Package: selectTWFE (via r-universe)

June 2, 2026

Type Package

Title Model Selection Between TWFE and ETWFE

Version 0.2.1

Description Estimates both a vanilla two-way fixed effects (TWFE) model and an extended TWFE (ETWFE) model, then selects between them using Cochran's Q test for heterogeneity. When ETWFE wins, reports the heterogeneity fraction (I-squared) and cohort-time estimates with empirical Bayes shrinkage and Bonferroni multiplicity correction. Methods build on Wooldridge (2025) <doi:10.1007/s00181-025-02807-z> and Callaway and Sant'Anna (2021) <doi:10.1016/j.jeconom.2020.12.001>.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports etwfe, fixest, ggplot2, scales, stats

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

Config/roxygen2/version 8.0.0

NeedsCompilation no

Author Paul von Hippel [aut, cre]

Maintainer Paul von Hippel <ph3828@eid.utexas.edu>

Depends R (>= 3.5.0)

Repository https://cran.r-universe.dev

Date/Publication 2026-06-02 11:20:02 UTC

RemoteUrl https://github.com/cran/selectTWFE

RemoteRef HEAD

RemoteSha 86ec39a73ecd4a0240e6806eeb543c984dc118c3

Contents

mpdta 2
plot.select_twfe 3
print.select_twfe 3
print.summary.select_twfe 4
select_twfe 4
summary.select_twfe 6

Index **7**

`mpdta` *Minimum wage and teen employment data*

Description

County-level panel data on log teen employment and log population in 500 US counties from 2003 to 2007, used to study the effect of state minimum wage increases on teen employment. Treatment is staggered: counties are grouped into cohorts by the year their state first raised the minimum wage above the federal level.

Usage

```
mpdta
```

Format

A data frame with 2,500 rows and 6 variables:

countyreal County FIPS code (unit identifier)

year Calendar year (2003–2007)

first.treat Year of first minimum wage increase for the county's state; `0` for never-treated counties

lemp Log teen employment (outcome)

lpop Log county population (covariate)

treat Binary indicator equal to 1 if the county is treated in that year (i.e., `year >= first.treat` and `first.treat > 0`)

Source

Callaway, B. and Sant'Anna, P.H.C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230. Originally distributed with the **did** package.

plot.select_twfe *Plot method for select_twfe objects*

Description

When ETWFE wins, produces an event study plot of cohort-time estimates, optionally shrunk toward the grand mean via empirical Bayes. When TWFE wins, produces a simple ATT plot.

Usage

```
## S3 method for class 'select_twfe'
plot(x, ...)
```

Arguments

x	A select_twfe object
...	Ignored

Value

Invisibly returns the input x (a `select_twfe` object). Called for its side effect of drawing a `ggplot` to the active graphics device: an event study plot of cohort-time effects with Bonferroni-adjusted 95% confidence intervals when ETWFE is selected, or a single-point ATT plot when TWFE is selected.

print.select_twfe *Print method for select_twfe objects*

Description

Print method for select_twfe objects

Usage

```
## S3 method for class 'select_twfe'
print(x, digits = 4, ...)
```

Arguments

x	A select_twfe object
digits	Number of decimal places for estimates (default 4)
...	Ignored

Value

Invisibly returns the input x (a `select_twfe` object). Called for its side effect of printing a formatted summary of the model selection result to the console.

print.summary.select_twfe *Print method for summary.select_twfe objects*

Description

Print method for summary.select_twfe objects

Usage

```
## S3 method for class 'summary.select_twfe'
print(x, digits = 4, ...)
```

Arguments

x	A summary.select_twfe object
digits	Number of significant digits (default 4)
...	Ignored

Value

Invisibly returns the input x (a summary.select_twfe object). Called for its side effect of printing a formatted summary to the console.

select_twfe *Select a model for two-way fixed effects*

Description

Estimates both a vanilla two-way fixed effects (TWFE) model and an extended TWFE (ETWFE) model, then selects the best model using Cochran's Q test for heterogeneity.

When ETWFE wins, reports the heterogeneity fraction (I^2) and cohort-time estimates with empirical Bayes shrinkage and multiplicity correction.

Usage

```
select_twfe(
  fml,
  tvar,
  gvar,
  data,
  ivar = NULL,
  cgroup = c("notyet", "never"),
  vcov = NULL,
```

```
  selection_criterion = c("Q", "estimated_mse"),
  alpha = 0.05,
  shrink_heterogeneity = TRUE,
  ...
)
```

Arguments

fml	A two-sided formula: outcome ~ controls. Use 1 on RHS if no controls.
tvar	Time variable (unquoted).
gvar	Group/cohort variable (unquoted). Should be 0 or Inf for never-treated.
data	A data frame.
ivar	Unit ID variable (unquoted). Optional; inferred from gvar FE if NULL.
cgroup	Comparison group: "notyet" (default) or "never".
vcov	Variance-covariance specification passed to fixest. Recommended: ~unit_id.
selection_criterion	Character: "Q" (default) or "estimated_mse". Determines how the model is selected. "Q" uses Cochran's Q test for heterogeneity (selects ETWFE if Q significantly indicates heterogeneity). "estimated_mse" uses bias-corrected estimated mean squared error comparison (legacy option, not recommended; Monte Carlo studies show the Q criterion performs better).
alpha	Significance level for Cochran's Q test (default 0.05). Only used when selection_criterion="Q".
shrink_heterogeneity	If TRUE (default) and ETWFE wins, apply empirical Bayes shrinkage to cohort-time estimates in plot output.
...	Additional arguments passed to etwfe().

Value

A `select_twfe` object containing:

selected Character: "etwfe" or "twfe"

att_etwfe Aggregate ATT estimate from ETWFE

se_etwfe Standard error of ETWFE ATT

att_twfe ATT estimate from naive TWFE

se_twfe Standard error of TWFE ATT

bias_twfe Estimated bias of TWFE = ATT(TWFE) - ATT(ETWFE)

var_diff Var(TWFE) - Var(ETWFE); included for reference

cov_WS Score-based sandwich estimate of Cov(ATT_TWFE, ATT_ETWFE)

rho_est Estimated correlation between the two ATT estimators

mse_diff Bias-corrected MSE(TWFE) - MSE(ETWFE) when selection_criterion="estimated_mse"

Q Cochran's Q statistic when selection_criterion="Q"

Q_pval P-value from Cochran's Q test when selection_criterion="Q"

i2 Heterogeneity fraction I^2 (only if ETWFE wins)

estimates_shrunk Data frame of EB-shrunk cohort-time estimates (if ETWFE wins and shrink_heterogeneity=TRUE). Both the point estimates and the standard errors are shrunk: the SE is the naive EB posterior SD $\sqrt{(1 - B_i)\hat{\sigma}_i^2}$. The original (unshrunk) SE is retained in the `se_raw` column.

mod_etwfe The fitted etwfe model object

mod_twfe The fitted feols TWFE model object

agg_etwfe The emfx() aggregated ETWFE results

shrink_heterogeneity Logical: whether shrinkage was applied

selection_criterion The criterion used for model selection

alpha Significance level used for Cochran's Q (if applicable)

`summary.select_twfe` *Summary method for select_twfe objects*

Description

Returns a named list of key quantities from the model selection, suitable for programmatic use. This is distinct from `print()`, which formats results for human reading.

Usage

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

Arguments

`object`	A select_twfe object
`...`	Ignored

Value

A list with elements: `selected`, `att_etwfe`, `se_etwfe`, `att_twfe`, `se_twfe`, `bias_twfe`, `var_diff`, `mse_diff`, `Q`, `Q_pval`, `i2` (heterogeneity fraction I^2; NULL if TWFE wins), `selection_criterion`, and `Q_significance_level`.

Index

∗ **datasets**
    mpdta, 2

mpdta, 2

plot.select_twfe, 3
print.select_twfe, 3
print.summary.select_twfe, 4

select_twfe, 4
summary.select_twfe, 6