i,t}
\]
Apply the OVB formula to derive the expression for $\mu_\ell$ in terms of $CATT_{e,\ell'}$:

- finds its associated regressor in the saturated regression $D^\ell_{i,t} \cdot 1 \{E_i = e\}$
- multiplies it with the regression coefficients from

$$D^\ell_{i,t} \cdot 1 \{E_i = e\} = \alpha_i + \lambda_t + \sum_{\ell \in g^{incl}} \omega^\ell_{e,\ell'} D^\ell_{i,t} + u_{i,t}$$
Intuition for the weights

\[ D_{i,t}^{\ell'} \cdot 1 \{ E_i = e \} = \alpha_i + \lambda_t + \sum_{\ell \in g^{incl}} \omega_{e,\ell'} D_{i,t}^{\ell} + u_{i,t} \]

- We provide code for calculating these weights eventstudyweights
- \( \omega_{e,\ell'} \neq 0 \) even for \( \ell' \neq \ell \) because the panel cannot be balanced in both calendar and relative times when there are multiple cohorts
- Magnitude of \( \omega_{e,\ell'} \) determines how sensitive \( \mu_\ell \) is to \( CAT T_{e,\ell'} \)
- Can invalidate a test for pre-trend
Proposition 2.

*Under parallel trends and no anticipation, we can write a lead coefficient $\mu_{\ell}$ for $\ell < 0$ as a linear combination of post-treatment $\text{CATT}_{e, \ell'}$ for all $\ell' \geq 0$:

$$\mu_{\ell} = \sum_{\ell' \geq 0} \sum_{e} \omega_{e, \ell'} \text{CATT}_{e, \ell'} + \sum_{\ell' \in \text{g-excl}, \ell' \geq 0} \sum_{e} \omega_{e, \ell'} \text{CATT}_{e, \ell'}$$

• Even if $\text{CATT}_{e, \ell} = 0$ for all $\ell < 0$, can still get non-zero $\mu_{\ell}$ due to contamination from post-treatment periods
Proposition 3.

Under parallel trends and treatment effect homogeneity, we have

\[ CATT_{e,\ell} = ATT_\ell \] for a given \( \ell \) and

\[ \mu_\ell = ATT_\ell + \sum_{\ell' \in g^{excl}} \omega^{\ell}_{\ell'} ATT_{\ell'} \]

- Additional term drops out when excluded periods have zero effect, otherwise can be thought of as a type of “normalization”:

\[ \sum_{\ell' \in g^{excl}} \omega^{\ell}_{\ell'} = -1 \]
Alternative methods
Alternative methods for estimating dynamic treatment effects

Consider the estimand:

\[ \nu_\ell = \sum_e CATT_{e,\ell} Pr\{ E_i = e \mid E_i \in [-\ell, T - \ell] \} \]

We propose the interaction-weighted (IW) estimator à la Gibbons et al. (2018)

1. Estimate \( CATT_{e,\ell} \) by

\[ Y_{i,t} = \alpha_i + \lambda_t + \sum_{e \not\in C} \sum_{\ell \neq -1} \delta_{e,\ell} (1\{ E_i = e \} \cdot D_{i,t}^\ell) + \epsilon_{i,t} \]

2. Estimate the weights by sample shares

3. Form the IW estimator \( \hat{\nu}_\ell \) by

\[ \hat{\nu}_\ell = \sum_e \hat{\delta}_{e,\ell} \hat{Pr}\{ E_i = e \mid E_i \in [-\ell, T - \ell] \} \]
Proposition 4.

Under parallel trends, no anticipation and some regularity conditions, the IW estimator $\hat{\nu}_\ell$ is consistent and asymptotically normal:

$$\sqrt{N} (\hat{\nu}_\ell - \nu_\ell) \rightarrow_d N(0, \Sigma)$$

- Note that $\hat{\delta}_{e,\ell}$ is a difference-in-differences estimator for $CATT_{e,\ell}$
- $\Sigma$ accounts for asymptotic variance from both $\hat{\delta}_{e,\ell}$ and the sample shares
Empirical illustration
Consequences of hospitalization (Dobkin et al., 2018)

Dynamic two-way fixed effects regression

\[ Y_{i,t} = \alpha_i + \lambda_t + \mu_{-3} D_{i,t}^{\ell} + \mu_{-2} D_{i,t}^{\ell} + \mu_0 D_{i,t}^0 + \mu_1 D_{i,t}^1 + \mu_2 D_{i,t}^2 + \mu_3 D_{i,t}^3 + \nu_{i,t} \]

- \( Y_{i,t} \) out-of-pocket medical spending; \( D_{i,t}^{\ell} \) period relative to initial hospitalization
- Balanced panel of \( N = 656 \) over \( t \in \{0, \ldots, 4\} \) from Health and Retirement Study (HRS)
- Four cohorts \( E_i \in \{1, 2, 3, 4\} \) with \( \ell \in \{-3, -2, 0, 1, 2, 3\} \) included but \( \ell \in \{-4, -1\} \) excluded
Decomposition of $\mu_{-2}$

$$\mu_{-2} = \sum_{e=1}^{4} \omega_{e,-2}^2 CATT_{e,-2}$$

own period

$$+ \sum_{\ell \in \{-3,0,1,2,3\}} \sum_{e=1}^{4} \omega_{e,\ell}^2 CATT_{e,\ell}$$

other included period

$$+ \sum_{\ell' \in \{-4,-1\}} \sum_{e=1}^{4} \omega_{e,\ell'}^2 CATT_{e,\ell'}$$

excluded period

Weights on 1st cohort CATT
Weights on 2nd cohort CATT
Weights on 3rd cohort CATT
Weights on 4th cohort CATT
Decomposition of $\mu_{-2}$

$$
\mu_{-2} = \sum_{e=1}^{4} \omega_{e,-2} CATT_{e,-2} \\
+ \sum_{\ell \in \{-3,0,1,2,3\}} \sum_{e=1}^{4} \omega_{e,\ell} CATT_{e,\ell} \\
+ \sum_{\ell' \in \{-4,-1\}} \sum_{e=1}^{4} \omega_{e,\ell'} CATT_{e,\ell'}
$$
Decomposition of $\mu_{-2}$

$$\mu_{-2} = \sum_{e=1}^{4} \omega_{e,-2}^{-2} CATT_{e,-2}$$

$$+ \sum_{\ell \in \{-3,0,1,2,3\}} \sum_{e=1}^{4} \omega_{e,\ell}^{-2} CATT_{e,\ell}$$

other included period

$$+ \sum_{\ell' \in \{-4,-1\}} \sum_{e=1}^{4} \omega_{e,\ell'}^{-2} CATT_{e,\ell'}$$

excluded period

---

![Graph showing weights on different cohorts of CATT]

- **Weights on 1st cohort CATT**
- **Weights on 2nd cohort CATT**
- **Weights on 3rd cohort CATT**
- **Weights on 4th cohort CATT**
### Effect of Hospitalization on Out-of-pocket Medical Spending

| $\ell$ Relative to Hospitalization | $\hat{\mu}_\ell$ | $\hat{\nu}_\ell$ | $\hat{\delta}_{1,\ell}$ | $\hat{\delta}_{2,\ell}$ | $\hat{\delta}_{3,\ell}$ |
|-----------------------------------|------------------|------------------|---------------------------|---------------------------|---------------------------|
| -3                                | 149              | 591              | -                         | -                         | 591                       |
|                                  | (792)            | (1273)           |                           |                           | (1273)                    |
| -2                                | 203              | 353              | -                         | 299                       | 411                       |
|                                  | (480)            | (698)            |                           | (967)                     | (1030)                    |
| -1                                | 0                | 0                | 0                         | 0                         | 0                         |
| 0                                 | 3,013            | 2,960            | 2,826                     | 3,031                     | 3,092                     |
|                                  | (511)            | (543)            | (1038)                    | (704)                     | (998)                     |
| 1                                 | 888              | 530              | 825                       | 107                       | -                         |
|                                  | (664)            | (587)            | (912)                     | (653)                     |                           |
| 2                                 | 1,172            | 800              | 800                       | -                         | -                         |
|                                  | (983)            | (1010)           | (1010)                    |                           |                           |
| 3                                 | 1,914            | -                | -                         | -                         | -                         |
|                                  | (1426)           |                  |                           |                           |                           |
- Decompose the relative period coefficient $\mu_\ell$ from dynamic specification for event studies
- Demonstrate that under treatment effects heterogeneity $\mu_\ell$ may pick up spurious terms consisting of treatment effects from periods other than $\ell$
- Propose “interaction-weighted” (IW) estimator that is more robust toward heterogeneity
Conclusion

• Decompose the relative period coefficient $\mu_\ell$ from dynamic specification for event studies
• Demonstrate that under treatment effects heterogeneity $\mu_\ell$ may pick up spurious terms consisting of treatment effects from periods other than $\ell$
• Propose “interaction-weighted” (IW) estimator that is more robust toward heterogeneity
• Thank you!