Package: xtbhst (via r-universe)

November 1, 2026

Title Bootstrap Slope Heterogeneity Test for Panel Data

Version 1.1.0

Date 2026-09-30

Description Implements the bootstrap slope heterogeneity test for panel data of Blomquist and Westerlund (2016) <doi:10.1007/s00181-015-0978-z>. Tests the null hypothesis that slope coefficients are homogeneous across cross-sectional units using a block bootstrap of the Swamy-type statistic, with the unit-specific variance estimator of the paper or that of Pesaran and Yamagata (2008) <doi:10.1016/j.jeconom.2007.05.010>, whose Delta and adjusted Delta statistics are reported with asymptotic p-values. Supports partialling out of control variables and cross-sectional averages.

License GPL-3

URL https://github.com/muhammedalkhalaf/xtbhst

BugReports https://github.com/muhammedalkhalaf/xtbhst/issues

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Depends R (>= 3.5.0)

Imports stats, graphics, grDevices

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

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

Date/Publication 2026-10-02 22:48:13 UTC

RemoteUrl https://github.com/muhammedalkhalaf/xtbhst

RemoteRef HEAD

RemoteSha 6c1d4040343014884900743876 92f166835a9c30

Contents

plot.xtbhst 2
print.xtbhst 3
summary.xtbhst 3
xtbhst 4

Index **8**

`plot.xtbhst` *Plot method for xtbhst objects*

Description

Produces diagnostic plots for the bootstrap slope heterogeneity test.

Usage

```
## S3 method for class 'xtbhst'
plot(x, which = c(1L, 2L), ask = NULL, ...)
```

Arguments

`x`	An object of class `"xtbhst"`.
`which`	Integer vector specifying which plots to produce: 1 = Bootstrap distribution of the statistic S, 2 = Bootstrap distribution of S on the Delta scale (the fixed rescaling `sqrt(N) (S/N - K)/sqrt(2K)` of S and S*), 3+ = Individual coefficient distributions (plot 3 is the first regressor, plot 4 the second, and so on). Default is `c(1, 2)`.
`ask`	Logical. If `TRUE`, prompt before each plot (default: `TRUE` if multiple plots and interactive session).
`...`	Additional arguments passed to plotting functions.

Value

Invisibly returns `NULL`.

print.xtbhst *Print method for xtbhst objects*

Description

Print method for xtbhst objects

Usage

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

Arguments

x	An object of class "xtbhst".
digits	Number of digits to display (default: 4).
...	Additional arguments (ignored).

Value

Invisibly returns the input object.

summary.xtbhst *Summary method for xtbhst objects*

Description

Summary method for xtbhst objects

Usage

```
## S3 method for class 'xtbhst'
summary(object, digits = 4L, ...)
```

Arguments

object	An object of class "xtbhst".
digits	Number of digits to display (default: 4).
...	Additional arguments (ignored).

Value

Invisibly returns a list with summary statistics.

`xtbhst` *Bootstrap Slope Heterogeneity Test for Panel Data*

Description

Implements the panel bootstrap test of slope homogeneity of Blomquist and Westerlund (2016). The null hypothesis is that the slope coefficients are equal across all cross-sectional units.

Usage

```
xtbhst(
  formula,
  data,
  id,
  time,
  reps = 999L,
  blocklength = NULL,
  variance = c("bw", "py"),
  partial = NULL,
  csa = NULL,
  csa_lags = 0L,
  constant = TRUE,
  seed = NULL
)
```

Arguments

`formula`	A formula of the form `y ~ x1 + x2 + ...` specifying the dependent variable and regressors.
`data`	A data frame containing the panel data.
`id`	A character string specifying the name of the cross-sectional identifier variable.
`time`	A character string specifying the name of the time variable.
`reps`	Integer. Number of bootstrap replications (default: 999).
`blocklength`	Integer. Block length for the block bootstrap. If `NULL` (default), set to `round(2 * T^(1/3))`, the deterministic rule of Blomquist and Westerlund (2016, Section 4.2), where `T` is the number of time periods in the estimation sample. Must lie between 1 and `T`.
`variance`	Character. Which unit-specific error variance estimate to use in the weights of the weighted fixed-effects estimator and in the test statistic. `"bw"` (default) is the estimator of Blomquist and Westerlund (2016, Section 3.1); `"py"` is the estimator of Pesaran and Yamagata (2008). See Details.
`partial`	Optional formula specifying variables to be partialled out unit by unit. For example, `~ z1 + z2`.
`csa`	Optional formula specifying variables for which cross-sectional averages should be computed and partialled out, in the spirit of Pesaran (2006). See Details.

csa_lags	Integer. Number of lags of the cross-sectional averages to include (default: 0).
constant	Logical. If TRUE (default), a unit-specific constant is partialled out.
seed	Optional integer seed for reproducibility.

Details

Let $\hat{\beta}_i$ be the least-squares estimate of the slopes of unit i after partialling out the constant (and any further variables given in partial and csa), let M denote the corresponding projection matrix, and let

$$\hat{\beta}_{WFE} = \left(\sum_{i=1}^{N} \frac{x_i' M x_i}{\hat{\sigma}_i^2} \right)^{-1} \sum_{i=1}^{N} \frac{x_i' M y_i}{\hat{\sigma}_i^2}$$

be the weighted fixed-effects estimator. The test statistic of Blomquist and Westerlund (2016, Section 3.1) is

$$S = \sum_{i=1}^{N} (\hat{\beta}_i - \hat{\beta}_{WFE})' \frac{x_i' M x_i}{\hat{\sigma}_i^2} (\hat{\beta}_i - \hat{\beta}_{WFE}).$$

Two estimators of σ_i^2 are available:

variance = "bw" $\hat{\sigma}_i^2 = T^{-1} \sum_t \hat{\varepsilon}_{i,t}^2$, where $\hat{\varepsilon}_{i,t}$ are the residuals of the unit-specific least-squares regression. This is the estimator used by Blomquist and Westerlund (2016).

variance = "py" $\tilde{\sigma}_i^2 = (T - K - 1)^{-1} (y_i - x_i \hat{\beta}_{FE})' M (y_i - x_i \hat{\beta}_{FE})$, where $\hat{\beta}_{FE}$ is the pooled fixed-effects estimator. This is the estimator used by Pesaran and Yamagata (2008). When further variables are partialled out, the degrees of freedom are $T - K - K_{partial}$, which reduces to $T - K - 1$ in the standard case.

The same choice is used for the weights of $\hat{\beta}_{WFE}$, in S and in every bootstrap statistic S^*.

The bootstrap follows Algorithm BOOT of Blomquist and Westerlund (2016): the unit-specific least-squares residuals are resampled in blocks of length blocklength along the time dimension (keeping the cross-section intact), pseudo-data are generated under the null hypothesis with slopes $\hat{\beta}_{WFE}$, and S^* is computed exactly as S. The reported bootstrap p-value is the proportion of S^* that are at least as large as S.

For reference, the standardized statistics of Pesaran and Yamagata (2008) are also reported with their asymptotic standard normal p-values:

$$\Delta = \sqrt{N} \frac{N^{-1} S - K}{\sqrt{2K}}, \qquad \Delta_{adj} = \sqrt{N} \frac{N^{-1} S - K}{\sqrt{2K(T - K - 1)/(T + 1)}}.$$

These are always computed from S evaluated with the Pesaran and Yamagata (2008) variance estimator $\tilde{\sigma}_i^2$ (returned as S_py), whatever the value of variance, so that they are the statistics of that paper with their asymptotic null distribution; variance governs only S and the bootstrap. Because Δ and Δ_{adj} are fixed rescalings of S_py, a bootstrap p-value for them would coincide with the bootstrap p-value of S when variance = "py"; only the asymptotic p-values are therefore reported for them. Δ_{adj} is NA when $T - K - 1 \leq 0$.

Rows of data with missing values in any of the variables used (the model, partial, csa, id and time) are dropped before the panel structure is checked. The remaining panel must be strongly balanced: every unit must be observed in exactly the same set of time periods, and every (id, time) pair must occur once.

Partialling out cross-sectional averages (`csa`, `csa_lags`) is an extension not covered by Blomquist and Westerlund (2016). With `csa_lags > 0` the first `csa_lags` periods are dropped for all units, so that the estimation sample is a common sample of `T - csa_lags` periods. If the dependent variable appears in `csa`, its cross-sectional averages are recomputed from the bootstrap pseudo-data in each replication (observed values are used for lagged averages that fall before the estimation sample). The dependent variable must then be a plain column of `data`.

Value

An object of class `"xtbhst"` containing:

S The Swamy-type statistic S that is bootstrapped, computed with the variance estimator selected by `variance`.

S_py The same statistic computed with the variance estimator of Pesaran and Yamagata (2008); it underlies `delta` and `delta_adj`. Equal to `S` when `variance = "py"`.

pval Bootstrap p-value: the proportion of bootstrap statistics S^* that are at least as large as S.

delta The Delta statistic of Pesaran and Yamagata (2008), computed from `S_py` whatever the value of `variance`.

delta_adj The adjusted Delta statistic of Pesaran and Yamagata (2008), or `NA` if $T - K - 1 \leq 0$.

pval_delta_asy Asymptotic (standard normal, upper tail) p-value of `delta`.

pval_delta_adj_asy Asymptotic (standard normal, upper tail) p-value of `delta_adj`.

S_stars Vector of bootstrap statistics S^*.

delta_stars `S_stars` on the Delta scale, that is `sqrt(N) (S_stars/N - K)/sqrt(2K)`.

blocklength The block length used in the bootstrap.

reps Number of bootstrap replications.

variance The variance estimator used (`"bw"` or `"py"`).

N Number of cross-sectional units.

T Number of time periods in the estimation sample (after dropping the first `csa_lags` periods).

K Number of regressors.

Kpartial Number of partialled-out variables (including the constant).

sigma2 Vector of unit-specific error variance estimates.

beta_i Matrix of individual slope estimates (N x K).

beta_fe Vector of weighted fixed-effects (pooled) estimates.

n_dropped Number of rows of `data` dropped because of missing values.

call The matched call.

References

Blomquist, J. and Westerlund, J. (2016). Panel bootstrap tests of slope homogeneity. *Empirical Economics*, 50(4), 1359-1381. doi:10.1007/s001810150978z

Pesaran, M. H. and Yamagata, T. (2008). Testing slope homogeneity in large panels. *Journal of Econometrics*, 142(1), 50-93. doi:10.1016/j.jeconom.2007.05.010

Pesaran, M. H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4), 967-1012. doi:10.1111/j.14680262.2006.00692.x

Examples

```
# Generate example panel data
set.seed(123)
N <- 20  # cross-sectional units
T_periods <- 30  # time periods

# Homogeneous slopes (H0 is true)
data_hom <- data.frame(
  id = rep(1:N, each = T_periods),
  time = rep(1:T_periods, N),
  x = rnorm(N * T_periods)
)
data_hom$y <- 1 + 0.5 * data_hom$x + rnorm(N * T_periods)

# Test for slope heterogeneity
result <- xtbhst(y ~ x, data = data_hom, id = "id", time = "time",
                 reps = 199, seed = 42)
print(result)
summary(result)

# Pesaran and Yamagata (2008) variance estimator
result_py <- xtbhst(y ~ x, data = data_hom, id = "id", time = "time",
                    reps = 199, seed = 42, variance = "py")
```

Index

plot.xtbhst, 2
print.xtbhst, 3

summary.xtbhst, 3

xtbhst, 4