Difference-in-Differences for Nonbinary Treatments in Stata

2025 Stata Conference
Nashville, TN
31 July - 1 August, 2025

Jeff Wooldridge
Department of Economics
Michigan State University

1. Review of the Binary Case
2. Nonbinary Treatments with $T = 2$
3. A Heterogeneous Slopes Model with Staggered Assignment
4. Application to Walmart Store Sitings
5. Summary

1. Review of the Binary Case

- For staggered interventions with a binary treatment, many estimators now exist.
- The Stata 18+ command `xthdidregress` supports:
  - `twfe`: Extended TWFE, a “lags only” estimator: Wooldridge (2025, forthcoming Empirical Economics).
  - `ra`: The “leads and lags” (event study) regression adjustment estimator: Callaway and Sant’Anna (2021, Journal of Econometrics).
    - ▹ Same as ETWFE with all leads and lags and full flexibility in controls [Wooldridge (2025)].
    - ▹ Same as Sun and Abraham (2021, J of E) without controls.

- `ipw`: CS (2021) inverse probability weighting; Abadie (2005, REStat), Sant'Anna and Zhou (2020, J of E) as special cases.
- `aipw`: CS (2021) augmented IPW – doubly robust.
- Estimates can be aggregrated and graphed by exposure time.

- `jwdid` [Rios-Avila, Nagengast, Yotov] produces ETWFE with moderating effects and heterogeneous trends.
- `jwdid` with the `never` option is the same as `xthdidregress ra` with a balanced panel, time-constant controls.
- `csdid` [Rios-Avila, Callaway, Sant'Anna] reproduces `ra`, `ipw`, `aipw`.
- Get more information doing estimation "by hand": `reg`, `xtreg`, `teffects` (after transforming the data).
- `reg`, `xtreg` allow inclusion of cohort-specific trends.

- Lee and Wooldridge (2023, Working Paper) show how one can remove pre-treatment averages or pre-treatment (unit-specific) trends to obtain lags only estimators.
  - One can apply any TE estimator: IPWRA, matching, machine learning on top of the others.
  - Can use “long differencing” instead, as in CS (2021).
  - In both cases, `teffects` can be applied after simple transformation.
  - Can apply causal machine learning methods, too.

- What can we do with non-binary treatments?
  - Minimum wages deviate from a national minimum wage.
  - New tax rates are imposed at the county level.
  - Distance to a natural disaster or new garbage incinerator.
  - Different levels of participation in a training program.
- Here I focus on time-invariant controls, balanced panel data.
  - Unbalanced panels easy to handle.
- Spoiler: I suggest a flexible equation estimated by TWFE.

Setup and Results for Binary Treatment

- Summary of Wooldridge (2021, 2025).
- Treatment cohorts $g \in \{q, q+1, \ldots, T\}$.
- Potential outcomes $Y_t(g)$, with $Y_t(\infty)$ in the never treated state.
- Cohort dummy variables (time constant): D_g,

$g \in \{q, q+1, \ldots, T, \infty\}$.

  - For unit i, D_{ig}.

- Parameters of interest are ATTs (or ATETs) for each cohort g:

$$\tau_{gt} = E[Y_t(g) - Y_t(\infty) | D_g = 1], t = g, \ldots, T.$$

Assumption NBC (**No Bad Controls**): For time-constant covariates $\mathbf{X}(g)$,

$$\mathbf{X} = \mathbf{X}(g) = \mathbf{X}(\infty). \; \square$$

Assumption CNA (**Conditional No Anticipation**): For each treatment cohort $g \in \{q, \ldots, T\}$,

$$E[Y_t(g)|D_q, \ldots, D_T, \mathbf{X}] = E[Y_t(\infty)|D_q, \ldots, D_T, \mathbf{X}], \; t < g. \; \square$$

- No conditional pre-treatment effects.

Assumption CPT (Conditional Parallel Trends): For $t = 2,\ldots,T$ and time-constant controls $\mathbf{X}$,

$$E[Y_t(\infty) - Y_1(\infty)|D_q,\ldots,D_T,\mathbf{X}] = E[Y_t(\infty) - Y_1(\infty)|\mathbf{X}]. \ \square$$

- Let $\mathbf{D} \equiv (D_q,\ldots,D_T)$ be the vector of cohort ("treatment") indicators.
- Assumption CPT implies that, conditional on $\mathbf{X}$, $\mathbf{D}$ is unconfounded with respect to the trends

$$Y_t(\infty) - Y_1(\infty), t = 2,\ldots,T$$

- $\mathbf{D}$ is allowed to be confounded with $Y_1(\infty)$ even conditional on $\mathbf{X}$.

Assumption LIN (Linearity): For treatment cohort indicators D_g and control variables $\mathbf{X}$,

$$E[Y_1(\infty)|\mathbf{D},\mathbf{X}] = \alpha + \sum_{g=q}^{T} \beta_g D_g + \mathbf{X}\boldsymbol{\kappa} + \sum_{g=q}^{T} (D_g \cdot \mathbf{X})\boldsymbol{\xi}_g$$

$$E[Y_t(\infty)|\mathbf{D},\mathbf{X}] - E[Y_1(\infty)|\mathbf{D},\mathbf{X}] = \sum_{s=2}^{T} \gamma_s fs_t + \sum_{s=2}^{T} (fs_t \cdot \mathbf{X})\boldsymbol{\pi}_s, \; t = 2,\ldots,T. \; \square$$

- The second part implies CPT.

- For a random draw i, with time-varying treatment indicator $W_{it} \in \{0, 1\}$:

$$\begin{aligned} E(Y_{it}|D_{iq}, \ldots, D_{iT}, \mathbf{X}_i) = & \sum_{g=q}^{T} \sum_{s=g}^{T} \tau_{gs}(W_{it} \cdot D_{ig} \cdot fs_t) \\ & + \sum_{g=q}^{T} \sum_{s=g}^{T} (W_{it} \cdot D_{ig} \cdot fs_t \cdot \dot{\mathbf{X}}_{ig})\boldsymbol{\rho}_{gs} \\ & + \sum_{s=2}^{T} \gamma_s fs_t + \sum_{s=2}^{T} (fs_t \cdot \mathbf{X}_i)\boldsymbol{\pi}_s \\ & + \alpha + \sum_{g=q}^{T} \beta_g D_{ig} + \mathbf{X}_i \boldsymbol{\kappa} + \sum_{g=q}^{T} (D_{ig} \cdot \dot{\mathbf{X}}_{ig})\boldsymbol{\xi}_g \end{aligned}$$

- Obvious estimator is pooled OLS controlling for time dummies, cohort dummies, and lots of interactions.
- Many equivalent estimators:

POLS = RE = FE = Cohort Imputation = BJS Imputation

  - Imputation methods use an initial regression with $W_{it} = 0$ observations.
- No "bad comparisons" unless we impose restrictions on the τ_{gt}.
  - For example, $\tau_{gt} = \tau$, all g and t.

- POLS (and RE) give the most information: moderating effects, heterogeneous trends, selection effects.
  - `xthdidregress twfe` gives ATTs.
  - `jwdid` gives ATTs, moderating effects, heterogeneous trends.
- Can include cohort-specific (linear) trends; same as imputation with trends, same as unit-specific trends.

2. Non-Binary Treatments with $T = 2$

- Assume initially that the treament has quantitative meaning ("continuous").
  - $D = 0$ is control; $D > 0$ is treatment.
  - Treated unit i goes from zero to a positive amount:

$$
\begin{aligned}
W_{i1} &= 0 \;\; \text{(no treatment in first period)} \\
W_{i2} &= D_i \; \text{(some treated units)}
\end{aligned}
$$

- Define the dose-response function in $t = 2$:

$$TE(d) = Y_2(d) - Y_2(0), d > 0$$

- Callaway, Goodman-Bacon, Sant'Anna (2024, WP) define the average D-R function:

$$ATE(d) = E[Y_2(d) - Y_2(0)]$$

  - Does not reduce to the usual ATT with binary treatment.

- Instead, define the dose-response function for the treated:

$$ATT(d) = E[Y_2(d) - Y_2(0)|D > 0]$$

- Consider a linear model with random slope, B_2:

$$Y_2(d) = A + \gamma_2 + B_2 d + U_2,\ E(U_2) = 0$$

$$ATT(d) = E(B_2|D > 0) \cdot d = \delta_2 \cdot d$$
$$\delta_2 \equiv E(B_2|D > 0)$$

- δ_2 is identified under an assumption that limits selection:

$$E(B_2|D) = E(B_2|D > 0)$$

- The CG-BS'A strong PT assumption is no selection:

$$E(B_2|D) = E(B_2) = \beta_2$$

- Then

$$E(Y_2 - Y_1|D) = \gamma_2 + \beta_2 D$$

- Identical to applying TWFE to

$$Y_{it} = \gamma_2 f2_t + \beta_2(f2_t \cdot D_i) + C_i + U_{it}, t = 1,2$$

- TWFE identifies δ_2 under the weaker selection assumption.

- Can apply TWFE to more flexible models:

$$
\begin{aligned}
Y_{it} = {} & \delta_{21}(f2_t \cdot D_i) + \delta_{22}(f2_t \cdot D_i^2) \\
& + \delta_{23}(f2_t \cdot D_i^3) + \gamma_2 f2_t + C_i + U_{it}
\end{aligned}
$$

  - Same as

$$\Delta Y_i \text{ on } 1, D_i, D_i^2, D_i^3, i = 1, \ldots, N$$

- Currently, the fancier methods are nonparametric versions of simple TWFE based on first differencing.

3. A Heterogeneous Slopes Model with Staggered Assignment

- Borrows from Wooldridge (2005, REStat).
- Let $\mathbf{W}_{it}$ be a row vector of treatment variables; could be continuous or discrete.
  - Could include functions of underlying treatment variables.
  - Can vary over time.
- Return to notation where D_{ig} denotes cohort indicators.
  - Assumes a control periods and a well-defined first period of treatment.

- Think of a potential outcomes (dose-response) function:

$$Y_{it}(\mathbf{w}) = \mathbf{w}\mathbf{B}_{it} + \gamma_t + C_i + U_{it}(\mathbf{w})$$

  - $\mathbf{B}_{it}$ a vector of heterogenous (random) coefficients; vary by i and t.

- Assume strictly exogenous treatment with respect to shocks:

$$E[U_{it}(\mathbf{w})|\mathbf{W}_{i1},\ldots,\mathbf{W}_{iT},C_i,\mathbf{B}_{i1},\ldots,\mathbf{B}_{IT}] = 0$$

  - No feedback from shocks to Y_{it} to future treatment.

- Define the cohort/time dose response function:

$$E[Y_{it}(\mathbf{w})|D_{iq},\ldots,D_{iT}] = \mathbf{w}E(\mathbf{B}_{it}|\mathbf{D}_i) + \gamma_t + E(C_i|\mathbf{D}_i)$$

  - Can only identify this function for (g,t) pairs with $t \in \{g,\ldots,T\}$.

- The following representation imposes no assumption when treatments are zero in $t \in \{1,\ldots,g-1\}$ for cohort g:

$$E(\mathbf{B}_{it}|D_{iq},\ldots,D_{iT}) = \sum_{g=q}^{T}\sum_{s=g}^{T}(D_{ig} \cdot fs_t)\boldsymbol{\delta}_{gs}$$

- In terms of observable Y_{it}:

$$Y_{it} = Y_{it}(\mathbf{W}_{it}) = \mathbf{W}_{it}\mathbf{B}_{it} + \gamma_t + C_i + U_{it}$$

- Treatment assignment can be arbitrarily correlated with C_i.
- As usual, allow additive time effects.
  - Unrestricted trend in the mean of the untreated state.
  - Additivity in γ_t and A_i imposes a PT assumption.

- Key restriction for identification:

$$
\begin{aligned}
E(\mathbf{B}_{it}|\mathbf{W}_{i1},\ldots,\mathbf{W}_{iT}) &= E(\mathbf{B}_{it}|D_{iq},\ldots,D_{iT}) \\
&= \sum_{g=q}^{T}\sum_{s=g}^{T}(D_{ig}\cdot fs_t)\boldsymbol{\delta}_{gs} + \sum_{s=1}^{T}(D_{i\infty}\cdot fs_t)\boldsymbol{\delta}_{\infty s}
\end{aligned}
$$

  - The *timing* of treatment can be related to the per-unit gain, $\mathbf{B}_{it}$.
  - The *level* of treatment cannot be related to the per-unit gain.
  - Imposes no extra assumptions in the binary case.

• If $\mathbf{W}_{it} \cdot D_{i\infty} \cdot fs_t = \mathbf{0}$ – treatment is zero for the NT cohort – we can write

$$Y_{it} = \sum_{g=q}^{T}\sum_{s=g}^{T}(\mathbf{W}_{it} \cdot D_{ig} \cdot fs_t)\boldsymbol{\delta}_{gs} + \sum_{s=2}^{T}\gamma_s fs_t + A_i + U_{it}$$

$$E(U_{it}|\mathbf{D}_i, \mathbf{W}_{i1}, \ldots, \mathbf{W}_{iT}, A_i) = 0$$

• Consistently estimate the $\boldsymbol{\delta}_{gs}$ by TWFE.

• Contains the binary case when

$$W_{it} = D_{iq} \cdot pq_t + \cdots + D_{iT} \cdot pT_t$$
$$pg_t = fg_t + \cdots + fT_t$$

• Might aggregate the coefficients by exposure time, or obtain a single weighted $\boldsymbol{\delta}$:

$$\boldsymbol{\delta} = \sum_{g=q}^{T} \sum_{t=g}^{T} \omega_{gt} \boldsymbol{\delta}_{gt}$$

Examples:

1. W_{it} a scalar: continuous, count, corner (zero before intervention).
2. $\mathbf{W}_{it} = (R_{it}, R_{it}^2, R_{it}^3)$ for some quantitative treatment R_{it}.
3. $\mathbf{W}_{it} = (W_{it1}, W_{it2}, \ldots, W_{itJ})$ (different treatment levels).
4. $\mathbf{W}_{it} = (W_{it1}, W_{it2}, \ldots, W_{itJ})$ (different treatments).

- Add controls, $\dot{\mathbf{X}}_{ig} \equiv \mathbf{X}_i - \bar{\mathbf{X}}_g$:

$$
\begin{aligned}
Y_{it} = & \sum_{g=q}^{T} \sum_{s=g}^{T} (\mathbf{W}_{it} \cdot D_{ig} \cdot fs_t) \boldsymbol{\delta}_{gs} \\
& + \sum_{g=q}^{T} \sum_{s=g}^{T} (\mathbf{W}_{it} \otimes D_{ig} \cdot fs_t \cdot \dot{\mathbf{X}}_{ig}) \boldsymbol{\rho}_{gs} \\
& \sum_{s=2}^{T} \gamma_s fs_t + \sum_{s=2}^{T} (fs_t \cdot \mathbf{X}_i) \boldsymbol{\pi}_s + C_i + U_{it}
\end{aligned}
$$

- Allows heterogeneity in the “dose-response” function to vary by cohort and time.
- In the binary W_{it} case, reduces to Wooldridge (2021, 2025).

- Allows for the heterogeneous trends, $fs_t \cdot \mathbf{X}_i$ – relaxes PT.
- Conditioning on $\mathbf{X}_i$ may also help with selection on the return to treatment:

$$E(\mathbf{B}_{it}|\mathbf{W}_{i1},\ldots,\mathbf{W}_{iT},\mathbf{X}_i) = E(\mathbf{B}_{it}|D_{iq},\ldots,D_{iT},\mathbf{X}_i)$$

- The additive effects, C_i, may be arbitrarily correlated with the $\mathbf{W}_{it}$.
  - The usual additive FE selection is allowed.

- $\mathbf{W}_{it}$ can be a vector of treatment indicators of units whose treatment may spill over into the outcome for unit i.
- If unit i's neighbors do not change over time, TWFE estimation of the flexible equation eliminates not just C_i, but $C_{j(i)}$ for all neighbors $j(i)$ of unit i.
- To be completely general, should allow $W_{j(i)t}$ to interact with $D_{j(i),g(j(i))} \cdot fs_t$, where $g(j(i))$ is the treament cohort for unit $j(i)$.

Checking and Correcting for Violation of PT

- Add pre-treatment indicators $D_{ig} \cdot fs_t$ for $s \leq g - 2$ to the flexible equation.
  - Event study estimation with binary W_{it}.
  - Check for significance of these "lead" indicators.
- Can add cohort-specific linear trends, $D_{ig} \cdot t$.
  - Wooldridge (2025) in binary case.
  - Allows additional selection into treatment.
- Test for violation of strict exogeneity by including $\mathbf{W}_{i,t+1}$ in TWFE.

Non-Zero Treatment Level for Controls

- Suppose the "treatment" is never zero for any unit.
  - ▸ For example, distance from a house to a new garbarge incinerator.
- Define $W_{it} = 0$ in the periods before the intervention.
- W_{it} is the treatment variable post-intervention; may never be zero.
  - ▸ The stronger assumption on selection is needed because all units are "treated" at some level.

4. Application to Number of Walmart Stores on Retail Employment

- The treatment is the number of Walmarts open at a particular point in time.
- Can allow heterogeneous effects by exposure time.
- Can allow unrestricted heterogeneity by cohort/calendar time.
- Can also include heterogeneous trends.
- Equations only change from binary case by interacting the treatment dummies with the number of stores.
- See walmart_lw_retail_n_open_slides.do.

```
. use walmart_lw, clear

. bysort year: tab n_open

-> year = 1986

     n_open |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      1,219       94.64       94.64
          1 |         62        4.81       99.46
          2 |          6        0.47       99.92
          4 |          1        0.08      100.00
------------+-----------------------------------
      Total |      1,288      100.00

-------------------------------------------------------------------------------------------
-> year = 1987

     n_open |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |      1,145       88.90       88.90
          1 |        125        9.70       98.60
          2 |         17        1.32       99.92
          5 |          1        0.08      100.00
------------+-----------------------------------
      Total |      1,288      100.00
```

...

-> year = 1999

n_open	Freq.	Percent	Cum.
0	395	30.67	30.67
1	664	51.55	82.22
2	137	10.64	92.86
3	47	3.65	96.51
4	24	1.86	98.37
5	9	0.70	99.07
6	1	0.08	99.15
7	3	0.23	99.38
8	2	0.16	99.53
9	1	0.08	99.61
10	1	0.08	99.69
11	2	0.16	99.84
12	1	0.08	99.92
13	1	0.08	100.00
Total	1,288	100.00	

```
. gen w = n_open

. * Each new store is estimated to increase retail employment by about 3.4%.

. xtreg log_retail_emp w i.year, fe vce(cluster fips)

Fixed-effects (within) regression               Number of obs     =     29,624
Group variable: fips                            Number of groups  =      1,288

R-squared:                                      Obs per group:
     Within  = 0.5100                                         min =         23
     Between = 0.2818                                         avg =       23.0
     Overall = 0.0255                                         max =         23

                                                F(23, 1287)       =     236.15
corr(u_i, Xb) = 0.0573                          Prob > F          =     0.0000

                               (Std. err. adjusted for 1,288 clusters in fips)
------------------------------------------------------------------------------
             |               Robust
log_retail~p | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
           w |    .034412   .0056512     6.09   0.000     .0233255    .0454986
```

year						
1978	.0730682	.0019088	38.28	0.000	.0693234	.076813
1979	.1090979	.0025137	43.40	0.000	.1041666	.1140292
1980	.0852645	.0028303	30.13	0.000	.0797121	.0908169
1981	.0708761	.003318	21.36	0.000	.0643668	.0773853
1982	.0733567	.0039143	18.74	0.000	.0656776	.0810359
1983	.0476426	.0041387	11.51	0.000	.0395233	.055762
1984	.0904116	.0045211	20.00	0.000	.081542	.0992811
1985	.1263614	.0049682	25.43	0.000	.1166148	.136108
1986	.1502841	.005596	26.86	0.000	.1393057	.1612625
1987	.1953998	.0061582	31.73	0.000	.1833186	.2074809
1988	.2301552	.0066879	34.41	0.000	.2170348	.2432757
1989	.2576538	.0070179	36.71	0.000	.2438861	.2714215
1990	.2850406	.0073995	38.52	0.000	.2705243	.2995569
1991	.2735001	.0076954	35.54	0.000	.2584032	.288597
1992	.2674434	.0079728	33.54	0.000	.2518022	.2830846
1993	.2722996	.0085955	31.68	0.000	.255437	.2891623
1994	.3012396	.0089872	33.52	0.000	.2836086	.3188707
1995	.331336	.0094764	34.96	0.000	.3127452	.3499269
1996	.3449819	.0099649	34.62	0.000	.3254325	.3645312
1997	.3711286	.0103622	35.82	0.000	.3507999	.3914573
1998	.3710202	.0106592	34.81	0.000	.3501089	.3919315
1999	.3878193	.010936	35.46	0.000	.3663648	.4092737
_cons	7.547977	.0051162	1475.32	0.000	7.53794	7.558014

```
. xtreg log_retail_emp c.exp0#c.w c.exp1#c.w c.exp2#c.w c.exp3#c.w c.exp4#c.w ///
>         c.exp5#c.w c.exp6#c.w c.exp7#c.w c.exp8#c.w c.exp9#c.w c.exp10#c.w ///
>         c.exp11#c.w c.exp12#c.w c.exp13#c.w i.year, fe vce(cluster fips)

------------------------------------------------------------------------------
             |               Robust
log_retail~p | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
  c.exp0#c.w |   .0260239   .0043407     6.00   0.000     .0175082    .0345395
             |
  c.exp1#c.w |   .0405417   .0044418     9.13   0.000     .0318278    .0492557
             |
  c.exp2#c.w |   .0327549   .0045172     7.25   0.000      .023893    .0416168
             |
  c.exp3#c.w |   .0281223   .0047973     5.86   0.000     .0187109    .0375337
             |
  c.exp4#c.w |    .026921    .005103     5.28   0.000     .0169097    .0369322
             |
  c.exp5#c.w |   .0294158   .0054844     5.36   0.000     .0186564    .0401751
             |
  c.exp6#c.w |   .0316174   .0060162     5.26   0.000     .0198147    .0434201
             |
  c.exp7#c.w |   .0371473   .0070739     5.25   0.000     .0232697     .051025
             |
  c.exp8#c.w |   .0406383   .0083609     4.86   0.000     .0242358    .0570407
```

c.exp9#c.w	.0418278	.0094813	4.41	0.000	.0232273	.0604283
c.exp10#c.w	.0593063	.0098245	6.04	0.000	.0400324	.0785801
c.exp11#c.w	.0658462	.0105913	6.22	0.000	.045068	.0866243
c.exp12#c.w	.0734672	.0144296	5.09	0.000	.0451591	.1017752
c.exp13#c.w	.087957	.0160304	5.49	0.000	.0565085	.1194056

year						
1978	.0730682	.0019093	38.27	0.000	.0693226	.0768138
1979	.1090979	.0025142	43.39	0.000	.1041655	.1140303
1980	.0852645	.0028309	30.12	0.000	.0797108	.0908181
1981	.0708761	.0033187	21.36	0.000	.0643654	.0773868
1982	.0733567	.0039152	18.74	0.000	.0656759	.0810376
1983	.0476426	.0041396	11.51	0.000	.0395215	.0557638
1984	.0904116	.0045221	19.99	0.000	.0815401	.0992831
1985	.1263614	.0049693	25.43	0.000	.1166126	.1361102
1986	.1507921	.0056007	26.92	0.000	.1398045	.1617797
1987	.1955097	.0061657	31.71	0.000	.1834139	.2076056
1988	.2302807	.00669	34.42	0.000	.2171563	.2434052
1989	.2584009	.0069906	36.96	0.000	.2446867	.2721151
1990	.2865656	.0073255	39.12	0.000	.2721945	.3009368
1991	.2750423	.0075585	36.39	0.000	.260214	.2898706
1992	.269186	.0078107	34.46	0.000	.2538628	.2845091
1993	.2743704	.0084198	32.59	0.000	.2578524	.2908883
1994	.3027812	.0088957	34.04	0.000	.2853295	.3202329
1995	.3328265	.0094779	35.12	0.000	.3142327	.3514204
1996	.3442326	.0100762	34.16	0.000	.324465	.3640002
1997	.3675389	.0105959	34.69	0.000	.3467517	.388326
1998	.3635317	.0111372	32.64	0.000	.3416827	.3853807
1999	.3750971	.011751	31.92	0.000	.3520439	.3981502
_cons	7.547977	.0051117	1476.62	0.000	7.537949	7.558005

```
. * Estimate separate effects by cohort/year, and then weight the coefficients.
. * Similar to imposing constant effects by exposure time.

. qui xtreg log_retail_emp c.w#c.d1986#(c.f1986 c.f1987 c.f1988 c.f1989 c.f1990 c.f1991 /
>         c.f1992 c.f1993 c.f1994 c.f1995 c.f1996 c.f1997 c.f1998 c.f1999) ///
>         c.w#c.d1987#(c.f1987 c.f1988 c.f1989 c.f1990 c.f1991 ///
>         c.f1992 c.f1993 c.f1994 c.f1995 c.f1996 c.f1997 c.f1998 c.f1999) ///
>         c.w#c.d1988#(c.f1988 c.f1989 c.f1990 c.f1991 ///
>         c.f1992 c.f1993 c.f1994 c.f1995 c.f1996 c.f1997 c.f1998 c.f1999) ///
>         c.w#c.d1989#(c.f1989 c.f1990 c.f1991 ///
>         c.f1992 c.f1993 c.f1994 c.f1995 c.f1996 c.f1997 c.f1998 c.f1999) ///
>         c.w#c.d1990#(c.f1990 c.f1991 c.f1992 c.f1993 c.f1994 c.f1995 c.f1996 ///
>         c.f1997 c.f1998 c.f1999) ///
>         c.w#c.d1991#(c.f1991 c.f1992 c.f1993 c.f1994 c.f1995 c.f1996 c.f1997 c.f1998 c.
>         c.w#c.d1992#(c.f1992 c.f1993 c.f1994 c.f1995 c.f1996 c.f1997 c.f1998 c.f1999) /
>         c.w#c.d1993#(c.f1993 c.f1994 c.f1995 c.f1996 c.f1997 c.f1998 c.f1999) ///
>         c.w#c.d1994#(c.f1994 c.f1995 c.f1996 c.f1997 c.f1998 c.f1999) ///
>         c.w#c.d1995#(c.f1995 c.f1996 c.f1997 c.f1998 c.f1999) ///
>         c.w#c.d1996#(c.f1996 c.f1997 c.f1998 c.f1999) ///
>         c.w#c.d1997#(c.f1997 c.f1998 c.f1999) ///
>         c.w#c.d1998#(c.f1998 c.f1999) ///
>         c.w#c.d1999#c.f1999 i.year, fe vce(cluster fips)
```

```
. lincom omega1986_0*c.w#c.d1986#c.f1986 + omega1987_0*c.w#c.d1987#c.f1987 ///
>         + omega1988_0*c.w#c.d1988#c.f1988 + omega1989_0*c.w#c.d1989#c.f1989 ///
>         + omega1990_0*c.w#c.d1990#c.f1990 + omega1991_0*c.w#c.d1991#c.f1991 ///
>         + omega1992_0*c.w#c.d1992#c.f1992 + omega1993_0*c.w#c.d1993#c.f1993 ///
>         + omega1994_0*c.w#c.d1994#c.f1994 + omega1995_0*c.w#c.d1995#c.f1995 ///
>         + omega1996_0*c.w#c.d1996#c.f1996 + omega1997_0*c.w#c.d1997#c.f1997 ///
>         + omega1998_0*c.w#c.d1998#c.f1998 + omega1999_0*c.w#c.d1999#c.f1999

------------------------------------------------------------------------------
log_retail~p | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0290702   .0040876     7.11   0.000     .0210512    .0370893
------------------------------------------------------------------------------

. lincom omega1986_1*c.w#c.d1986#c.f1987 + omega1987_1*c.w#c.d1987#c.f1988 ///
>         + omega1988_1*c.w#c.d1988#c.f1989 + omega1989_1*c.w#c.d1989#c.f1990 ///
>         + omega1990_1*c.w#c.d1990#c.f1991 + omega1991_1*c.w#c.d1991#c.f1992 ///
>         + omega1992_1*c.w#c.d1992#c.f1993 + omega1993_1*c.w#c.d1993#c.f1994 ///
>         + omega1994_1*c.w#c.d1994#c.f1995 + omega1995_1*c.w#c.d1995#c.f1996 ///
>         + omega1996_1*c.w#c.d1996#c.f1997 + omega1997_1*c.w#c.d1997#c.f1998 ///
>         + omega1998_1*c.w#c.d1998#c.f1999

------------------------------------------------------------------------------
log_retail~p | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0438954   .0042264    10.39   0.000     .0356041    .0521867
------------------------------------------------------------------------------
```

```
. lincom omega1986_11*c.w#c.d1986#c.f1997 + omega1987_11*c.w#c.d1987#c.f1998 ///
>         + omega1988_11*c.w#c.d1988#c.f1999

------------------------------------------------------------------------------
log_retail~p | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0670685   .0120642     5.56   0.000     .0434008    .0907361
------------------------------------------------------------------------------

. lincom omega1986_12*c.w#c.d1986#c.f1998 + omega1987_12*c.w#c.d1987#c.f1999

------------------------------------------------------------------------------
log_retail~p | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0705872   .0160293     4.40   0.000     .0391408    .1020336
------------------------------------------------------------------------------

. * One coefficient for 13 years exposure:
.
. lincom c.w#c.d1986#c.f1999

------------------------------------------------------------------------------
log_retail~p | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0894949   .0209551     4.27   0.000     .0483851    .1306048
------------------------------------------------------------------------------
```

```
. * Heterogeneous trends by cohort. Effects are uniformly smaller:

. xtreg log_retail_emp c.exp0#c.w c.exp1#c.w c.exp2#c.w c.exp3#c.w c.exp4#c.w ///
>         c.exp5#c.w c.exp6#c.w c.exp7#c.w c.exp8#c.w c.exp9#c.w c.exp10#c.w ///
>         c.exp11#c.w c.exp12#c.w c.exp13#c.w ///
>         c.year#(c.d1986 c.d1987 c.d1988 c.d1989 c.d1990 c.d1991 c.d1992 c.d1993 ///
>         c.d1994 c.d1995 c.d1996 c.d1997 c.d1998 c.d1999) i.year, fe vce(cluster fips)

--------------------------------------------------------------------------------
               |               Robust
log_retail_emp | Coefficient  std. err.      t    P>|t|     [95% conf. interval]
---------------+----------------------------------------------------------------
    c.exp0#c.w |   .0034236   .0036042     0.95   0.342    -.0036471    .0104942
               |
    c.exp1#c.w |   .0184111   .0037281     4.94   0.000     .0110973    .0257248
               |
    c.exp2#c.w |   .0112132   .0037578     2.98   0.003     .0038411    .0185853
               |
    c.exp3#c.w |   .0069585    .003944     1.76   0.078    -.0007788    .0146959
               |
    c.exp4#c.w |   .0059511   .0042638     1.40   0.163    -.0024137    .0143158
               |
    c.exp5#c.w |   .0065552   .0046092     1.42   0.155    -.0024873    .0155976
               |
    c.exp6#c.w |   .0084336    .005066     1.66   0.096    -.0015049     .018372
               |
    c.exp7#c.w |   .0129591   .0058303     2.22   0.026     .0015212     .024397
```

c.exp8#c.w	.0155207	.0068423	2.27	0.023	.0020975	.0289439
c.exp9#c.w	.0166873	.0079335	2.10	0.036	.0011233	.0322513
c.exp10#c.w	.0250665	.0080798	3.10	0.002	.0092155	.0409175
c.exp11#c.w	.0280936	.0095083	2.95	0.003	.0094402	.0467471
c.exp12#c.w	.0404849	.0120386	3.36	0.001	.0168675	.0641023
c.exp13#c.w	.0527214	.0137155	3.84	0.000	.0258142	.0796286

c.year#c.d1986	.008707	.0026515	3.28	0.001	.0035052	.0139088
c.year#c.d1987	.007829	.0023175	3.38	0.001	.0032825	.0123755
c.year#c.d1988	.0121625	.002025	6.01	0.000	.0081899	.0161351
c.year#c.d1989	.0096362	.0017829	5.40	0.000	.0061384	.013134
c.year#c.d1990	.0085395	.0014877	5.74	0.000	.0056208	.0114581
c.year#c.d1991	.0097681	.0016158	6.05	0.000	.0065983	.0129379
c.year#c.d1992	.0095503	.0017016	5.61	0.000	.006212	.0128885
c.year#c.d1993	.0087822	.0016119	5.45	0.000	.00562	.0119444
c.year#c.d1994	.0103002	.0018614	5.53	0.000	.0066486	.0139518
c.year#c.d1995	.0063531	.0017904	3.55	0.000	.0028406	.0098656
c.year#c.d1996	.0135357	.0033423	4.05	0.000	.0069789	.0200926
c.year#c.d1997	.0133329	.0030466	4.38	0.000	.0073561	.0193097
c.year#c.d1998	.0096277	.0031339	3.07	0.002	.0034795	.0157759
c.year#c.d1999	.0096849	.0036959	2.62	0.009	.0024343	.0169356

...

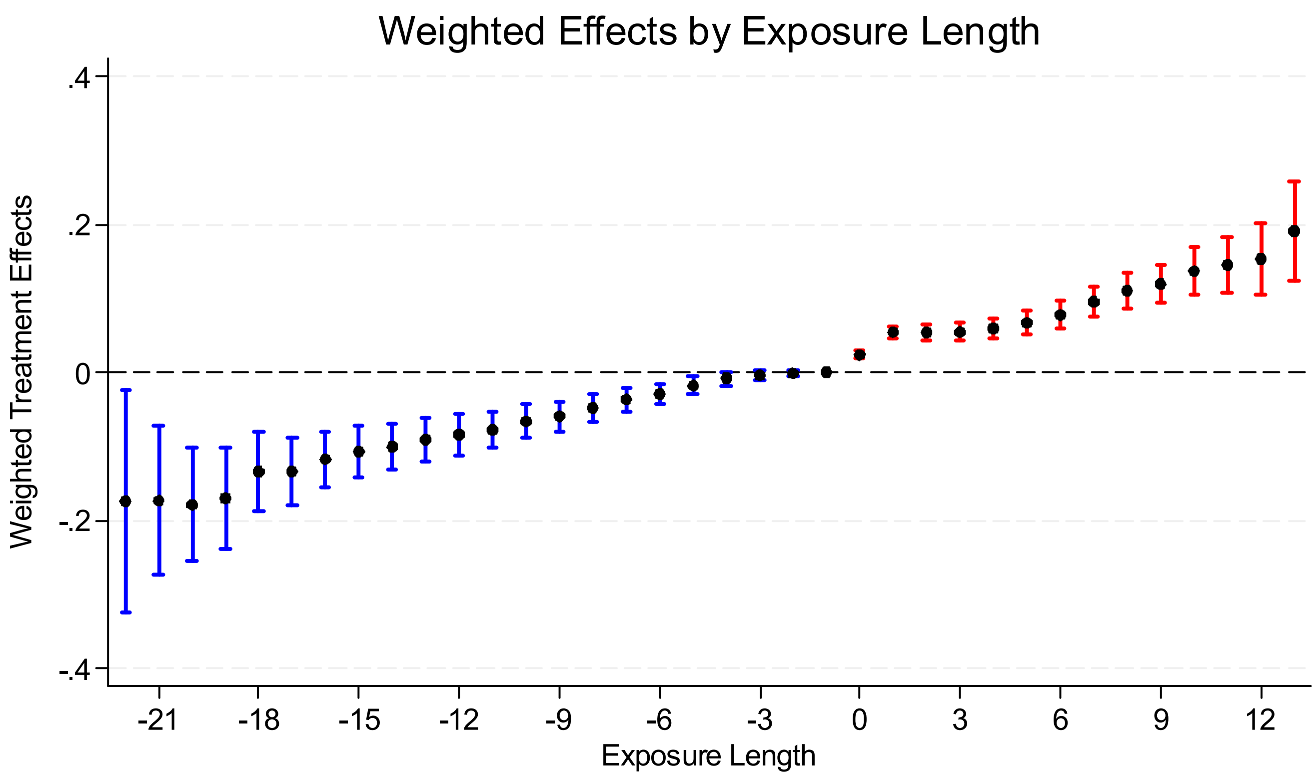
- See walmart_lw_retail_n_open_slides.do for full heterogeneity in slopes with trend.
- For the immediate effect:

```
. lincom omega1986_0*c.w#c.d1986#c.f1986 + omega1987_0*c.w#c.d1987#c.f1987 ///
>          + omega1988_0*c.w#c.d1988#c.f1988 + omega1989_0*c.w#c.d1989#c.f1989 ///
>          + omega1990_0*c.w#c.d1990#c.f1990 + omega1991_0*c.w#c.d1991#c.f1991 ///
>          + omega1992_0*c.w#c.d1992#c.f1992 + omega1993_0*c.w#c.d1993#c.f1993 ///
>          + omega1994_0*c.w#c.d1994#c.f1994 + omega1995_0*c.w#c.d1995#c.f1995 ///
>          + omega1996_0*c.w#c.d1996#c.f1996 + omega1997_0*c.w#c.d1997#c.f1997 ///
>          + omega1998_0*c.w#c.d1998#c.f1998 + omega1999_0*c.w#c.d1999#c.f1999

------------------------------------------------------------------------------
log_retail~p | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0049229   .0033239     1.48   0.139     -.001598    .0114437
------------------------------------------------------------------------------
```

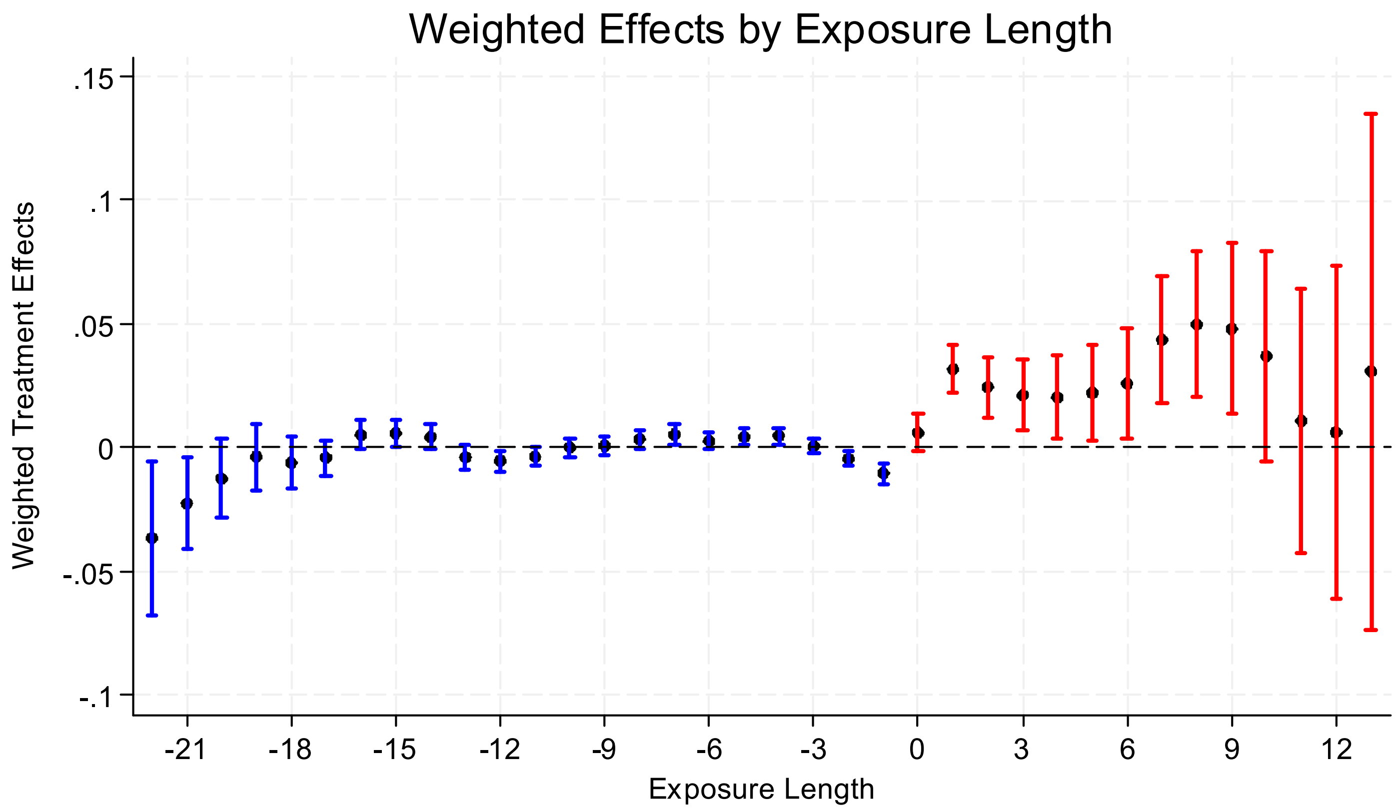
```
. lincom omega1986_1*c.w#c.d1986#c.f1987 + omega1987_1*c.w#c.d1987#c.f1988 ///
>         + omega1988_1*c.w#c.d1988#c.f1989 + omega1989_1*c.w#c.d1989#c.f1990 ///
>         + omega1990_1*c.w#c.d1990#c.f1991 + omega1991_1*c.w#c.d1991#c.f1992 ///
>         + omega1992_1*c.w#c.d1992#c.f1993 + omega1993_1*c.w#c.d1993#c.f1994 ///
>         + omega1994_1*c.w#c.d1994#c.f1995 + omega1995_1*c.w#c.d1995#c.f1996 ///
>         + omega1996_1*c.w#c.d1996#c.f1997 + omega1997_1*c.w#c.d1997#c.f1998 ///
>         + omega1998_1*c.w#c.d1998#c.f1999

------------------------------------------------------------------------------
log_retail~p | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0195912   .0033805     5.80   0.000     .0129594     .026223
------------------------------------------------------------------------------
```

```
. lincom omega1986_11*c.w#c.d1986#c.f1997 + omega1987_11*c.w#c.d1987#c.f1998 ///
>         + omega1988_11*c.w#c.d1988#c.f1999

------------------------------------------------------------------------------
log_retail~p | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0422307   .0146472     2.88   0.004     .0134956    .0709657
------------------------------------------------------------------------------

.
. lincom omega1986_12*c.w#c.d1986#c.f1998 + omega1987_12*c.w#c.d1987#c.f1999

------------------------------------------------------------------------------
log_retail~p | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0587202   .0200277     2.93   0.003     .0194296    .0980108
------------------------------------------------------------------------------

. lincom c.w#c.d1986#c.f1999

------------------------------------------------------------------------------
log_retail~p | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
         (1) |   .0925373   .0301741     3.07   0.002     .0333414    .1517332
------------------------------------------------------------------------------
```

- Collapsing to a binary treatment:



- Cohort-specific linear trends:



5. Summary

- TWFE has been applied to flexible equations to estimate ATTs by cohort/time.
  - ▸ Can obtain "lags only" or "leads and lags."
  - ▸ Aggregation is straightforward, with valid standard errors.
- Can extend to non-binary treatments in a heterogeneous coefficients setting.
  - ▸ Dose-response function for the treated.
  - ▸ Impose a restriction on selection into treatment.

- Can apply TWFE to a flexible equation, and aggregate the coefficients.
  - Use leads or heterogeneous trends to test for violation of PT.
- Would be easy to modify `xthdidregress` to allow non-binary treatments.
  - Can reduce treatment to binary indicator to check `xtreg` commands.
- `wooldid` (Thomas Hegland) allows a single "continuous" treatment, but lots of moving parts.
- Could even allow for exit: keep estimating effects after the policy goes away.