Package 'staggeredGMM'

November 2, 2026

Title GMM Estimation of Treatment Effects Under Staggered Adoption

Version 0.2.0

Description Estimates cohort-by-time average treatment effects under staggered treatment adoption by the generalized method of moments. Three weighting schemes are provided, corresponding to a pooled stationary covariance, a cohort-specific stationary covariance, and an unrestricted within-cohort covariance. Optional adjustment for baseline covariates by outcome regression, and a serial-correlation robust over-identification test of unconditional parallel trends and no anticipation, are also supported. The methods are described in Arora and Bijani (2026) <doi:10.2139/ssrn.6558759>.

License MIT + file LICENSE

Encoding UTF-8

Depends R (>= 4.1.0)

Imports fixest, MASS, stats, utils

Suggests covr, data.table, knitr, rmarkdown, testthat (>= 3.1.5)

Config/testthat/edition 3

VignetteBuilder knitr

LazyData true

URL https://github.com/RishabhBijani/staggeredGMM

BugReports https://github.com/RishabhBijani/staggeredGMM/issues

Config/roxygen2/version 8.0.0

NeedsCompilation no

Author Rishabh Bijani [aut, cre, cph] (ORCID: <https://orcid.org/0009-0003-9636-1774>), Parush Arora [aut, cph] (ORCID: <https://orcid.org/0009-0004-6571-3850>)

Maintainer Rishabh Bijani <rishabhbijani@gmail.com>

Repository CRAN

Date/Publication 2026-11-02 15:40:02 UTC

Contents

coef.staggered_gmm 2
confint.staggered_gmm 3
gmm_j_test 3
gmm_staggered 5
print.staggered_gmm 8
print.staggered_gmm_jtest 8
sim_panel 9
summary.staggered_gmm 10
vcov.staggered_gmm 11

Index **12**

`coef.staggered_gmm` *Extract cohort-by-time effect estimates*

Description

Extract cohort-by-time effect estimates

Usage

```
## S3 method for class 'staggered_gmm'
coef(object, ...)
```

Arguments

`object`	A `staggered_gmm` object, as returned by `gmm_staggered()`.
`...`	Ignored.

Value

A named numeric vector of cohort-by-time effects, one per cell, named `"g<cohort>:t<period>"`. Cells with no clean comparison are `NA`.

Examples

```
fit <- gmm_staggered(sim_panel, yname = "y", tname = "year",
                     idname = "unit_id", gname = "cohort")
head(coef(fit))
```

confint.staggered_gmm *Confidence intervals for cohort-by-time effects*

Description

Pointwise normal-approximation intervals. They are not adjusted for multiplicity, so reading a whole event-study path off them overstates confidence.

Usage

```
## S3 method for class 'staggered_gmm'
confint(object, parm, level = 0.95, ...)
```

Arguments

object	A staggered_gmm object, as returned by gmm_staggered().
parm	Optional character vector of coefficient names. Defaults to all of them.
level	Confidence level, strictly between 0 and 1.
...	Ignored.

Value

A two-column matrix of lower and upper limits, with one row per requested coefficient. Unidentified cells are NA.

Examples

```
fit <- gmm_staggered(sim_panel, yname = "y", tname = "year",
                     idname = "unit_id", gname = "cohort")
head(confint(fit))
```

gmm_j_test *Specification test for unconditional parallel trends and no anticipation*

Description

A serial-correlation robust joint test of the identifying assumptions, following Section 4.4 of Arora and Bijani (2026).

Usage

```
gmm_j_test(object, type = c("local", "full"), window = 3L)
```

Arguments

`object`	A fitted `gmm_staggered()` object.
`type`	Character. `"local"` (default) restricts the test to placebo moments within `window` periods before each cohort's adoption; `"full"` uses all of them.
`window`	Integer. Width of the local pre-window, in periods. Ignored when `type = "full"`.

Details

The test is built on the never-treated event-study basis of Eq. (27) of the paper. Fixing a base pre-period t_0, each treated cohort g and period $t \neq t_0$ defines

$$\Delta_{g,t} = (\bar{Y}_{g,t} - \bar{Y}_{g,t_0}) - (\bar{Y}_{\infty,t} - \bar{Y}_{\infty,t_0}).$$

Those with $t \geq g$ identify the cohort-by-time effects. Those with $t < g$ are placebo restrictions whose expectation is zero only under parallel trends and no anticipation. Collecting a set of J placebo moments into $\hat{p}$ with covariance $\hat{\Omega}_p$ implied by the fitted serial-correlation model,

$$J = \hat{p}' \hat{\Omega}_p^{-1} \hat{p} \;\rightarrow\; \chi^2_J$$

under the null. Because the placebo moments do not involve the treatment effects, no degrees of freedom are subtracted for estimated parameters.

The base period is $t_0 = g_{\min} - 1$, the last period before any cohort adopts, so that every moment has an untreated reference.

A large number of restrictions degrades the chi-squared approximation in small panels. `type = "local"` therefore uses only the placebo moments within `window` periods of each cohort's adoption, which holds its nominal size across sample sizes at essentially no cost in power, since a differential trend or an anticipation effect is already visible just before treatment. `type = "full"` uses every placebo moment and is well-sized only when the number of units is comfortably large.

The test is unconditional. The placebo moments are unadjusted group means, whether or not the fitted model adjusts for covariates. After a fit with `covar`, the test therefore checks unconditional parallel trends and no anticipation; it does not test the covariate-conditional parallel trends assumption under which the effects were estimated. If covariates are distributed differently across cohorts, the test can reject although conditional parallel trends hold, and a non-rejection does not validate the conditional assumption. The printed result repeats this warning whenever the fit used covariates.

Value

An object of class `staggered_gmm_jtest`, a list with components `statistic`, `df`, `p_value`, `type`, `window`, `base_period`, `n_moments`, `moments` (the placebo moment table), `scope` (always `"unconditional"`) and `covar` (the covariates of the fitted model, or `NULL`).

References

Arora, P. and Bijani, R. (2026). "Estimating Treatment Effects under Staggered Timing and Non-Spherical Errors." doi:10.2139/ssrn.6558759

Examples

```
fit <- gmm_staggered(sim_panel, yname = "y", tname = "year",
                     idname = "unit_id", gname = "cohort")
gmm_j_test(fit)
```

`gmm_staggered` *GMM estimation of treatment effects under staggered adoption*

Description

Estimates cohort-by-time average treatment effects (CATTs) under staggered treatment adoption by the generalized method of moments, following Arora and Bijani (2026).

Usage

```
gmm_staggered(
  data,
  yname,
  tname,
  idname,
  gname,
  weighting = c("pooled_toeplitz", "cohort_toeplitz", "unrestricted"),
  never_treated = NULL,
  covar = NULL,
  max_iter = 100L,
  tol = 1e-06
)
```

Arguments

`data`	A data frame with one row per unit-period.
`yname`	Character. Name of the outcome column. Must be numeric; factor, character and logical outcomes are rejected rather than coerced.
`tname`	Character. Name of the calendar-time column. Must be integer-valued, and the set of periods present in `data` (pooled across units) must be consecutive.
`idname`	Character. Name of the unit identifier column.
`gname`	Character. Name of the cohort column, giving each unit's first treated period, with `0` for units that are never treated. Must be constant within unit.
`weighting`	Character. The covariance model used to form the optimal GMM weight. One of `"pooled_toeplitz"` (the default: a single stationary autocovariance sequence shared by all groups), `"cohort_toeplitz"` (a separate stationary sequence per cohort), or `"unrestricted"` (a full, unrestricted within-cohort covariance). These correspond to GMM-T, GMM-HT and GMM-U in the paper.

never_treated	Logical or NULL. Whether to use never-treated units as controls in addition to not-yet-treated cohorts. NULL (the default) detects them automatically. TRUE requires that at least one exists and errors otherwise. FALSE excludes them, and warns if any are present.
covar	Optional character vector of baseline (pre-treatment, time-invariant) covariate columns. When supplied, the outcome-regression adjustment of Section 4.5 of the paper is applied, so that parallel trends need hold only conditional on covar. Note that gmm_j_test() does not adjust for covar: it tests unconditional parallel trends.
max_iter	Integer. Maximum number of iterated-GMM reweighting steps.
tol	Numeric. Convergence tolerance on the largest absolute change in any CATT estimate between successive iterations.

Value

An object of class staggered_gmm: a list with components call, weighting, coefficients, vcov, catt, aggregate, weights, convergence, design, controls and validation. See summary.staggered_gmm() and vignette("staggeredGMM").

Data requirements

The following conditions are checked before any estimation, and each raises an error naming the offending rows, units, periods or cohorts:

- data is a data frame and every named column exists.
- The outcome is numeric; time and cohort are integer-valued.
- No missing unit id, period or cohort. Missing outcomes are permitted.
- Each unit-period combination appears exactly once.
- The cohort is constant within unit.
- No period is absent from the whole panel. Individual units may be missing individual periods.
- No unit is treated at or before the first observed period. Such units have no pre-treatment period, so none of their effects is identified, and silently including them contaminates the aggregate.
- If never-treated units are coded 0, period 0 does not lie after the first observed period; otherwise 0 would also mean "first treated in period 0". Code never-treated units with a value after the last period instead.
- At least one cohort adopts strictly inside the observed window.
- With covar, every covariate is observed for at least one unit of each treated cohort, so that the cohort's covariate mean is defined.

Units whose cohort falls after the last observed period are never treated within the window and are therefore pooled with the never-treated group as clean controls, rather than discarded.

Unbalanced panels

Missing outcomes are supported. Group-period means are taken over the units observed at that period; autocovariances are divided by the number of residual pairs actually observed at each lag; and the variance of a group-period mean accounts for the overlap between the unit sets contributing at each pair of periods. When the panel is balanced all three reduce exactly to the complete-data formulas. Validity requires that missingness be independent of the outcome given group and period; selection on outcomes is not addressed by any weighting scheme.

Partial identification

A cohort-by-time effect with no clean comparison is not estimable. Such cells are reported with `estimate` = `NA` and `identified` = `FALSE`. Where an unidentified cell carries positive aggregation weight the corresponding aggregate is `NA`, and an identified-subset aggregate renormalised over the estimable cells is reported alongside it, with the share of weight it covers.

Convergence

Three states are reported separately. `converged` is `TRUE` only when the largest change in any CATT fell below `tol`. `solve_ok` records the weaker fact that at least one weighted solve succeeded. `termination` gives the reason the iteration stopped.

When `solve_ok` is `FALSE`, no weighted step was ever completed and the returned estimates are the identity-weighted seed rather than a GMM estimate under the requested weighting. That always raises a warning.

When `solve_ok` is `TRUE` but `converged` is `FALSE`, the iteration ran and did not settle; the estimates are the last valid iterate. That warns for the two Toeplitz weightings but not for `"unrestricted"`, where slow convergence is routine.

References

Arora, P. and Bijani, R. (2026). "Estimating Treatment Effects under Staggered Timing and Non-Spherical Errors." doi:10.2139/ssrn.6558759

See Also

`gmm_j_test()` for the over-identification test of parallel trends and no anticipation.

Examples

```
fit <- gmm_staggered(sim_panel, yname = "y", tname = "year",
                     idname = "unit_id", gname = "cohort")
fit
```

print.staggered_gmm *Print a fitted staggered GMM model*

Description

Print a fitted staggered GMM model

Usage

```
## S3 method for class 'staggered_gmm'
print(x, ...)
```

Arguments

x	A staggered_gmm object, as returned by gmm_staggered().
...	Ignored.

Value

x, invisibly. Called for the side effect of printing.

Examples

```
fit <- gmm_staggered(sim_panel, yname = "y", tname = "year",
                     idname = "unit_id", gname = "cohort")
print(fit)
```

print.staggered_gmm_jtest *Print a specification test result*

Description

Print a specification test result

Usage

```
## S3 method for class 'staggered_gmm_jtest'
print(x, ...)
```

Arguments

x	A staggered_gmm_jtest object, as returned by gmm_j_test().
...	Ignored.

Value

x, invisibly. Called for the side effect of printing.

Examples

```
fit <- gmm_staggered(sim_panel, yname = "y", tname = "year",
                     idname = "unit_id", gname = "cohort")
print(gmm_j_test(fit))
```

---

sim_panel *Simulated staggered adoption panel*

Description

A balanced synthetic panel used in the examples, the vignette and the tests: 60 units observed over 33 periods, with five treatment cohorts (first treated in periods 10, 13, 16, 19 or 22; 10 units each) and a never-treated group of 10 units. True effects are deterministic given (cohort, period), so both aggregate targets are known exactly: `ATT_CW` = `-16.79375` under treated-observation weighting and `ATT_EW` = `-15.82020` under cohort-equal weighting. The two differ because earlier cohorts are observed for more post-treatment periods; see `gmm_staggered()` for the definitions.

Usage

```
sim_panel
```

Format

A data frame with 1980 rows and 6 variables:

unit_id Unit identifier, integer 1–60.

year Period, integer 1–33.

cohort 0 = never treated; 10, 13, 16, 19 or 22 = first treated period.

y Observed outcome.

x1 Continuous baseline covariate.

x2 Binary baseline covariate.

Details

The data-generating process, for unit i in cohort g_i at period t:

$$y_{it} = \alpha_i + \lambda_t + \tau_{it} + x_{1i}\theta_{1t} + x_{2i}\theta_{2t} + \varepsilon_{it}$$

- $\alpha_i \sim N(0,1)$: unit fixed effect.
- $\lambda_t \sim N(0,1)$: period fixed effect.
- $\tau_{it} = \beta_{g_i}(1 + r_{g_i})^{t-g_i}$ for $t \geq g_i$ and 0 otherwise, with $\beta = (-16, -12, -10, -9, -2)$ and $r = (0.01, 0.04, 0.08, 0.10, 0.07)$ for cohorts $(10, 13, 16, 19, 22)$.

- ε_{it}: AR(1) within unit, $\rho = 0.5$.
- x_{1i}, x_{2i}: baseline covariates with cohort-correlated means and linear-in-time loadings (θ_{1t}, θ_{2t}), so an unconditional difference-in-differences is genuinely confounded by them.

`set.seed(312844)` throughout. See `data-raw/sim_panel.R` for the exact implementation.

This is the design of Appendix E of the paper, so the specification test has 70 full-set and 14 local-window restrictions on it.

Source

Simulated for Arora, P. and Bijani, R. (2026). "Estimating Treatment Effects under Staggered Timing and Non-Spherical Errors." doi:10.2139/ssrn.6558759

`summary.staggered_gmm` *Summarise a fitted staggered GMM model*

Description

Prints everything `print.staggered_gmm()` shows, followed by the full table of cohort-by-time effects and the covariate-adjustment status.

Usage

```
## S3 method for class 'staggered_gmm'
summary(object, ...)
```

Arguments

`object`	A `staggered_gmm` object, as returned by `gmm_staggered()`.
`...`	Ignored.

Value

`object`, invisibly. Called for the side effect of printing.

Examples

```
fit <- gmm_staggered(sim_panel, yname = "y", tname = "year",
                     idname = "unit_id", gname = "cohort")
summary(fit)
```

`vcov.staggered_gmm` *Covariance matrix of the estimated cohort-by-time effects*

Description

Covariance matrix of the estimated cohort-by-time effects

Usage

```
## S3 method for class 'staggered_gmm'
vcov(object, ...)
```

Arguments

`object`	A `staggered_gmm` object, as returned by `gmm_staggered()`.
`...`	Ignored.

Value

A square numeric matrix covering the identified cells only, with row and column names matching the corresponding entries of `coef()`.

Examples

```
fit <- gmm_staggered(sim_panel, yname = "y", tname = "year",
                     idname = "unit_id", gname = "cohort")
dim(vcov(fit))
```

Index

∗ **datasets**
sim_panel, 9

coef.staggered_gmm, 2
confint.staggered_gmm, 3

gmm_j_test, 3
gmm_j_test(), *6–8*
gmm_staggered, 5
gmm_staggered(), *2–4*, *8–11*

print.staggered_gmm, 8
print.staggered_gmm(), *10*
print.staggered_gmm_jtest, 8

sim_panel, 9
summary.staggered_gmm, 10
summary.staggered_gmm(), *6*

vcov.staggered_gmm, 11