# Heterogeneous difference-in-differences estimation

StataCorp LLC

July 20, 2023

# Context: Treatment-effects estimation in Stata

The effect of a treatment or exposure on an outcome

# Context: Treatment-effects estimation in Stata

The effect of a treatment or exposure on an outcome

- Average treatment effect (ATE) and average treatment effect on the treated (ATET)

## Context: Treatment-effects estimation in Stata

The effect of a treatment or exposure on an outcome

- Average treatment effect (ATE) and average treatment effect on the treated (ATET)
- `teffects`: cross-sectional data selection on observables
- `didregress` and `xtdidregress`: repeated-measures selection on unobservables
- Estimation, inference, visualization, diagnostics, and tests

## Context: Treatment-effects estimation in Stata

The effect of a treatment or exposure on an outcome

- Average treatment effect (ATE) and average treatment effect on the treated (ATET)
- `teffects`: cross-sectional data selection on observables
- `didregress` and `xtdidregress`: repeated-measures selection on unobservables
- Estimation, inference, visualization, diagnostics, and tests
- One treatment effect. That assumes that the treatment does not change over groups or time. Homogeneous treatment effects
- New: Heterogeneous treatment effects for selection on unobservables. Heterogeneous Difference in differences (DID). Multiple ATETs

# Starting point: Why heterogeneous treatment effects?

- DID with different groups treated at different times

## Starting point: Why heterogeneous treatment effects?

- DID with different groups treated at different times
- With multiple treatment times the ATET for DID was obtained via

$$y_{it} = \beta_0 + D_{it}\beta_1 + \gamma_t + \gamma_g + e_{it}$$

- $\beta_1$ is the ATET
- A generalization of the well understood 2 by 2 model.

# Starting point: Why heterogeneous treatment effects?

- DID with different groups treated at different times
- With multiple treatment times the ATET for DID was obtained via

$$y_{it} = \beta_0 + D_{it}\beta_1 + \gamma_t + \gamma_g + e_{it}$$

- $\beta_1$ is the ATET
- A generalization of the well understood 2 by 2 model.
- Two-way fixed-effects (TWFE) was criticized in the last six years
- Why:
    1. Homogeneity (well understood)
    2. Hidden cost of generalizing of 2 by 2

# What happens when we go beyond the 2 by 2 model?

- ATET might be different for groups treated at different times
- Within groups, ATET might change over time
- If this is the case:

$$y_{it} = \beta_0 + D_{it}\beta_1 + \gamma_t + \gamma_g + e_{it}$$

# What happens when we go beyond the 2 by 2 model?

- ATET might be different for groups treated at different times
- Within groups, ATET might change over time
- If this is the case:

$$y_{it} = \beta_0 + D_{it}\beta_1 + \gamma_t + \gamma_g + e_{it}$$

- Weighted average
  - Borusyak, Jaravel, and Spiess (2018)
  - de Chaisemartin and D'Haultfoeuille (2020)
  - Goodman-Bacon (2021) (estat bdecomp)

# Bacon decomposition

- $\beta_1$ is a weighted average of (2 by 2) estimates

$$y_{it} = \beta_0 + D_{it}\beta_1 + \gamma_t + \gamma_g + e_{it}$$

- 2 by 2 estimates using:
  1. Never treated groups
  2. Early treated groups
  3. Later treated groups

## Bacon decomposition

- $\beta_1$ is a weighted average of (2 by 2) estimates

$$y_{it} = \beta_0 + D_{it}\beta_1 + \gamma_t + \gamma_g + e_{it}$$

- 2 by 2 estimates using:
  1. Never treated groups
  2. Early treated groups
  3. Later treated groups
- Some of the comparisons are valid comparisons. Some are not.
- Validity: satisfying a parallel-trends assumption.

## Let's see it work

- What a 2 by 2 model looks like
- What a DID model with multiple treatment times looks like graphically
  - homogeneous treatment
  - heterogeneous treatment
- What the Bacon decomposition tells us
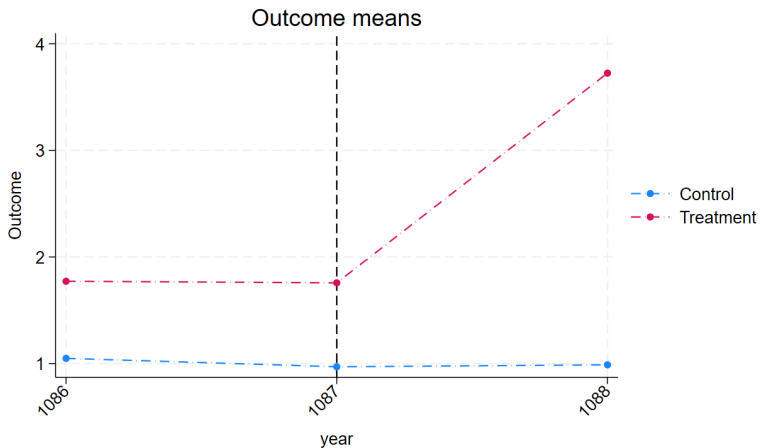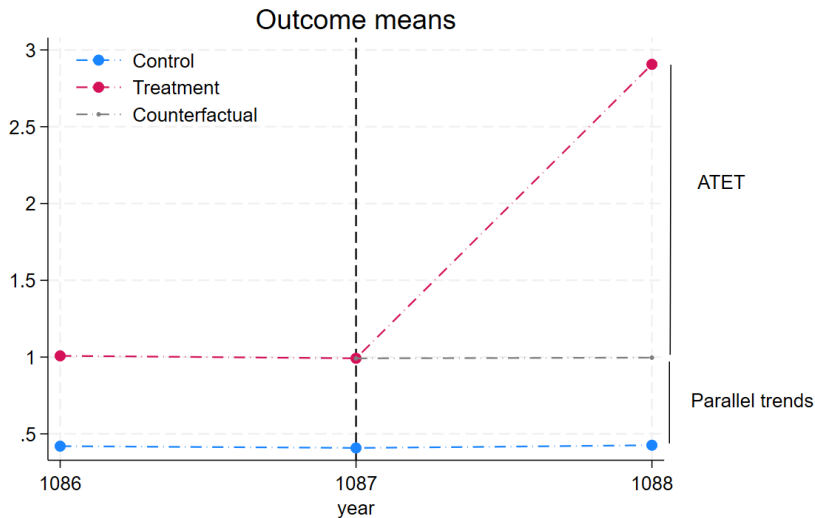  - homogeneous treatment
  - heterogeneous treatment

## Let's see it work

- What a 2 by 2 model looks like
- What a DID model with multiple treatment times looks like graphically
  - homogeneous treatment
  - heterogeneous treatment
- What the Bacon decomposition tells us
  - homogeneous treatment
  - heterogeneous treatment
- Build our understanding of heterogeneous DID

# 2 by 2 framework: `estat trendplots`

# 2 by 2 framework: Parallel(Common)-trends assumption

# Homogeneous treatment with multiple treatment times I

- 10 time periods
- Treatment occurs for some groups at time 3 (Earlier treated)
- For others at time 6 (Later treated)
- Some units remain untreated

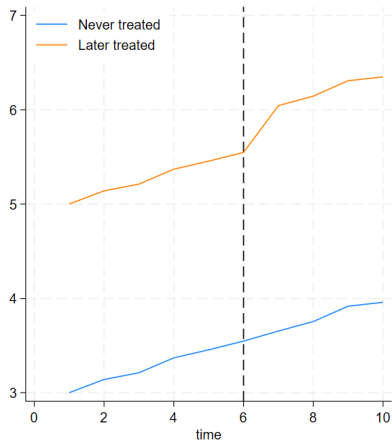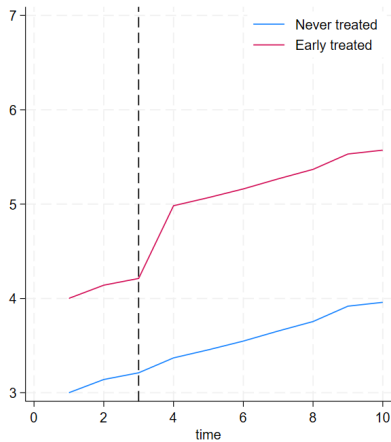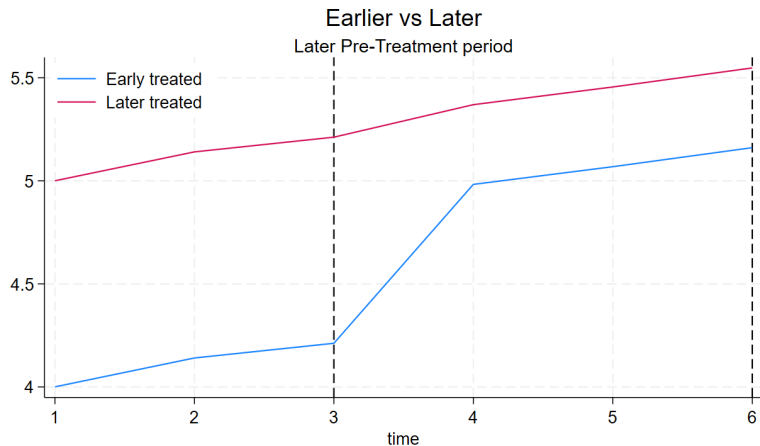# Homogeneous treatment with multiple treatment times I

- 10 time periods
- Treatment occurs for some groups at time 3 (Earlier treated)
- For others at time 6 (Later treated)
- Some units remain untreated
- What are the comparisons used by TWFE

# Homogeneous treatment with multiple treatment times I



Never treated vs. Later and Earlier

Earlier vs Later

# Homogeneous treatment: `estat bdecomp`
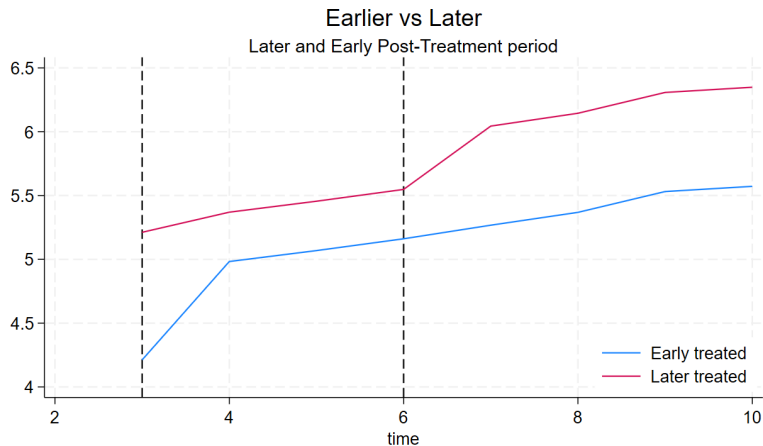
```
. estat bdecomp, summaryonly
DID treatment-effect decomposition
ATET = 4.041694                                    Number of obs     = 100,000
                                                   Number of groups  =  10,000
                                                   Number of cohorts =       3
ATET decomposition summary              ATET component                 Weight

Treated vs never treated                     4.0435917               0.865605
Treated earlier vs later                     4.0245287               0.057598
Treated later vs earlier                     4.0331799               0.076797
```

Note: Number of cohorts includes never treated.
Note: The ATET reported by xtdidregress is a weighted average of the ATET
      components. If any component is substantially different from the ATET
      reported by xtdidregress and the weight is large, consider accounting
      for treatment-effect heterogeneity by using xthdidregress.

# Heterogeneous treatment effect I



Never treated vs. Later and Earlier
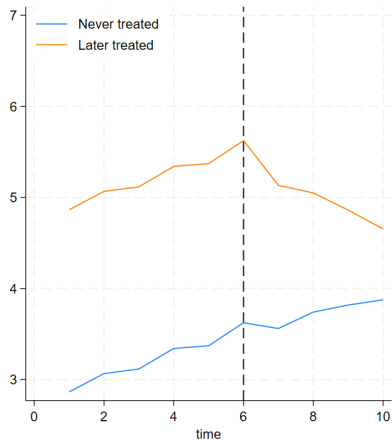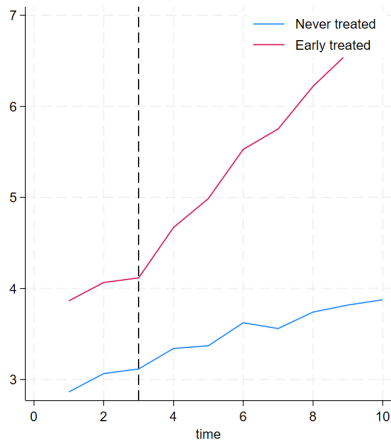
# Heterogeneous treatment effect II
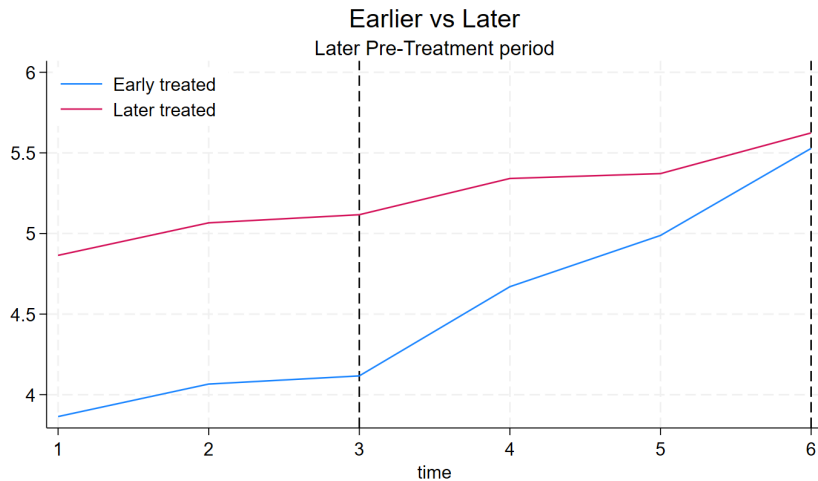
# Heterogeneous treatment effect III
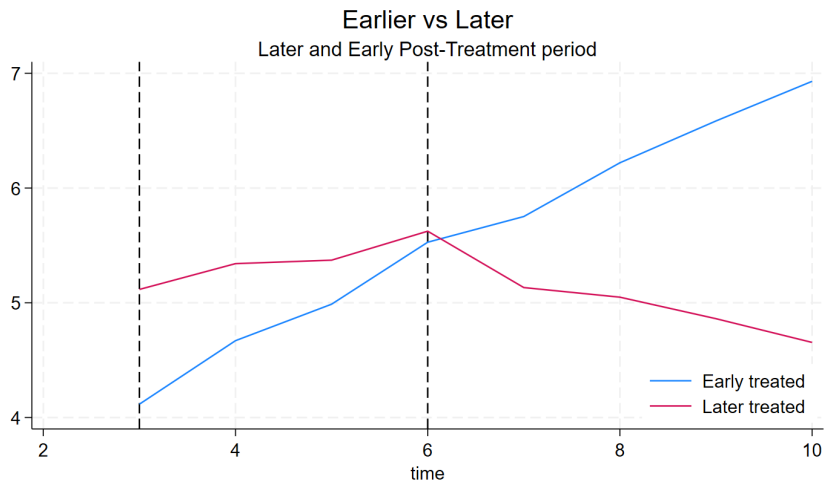
# Heterogeneous treatment: estat bdecomp

```
  . estat bdecomp, summaryonly
DID treatment-effect decomposition
ATET = 1.389688                                    Number of obs     = 50,000
                                                   Number of groups  =  5,000
                                                   Number of cohorts =      3
ATET decomposition summary              ATET component                 Weight

Treated vs never treated                     1.9728088               0.860939
Treated earlier vs later                     4.3441312               0.059597
Treated later vs earlier                    -7.1439268               0.079463

Note: Number of cohorts includes never treated.
Note: The ATET reported by xtdidregress is a weighted average of the ATET
      components. If any component is substantially different from the ATET
      reported by xtdidregress and the weight is large, consider accounting
      for treatment-effect heterogeneity by using xthdidregress.
```

## What would the aggregated effect be with "good" comparison groups?

```
. estat aggregation
Overall ATET                                            Number of obs = 50,000
                              (Std. err. adjusted for 50 clusters in state)
```

| y | ATET | Robust std. err. | z | P>\|z\| | [95% conf. interval] | |
|---|---|---|---|---|---|---|
| tr (1 vs 0) | 3.654175 | .962046 | 3.80 | 0.000 | 1.768599 | 5.53975 |

## What have we learned?

- With multiple treatment times, traditional DID assumes homogeneity
- 2 by 2 comparisons are well defined, but not all are useful
    - Comparison group matters
    - Time of comparison matters
- With multiple treatment times, it is important to assess heterogeneity
- If we suspect heterogeneity or do not want to assume homogeneity,

# What have we learned?

- With multiple treatment times, traditional DID assumes homogeneity
- 2 by 2 comparisons are well defined, but not all are useful
    - Comparison group matters
    - Time of comparison matters
- With multiple treatment times, it is important to assess heterogeneity
- If we suspect heterogeneity or do not want to assume homogeneity, use hdidregress and xthdidregress

# Heterogeneous treatment effects approaches

- Solution 1: Callaway and Sant'Anna (2021)
  - Use valid comparison groups
  - Split the problem into 2 by 2 comparisons
  - Compute $ATET(g, t)$, where $g$ is the cohort and $t$ is the time
- Solution 2: Wooldridge (2021)
  - (Do not blame the messenger) Include heterogeneity in your regression
  - Add the adequate interactions with cohort and time
  - Compute $ATET(g, t)$, where $g$ is the cohort and $t$ is the time
  - Computing $ATET(g)$ or $ATET(t)$ is also possible
- Other solutions exist. Good surveys are Roth et al. (2022) and de Chaisemartin and D'Haultfoeuille (forthcoming)

## Framework I

$$Y_{it} = Y_{it}(0) + \sum_{g=2}^{T} [Y_{it}(g) - Y_{it}(0)] \, G_{ig}$$

- $Y_{it}$ observed outcome
- $Y_{it}(0)$ potential outcome of not being treated
- $G_{ig}$ is an indicator for treatment group
- $g$ is the time at which a group of individuals is treated (cohort)
- $Y_{it}(g)$ potential outcome for cohort $g$

# Framework II

- Treatment is staggered
- Parallel trends
- No anticipation
- Overlap

# Callaway and Sant'Anna

- Regression adjustment (RA)
- Inverse-probability weighting (IPW)
- Augmented inverse-probability weighting (AIPW)

## RA

$$ATET(g,t) = E\left\{\frac{G_g}{E(G_g)}\left[Y_t - Y_{g-1} - m_{gt}(X)\right]\right\}$$

- $m_{gt}(X) = E(Y_t - Y_{g-1}|X, C = 1)$
- $C = 1$ is the never treated group ($G_g = \infty$)

## RA

$$ATET(g, t) = E \left\{ \frac{G_g}{E(G_g)} \left[ Y_t - Y_{g-1} - m_{gt}(X) \right] \right\}$$

- $ATET(g, t)$ is calculated using two groups: $g$ and $C = 1$, never treated
- Outcomes are computed for two points in time

## RA

$$ATET(g, t) = E\left\{ \frac{G_g}{E(G_g)} \left[ Y_t - Y_{g-1} - m_{gt}(X) \right] \right\}$$

- $ATET(g, t)$ is calculated using two groups: $g$ and $C = 1$, never treated
- Outcomes are computed for two points in time
- 2 by 2 idea
- This is done for all $g$ and all $t$
- We could have other 2 by 2 comparisons, i.e, using the not yet treated
- Identification assumptions are the same but need to hold for each 2 by 2

# Heuristically

1. keep if time is $t$ or $g-1$
2. keep if cohort is $g$ or cohort is $C=1$
3. generate $Y_t - Y_{g-1} \equiv \Delta Y$
4. regress $\Delta Y$ on $X$ for $C=1$ and predict, $\widehat{m}_{gt}(X)$
5. generate $\Delta Y - \widehat{m}_{gt}(X) = \widehat{TE}$
6. summarize $\widehat{TE}$ if cohort is $g$

- This is done for each $g$ and $t$
- Doing it for all is a GMM problem, Callaway and Sant'Anna use influence functions.

## IPW

$$ATET(g,t) = E\left\{ \left( \frac{G_g}{E\left(G_g\right)} - \frac{\frac{p_g(X)}{1-p_g(X)}}{E\left[\frac{p_g(X)}{1-p_g(X)}\right]} \right) [Y_t - Y_{g-1}] \right\}$$

- $p_g(X) = P(G_g = 1 | X, G_g + C = 1)$, i.e., conditional on the sample we keep
- Steps are similar to before with the additional computation of $\widehat{p}_g(X)$ and the quotient $\dfrac{\frac{\widehat{p}_g(X)}{1-\widehat{p}_g(X)}}{\widehat{E}\left[\frac{\widehat{p}_g(X)}{1-\widehat{p}_g(X)}\right]}$

## AIPW

$$ATET(g,t) = E\left\{ \left( \frac{G_g}{E(G_g)} - \frac{\frac{p_g(X_1)}{1-p_g(X_1)}}{E\left[\frac{p_g(X_1)}{1-p_g(X_1)}\right]} \right) [Y_t - Y_{g-1} - m_{gt}(X_2)] \right\}$$

- Notice $m_{gt}(X_2)$. Emphasizes we could have different covariates
- AIPW is doubly robust. You may incorrectly specify at least one of $m_{gt}(X_2)$ or $p_g(X_1)$ and still recover $ATET(g,t)$

## Wooldridge

$$Y_{it} = \eta + \sum_{g=q}^{T} G_{ig} \alpha_g + \sum_{s=q}^{T} f_s \alpha_s + \sum_{g=q}^{T} \sum_{s=g}^{T} D_{it} G_{ig} f_s \tau_{gs}$$

- $q$ is the first treatment time and $q \ldots T$ the post period
- $f_s$ is 1 if the time period is $s$ and 0 otherwise
- We have group and time effects $\alpha_g$ and $\alpha_s$

## Wooldridge

$$Y_{it} = \eta + \sum_{g=q}^{T} G_{ig}\alpha_g + \sum_{s=q}^{T} f_s\alpha_s + \sum_{g=q}^{T}\sum_{s=g}^{T} D_{it}G_{ig}f_s\tau_{gs}$$

- $q$ is the first treatment time and $q \ldots T$ the post period
- $f_s$ is 1 if the time period is $s$ and 0 otherwise
- We have group and time effects $\alpha_g$ and $\alpha_s$
- Heterogeneity is captured by interacting group and time effects
- $\tau_{gs} \equiv ATET(g, t)$

## Wooldridge

$$Y_{it} = \eta + \sum_{g=q}^{T} G_{ig}\alpha_g + \sum_{s=q}^{T} f_s\alpha_s + \sum_{g=q}^{T}\sum_{s=g}^{T} D_{it}G_{ig}f_s\tau_{gs}$$

- $q$ is the first treatment time and $q \ldots T$ the post period
- $f_s$ is 1 if the time period is $s$ and 0 otherwise
- We have group and time effects $\alpha_g$ and $\alpha_s$
- Heterogeneity is captured by interacting group and time effects
- $\tau_{gs} \equiv ATET(g, t)$
- We use all our data and do not partition them
- If I have covariates, they enter fully interacted

## Wooldridge

$$Y_{it} = \eta + \sum_{g=q}^{T} G_{ig}\alpha_g + \sum_{s=q}^{T} f_s\alpha_s + \sum_{g=q}^{T}\sum_{s=g}^{T} D_{it}G_{ig}f_s\tau_{gs}$$

- $q$ is the first treatment time and $q \ldots T$ the post period
- $f_s$ is 1 if the time period is $s$ and 0 otherwise
- We have group and time effects $\alpha_g$ and $\alpha_s$
- Heterogeneity is captured by interacting group and time effects
- $\tau_{gs} \equiv ATET(g, t)$
- We use all our data and do not partition them
- If I have covariates, they enter fully interacted
- Extended two-way fixed effects

# Let's see it work

```
. webuse akc, clear
(Fictional dog breed and AKC registration data)
. describe
Contains data from https://www.stata-press.com/data/r18/akc.dta
 Observations:          1,410                 Fictional dog breed and AKC registration data
    Variables:              5                 1 Feb 2023 14:23

Variable      Storage   Display    Value
    name         type    format    label      Variable label

year            int     %10.0g                Year
breed           int     %34.0g     Breed      Dog breed
movie           byte    %9.0g                 Was a movie protagonist
best            byte    %9.0g                 Won best in show in past 10 years
registered      int     %9.0g                 Number of AKC registrations

Sorted by: breed
```

## Let's see it work

```
. webuse akc, clear
(Fictional dog breed and AKC registration data)
. describe
Contains data from https://www.stata-press.com/data/r18/akc.dta
 Observations:          1,410                Fictional dog breed and AKC registration data
    Variables:              5                1 Feb 2023 14:23

Variable      Storage   Display    Value
    name         type    format    label     Variable label

year            int      %10.0g              Year
breed           int      %34.0g     Breed    Dog breed
movie           byte     %9.0g               Was a movie protagonist
best            byte     %9.0g               Won best in show in past 10 years
registered      int      %9.0g               Number of AKC registrations

Sorted by: breed
```

- Out of the 113 contests, the Terrier group has won 47 times

# Staggered treatment

```
. tabulate year movie
                Was a movie
                protagonist
       Year        0        1 |    Total
      2031       141        0 |      141
      2032       141        0 |      141
      2033       141        0 |      141
      2034       137        4 |      141
      2035       137        4 |      141
      2036       134        7 |      141
      2037       119       22 |      141
      2038       119       22 |      141
      2039       119       22 |      141
      2040       119       22 |      141
      Total     1,307      103 |    1,410
```

# Specification: Panel/Longitudinal

`xthdidregress` *estimator* `(registered best) (movie), group(breed)`

# Specification: Panel/Longitudinal

```
xthdidregress estimator (registered best) (movie), group(breed)
```

- You need to xtset with panel ID and time variable
- *estimator* is one of:
    - ra
    - twfe
    - ipw
    - aipw

## Output I

```
. xtset breed year
. xthdidregress ra (registered best) (movie), group(breed)
note: variable _did_cohort, containing cohort indicators formed by treatment variable
      movie and group variable breed, was added to the dataset.
Computing ATET for each cohort and time:
Cohort 2034 (9): ......... done
Cohort 2036 (9): ......... done
Cohort 2037 (9): ......... done
Treatment and time information
Time variable: year
Time interval: 2031 to 2040
Control:     _did_cohort = 0
Treatment:   _did_cohort > 0
```

|                  | _did_cohort |
|------------------|------------:|
| Number of cohorts |           4 |
| Number of obs    |             |
| Never treated    |        1190 |
| 2034             |          40 |
| 2036             |          30 |
| 2037             |         150 |

# Output II

```
. xthdidregress ra (registered best) (movie), group(breed)
  (output omitted)
Heterogeneous-treatment-effects regression          Number of obs   = 1,410
                                                     Number of panels =   141

Estimator:        Regression adjustment
Panel variable:   breed
Treatment level:  breed
Control group:    Never treated
                               (Std. err. adjusted for 141 clusters in breed)
```
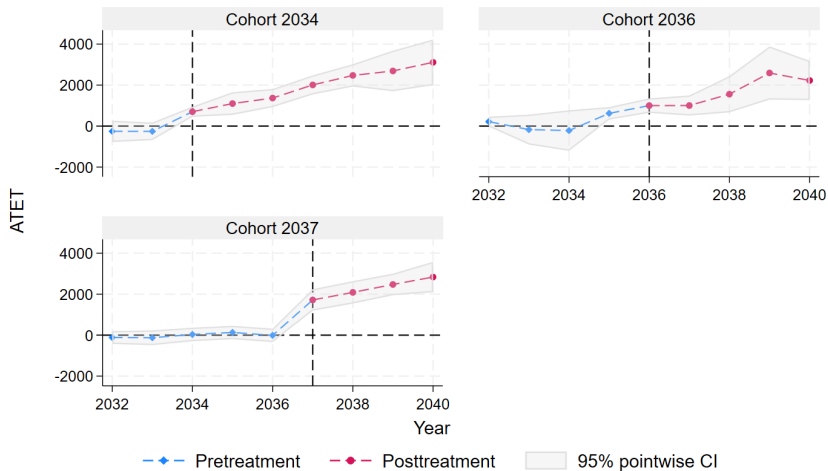
| Cohort | ATET | Robust std. err. | z | P>\|z\| | [95% conf. interval] | |
|--------|------|------------------|---|---------|------|------|
| **2034** | | | | | | |
| year | | | | | | |
| 2032 | -254.8927 | 266.1024 | -0.96 | 0.338 | -776.4439 | 266.6584 |
| 2033 | -257.5329 | 217.9389 | -1.18 | 0.237 | -684.6852 | 169.6194 |
| 2034 | 701.1318 | 127.0935 | 5.52 | 0.000 | 452.0331 | 950.2304 |
| 2035 | 1099.044 | 282.0704 | 3.90 | 0.000 | 546.196 | 1651.892 |
| 2036 | 1367.632 | 225.8702 | 6.05 | 0.000 | 924.9343 | 1810.329 |
| 2037 | 2008.294 | 237.2396 | 8.47 | 0.000 | 1543.313 | 2473.275 |
| 2038 | 2472.624 | 278.2949 | 8.88 | 0.000 | 1927.176 | 3018.072 |
| 2039 | 2689.615 | 504.3324 | 5.33 | 0.000 | 1701.142 | 3678.088 |
| 2040 | 3110.97 | 568.916 | 5.47 | 0.000 | 1995.915 | 4226.025 |

```
  (output omitted)
Note: ATET computed using covariates.
```

# Graphical representation: `estat atetplot`

## How to interpret my results

- I might want an average over all the $ATET(g, t)$
- I might want to know the effect of treatment within each group
- I might want to know the effect of treatment within a particular year
- I might want to see how the effect evolves with the duration of treatment

# Overall

```
. estat aggregation, overall
Overall ATET                                          Number of obs = 1,410
                                 (Std. err. adjusted for 141 clusters in breed)
```

| registered | ATET | Robust std. err. | z | P>\|z\| | [95% conf. interval] | |
|---|---|---|---|---|---|---|
| movie (1 vs 0) | 2093.318 | 122.5752 | 17.08 | 0.000 | 1853.075 | 2333.561 |

# Cohort Tabular

```
. estat aggregation, cohort graph sci
ATET over cohort                              Number of obs = 1,410
                                              Replications =    999
                (Std. err. adjusted for 141 clusters in breed)

                     Observed   Bootstrap         Simultaneous
         Cohort          ATET   std. err.    [95% conf. interval]

           2034       1921.33    126.9843     1570.576    2272.084
           2036      1675.093     127.749     1322.227    2027.959
           2037      2278.136    173.1852     1799.767    2756.504

Note: Simultaneous confidence intervals provide inference
      across all aggregations simultaneously.
```
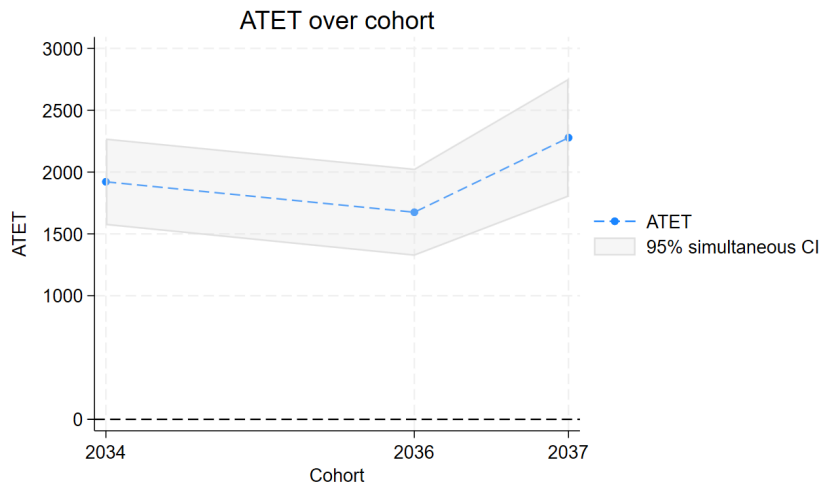
# Cohort Graphical



ATET over cohort

# Cohort list

```
. estat aggregation, cohort(2036 2034)
ATET over cohort                                      Number of obs = 1,410
                              (Std. err. adjusted for 141 clusters in breed)
```

|        |          | Robust    |       |       |          |           |
| Cohort |     ATET | std. err. |     z | P>\|z\| | [95% conf. interval] |           |
|--------|----------|-----------|-------|-------|----------|-----------|
|   2034 |  1921.33 |  187.2787 | 10.26 | 0.000 | 1554.271 |  2288.389 |
|   2036 | 1675.093 |  130.4929 | 12.84 | 0.000 | 1419.332 |  1930.855 |

# Time Tabular

```
. estat aggregation, time graph sci
ATET over time                                    Number of obs =   1,410
                                                  Replications  =     999
                   (Std. err. adjusted for 141 clusters in breed)
```

|          | Observed | Bootstrap  | Simultaneous |          |
| Time     | ATET     | std. err.  | [95% conf. interval] |          |
|----------|----------|------------|--------------|----------|
| 2034     | 701.1318 | 135.4942   | 360.2741     | 1041.989 |
| 2035     | 1099.044 | 247.3894   | 476.6955     | 1721.392 |
| 2036     | 1209.68  | 177.7654   | 762.4823     | 1656.878 |
| 2037     | 1672.655 | 206.0035   | 1154.42      | 2190.891 |
| 2038     | 2084.658 | 214.6857   | 1544.581     | 2624.735 |
| 2039     | 2528.847 | 225.004    | 1962.813     | 3094.882 |
| 2040     | 2802.171 | 308.5624   | 2025.932     | 3578.41  |

Note: Simultaneous confidence intervals provide inference
      across all aggregations simultaneously.

# Time Graphical



ATET over time
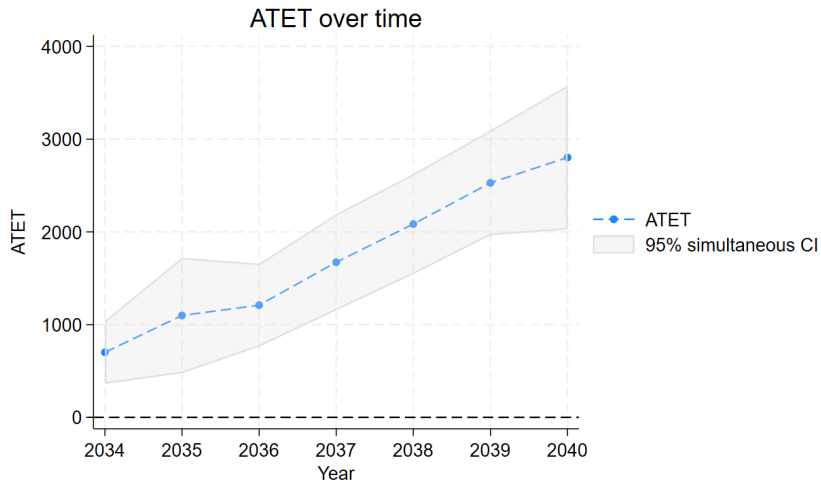
# Duration Tabular

```
. estat aggregation, dynamic graph
Duration of exposure ATET                                Number of obs = 1,410
                                (Std. err. adjusted for 141 clusters in breed)
```

| Exposure | ATET | Robust std. err. | z | P>\|z\| | [95% conf. interval] | |
|---|---|---|---|---|---|---|
| -5 | -114.582 | 160.0972 | -0.72 | 0.474 | -428.3668 | 199.2028 |
| -4 | -70.65034 | 156.3185 | -0.45 | 0.651 | -377.029 | 235.7283 |
| -3 | -.9117242 | 153.0999 | -0.01 | 0.995 | -300.982 | 299.1585 |
| -2 | 12.79653 | 144.8216 | 0.09 | 0.930 | -271.0486 | 296.6417 |
| -1 | 30.71473 | 132.8508 | 0.23 | 0.817 | -229.668 | 291.0975 |
| 0 | 1434.409 | 206.3277 | 6.95 | 0.000 | 1030.014 | 1838.804 |
| 1 | 1759.461 | 224.0229 | 7.85 | 0.000 | 1320.385 | 2198.538 |
| 2 | 2147.486 | 221.903 | 9.68 | 0.000 | 1712.564 | 2582.408 |
| 3 | 2651.452 | 284.8928 | 9.31 | 0.000 | 2093.073 | 3209.832 |
| 4 | 2366.805 | 267.4253 | 8.85 | 0.000 | 1842.661 | 2890.949 |
| 5 | 2689.615 | 504.3324 | 5.33 | 0.000 | 1701.142 | 3678.088 |
| 6 | 3110.97 | 568.916 | 5.47 | 0.000 | 1995.915 | 4226.025 |

Note: Exposure is the number of periods since the first treatment time.

# Duration Graphical
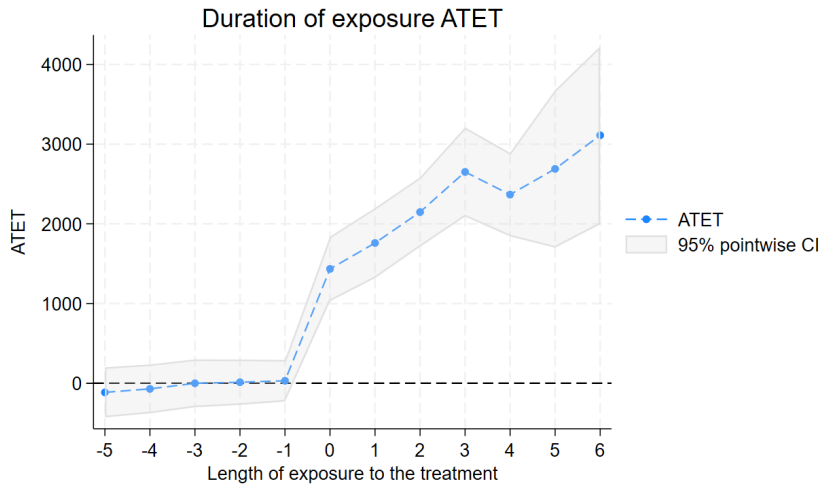


Duration of exposure ATET

## Costs

$$Y_{it} = \eta + \sum_{g=q}^{T} G_{ig}\alpha_g + \sum_{s=q}^{T} f_s\alpha_s + \sum_{g=q}^{T}\sum_{s=g}^{T} D_{it}G_{ig}f_s\tau_{gs}$$

- Number of parameters increases with $t$, $g$, and covariates
- This is true for Callaway and Sant'Anna for each 2 by 2 subset

## Costs

$$Y_{it} = \eta + \sum_{g=q}^{T} G_{ig}\alpha_g + \sum_{s=q}^{T} f_s\alpha_s + \sum_{g=q}^{T}\sum_{s=g}^{T} D_{it}G_{ig}f_s\tau_{gs}$$

- Number of parameters increases with $t$, $g$, and covariates
- This is true for Callaway and Sant'Anna for each 2 by 2 subset
- `twfe` gives us a chance to address this at estimation with `hettype()`

# Cohort heterogeneity

```
. xthdidregress twfe (registered best) (movie), group(breed) hettype(cohort)
 (output omitted )
Computing ATETs using margins ...
 (output omitted )
Heterogeneous-treatment-effects regression          Number of obs   = 1,410
                                                     Number of panels =   141

Estimator:       Two-way fixed effects
Panel variable:  breed
Treatment level: breed
Control group:   Never treated
Heterogeneity:   Cohort
                              (Std. err. adjusted for 141 clusters in breed)
```

| Cohort | ATET | Robust std. err. | t | P>|t| | [95% conf. interval] | |
|---|---|---|---|---|---|---|
| 2034 | 1662.492 | 108.002 | 15.39 | 0.000 | 1448.966 | 1876.017 |
| 2036 | 1978.645 | 54.21043 | 36.50 | 0.000 | 1871.468 | 2085.822 |
| 2037 | 2276.223 | 70.63244 | 32.23 | 0.000 | 2136.579 | 2415.867 |

Note: ATET computed using covariates.

# Time heterogeneity

```
. xthdidregress twfe (registered best) (movie), group(breed) hettype(time)
 (output omitted)
Computing ATETs using margins ...
 (output omitted)
Heterogeneous-treatment-effects regression        Number of obs   = 1,410
                                                   Number of panels =   141

Estimator:       Two-way fixed effects
Panel variable:  breed
Treatment level: breed
Control group:   Never treated
Heterogeneity:   Time
                               (Std. err. adjusted for 141 clusters in breed)
```

|       Time |     ATET | Robust std. err. |     t | P>\|t\| | [95% conf. | interval] |
|-----------:|---------:|-----------------:|------:|--------:|-----------:|----------:|
|       year |          |                  |       |         |            |           |
|       2034 | 368.7667 |         177.8305 |  2.07 |   0.040 |   17.18625 |  720.3472 |
|       2035 | 825.1479 |         269.4845 |  3.06 |   0.003 |   292.3625 |  1357.933 |
|       2036 | 1262.367 |         85.71962 | 14.73 |   0.000 |   1092.895 |  1431.839 |
|       2037 | 1692.904 |          165.318 | 10.24 |   0.000 |   1366.061 |  2019.746 |
|       2038 | 2110.763 |         162.5539 | 12.99 |   0.000 |   1789.386 |  2432.141 |
|       2039 | 2554.188 |         241.3463 | 10.58 |   0.000 |   2077.033 |  3031.342 |
|       2040 |  2829.94 |         243.2751 | 11.63 |   0.000 |   2348.972 |  3310.908 |

Note: ATET computed using covariates.

## Parting thoughts

- Continued interest and development in causal inference/treatment effects
- Heterogeneous DID
- Estimation and postestimation tools displayed in tabular and graphical forms
- Different ways of thinking about heterogeneity