

Package: semTests (via r-universe)

October 17, 2026

Type Package

Title Robust Test Statistics for Structural Equation Models

Description Computes robust p-values for overall fit and nested comparisons of structural equation models fitted with 'lavaan'. Implements penalized eigenvalue block averaging and penalized regression (Foldnes, Moss, Grønneberg, 2025) [doi:10.1080/10705511.2024.2372028](https://doi.org/10.1080/10705511.2024.2372028), including their extension to nested models (Foldnes, Grønneberg, Moss, 2026) [doi:10.3758/s13428-026-02968-4](https://doi.org/10.3758/s13428-026-02968-4), alongside familiar corrections such as Satorra-Bentler. Supported settings include complete-data ML, GLS, and ULS, categorical DWLS and ULS, and full-information maximum likelihood with one or several groups.

Version 1.0.0

License GPL (>= 3)

Encoding UTF-8

URL <https://github.com/JonasMoss/semTests>

Depends R (>= 3.5.0)

Imports lavaan (>= 0.7-2), methods

Suggests testthat (>= 3.0.0), knitr, rmarkdown

VignetteBuilder knitr

Config/testthat/edition 3

Roxygen list(markdown = TRUE)

Config/roxygen2/version 8.0.0

Repository <https://jonasmoss.r-universe.dev>

Date/Publication 2026-07-19 18:36:17 UTC

RemoteUrl <https://github.com/jonasmoss/semtests>

RemoteRef HEAD

RemoteSha e54d90b27c03f151a73ad57a456308ad44862328

Contents

print.semTests_pvalues	2
pvalues	2
semTests-support	6
Index	9

print.semTests_pvalues	<i>Print method for p-values from pvalues() / pvalues_nested().</i>
------------------------	---

Description

Prints the p-values, then a one-line provenance footer (estimator, data type, information, df) recording the options used.

Usage

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

Arguments

- x A semTests_pvalues object.
- ... Passed to the default print method.

Value

x, invisibly.

pvalues	<i>Compute robust p-values for one or two lavaan objects</i>
---------	--

Description

Compute one or several *p*-values for a fitted lavaan model. Available methods include penalized eigenvalue block averaging, penalized regression, and familiar robust corrections. The defaults select the recommended method for the fitted model.

Usage

```
pvalues(
  object,
  tests = if (is_classic_nt(object)) "pEBA4_RLS" else "pEBA4",
  fiml.convention = c("observed", "lavaan")
)

pvalues_nested(
  m0,
  m1,
  method = c("2000", "2001"),
  tests = if (is_classic_nt(m0)) "PALL_UG_ML" else "PALL",
  A.method = c("delta", "exact"),
  fiml.convention = c("observed", "lavaan")
)
```

Arguments

<code>object, m0, m1</code>	One or two lavaan objects. <code>pvalues</code> does goodness-of-fit testing on one object, while <code>pvalues_nested</code> does hypothesis testing on two nested models.
<code>tests</code>	A non-empty character vector of tests. Each element uses one of TEST, TEST_UG, TEST_ML, TEST_RLS, TEST_UG_ML, or TEST_UG_RLS. The defaults are the recommended options. EBA and pEBA take a positive integer number of blocks (for example, "pEBA4"), no greater than the test df. pOLS takes a finite positive penalty (for example, "pOLS2").
<code>fiml.convention</code>	For FIML fits, use the fully observed-information convention ("observed", the default) or reproduce lavaan 0.7-2's inspected robust-test spectrum ("lavaan").
<code>method</code>	Nested reduction method. Only Satorra's "2000" construction (the paper-recommended default) is available. The Satorra–Bentler "2001" construction has been withdrawn because it performs poorly, and requesting it points you to "2000".
<code>A.method</code>	For nested FIML or categorical models, choose "exact" for the literal parameter restriction map or "delta" for the local moment-tangent restriction map. "delta" is the default and applies to a wider range of equivalent model parameterizations. Categorical nested models support "delta" only.

Details

The `tests` argument is a character vector. Each element has one of the forms TEST, TEST_UG, TEST_ML, TEST_RLS, TEST_UG_ML, or TEST_UG_RLS. For example, SB_UG_RLS.

- The first part names the test. The available names are listed below.
- Add UG to use the unbiased estimator of the fourth-order moment matrix from Du and Bentler (2022). Leave it out to use the standard biased matrix. Performance depends on the data-generating process, so neither gamma choice dominates in every setting.
- The final part chooses the chi-square statistic. ML uses the normal-discrepancy statistic (Bollen, 1989). RLS uses Browne's (1974) reweighted least squares statistic.

The peba method is the recommended default. It partitions the eigenvalues into j equally sized sets (if not possible, the smallest set is incomplete), shrinks them towards their common mean, and averages within each set. Provide a positive integer j no greater than the test df. Values from 2 through 6 are usually good candidates. The method was introduced by Foldnes, Moss, and Grønneberg (2025).

pols is a penalized regression method with a finite positive penalization term. Foldnes, Moss, and Grønneberg (2025) studied $\text{pols}=2$, which has good performance in a variety of contexts.

pall penalizes all eigenvalues in ugamm, while all uses all eigenvalues without penalization. pall is the recommended option for nested models, for which the penalized methods were extended and evaluated by Foldnes, Grønneberg, and Moss (2026).

The eba method is the unpenalized predecessor of peba (Foldnes and Grønneberg, 2018). It averages within the eigenvalue blocks without shrinkage. peba usually performs better. eba remains available for comparisons, and values from $j=2$ through $j=4$ tend to work best.

Familiar corrections are available too:

- std uses the ordinary chi-square reference distribution.
- sb gives the Satorra–Bentler p -value (Satorra and Bentler, 1994).
- ss gives the scaled and shifted p -value (Asparouhov and Muthén, 2010).
- sf gives the scaled F p -value (Wu and Lin, 2016).

Estimators and data types:

[semTests-support](#) (?semTests-support) gives the complete matrix of supported estimators, data types, and configurations. Here is the short version.

The limiting null law of the test statistic is a weighted sum of chi-squares for any minimum-discrepancy estimator. pvalues() supports ML/MLM/MLR, GLS, ULS, FIML (missing data), and categorical DWLS/ULS families, with single- and multi-group continuous, ordered, and mixed-indicator fits. pvalues_nested() supports continuous estimators and categorical Satorra-2000 comparisons with a delta restriction map. The fitted model must expose the asymptotic moment covariance. Robust choices such as `test = "satorra.bentler"` or `estimator = "MLM"`, "MLR", or "DWLS" provide it. The RLS statistic (`browne.residual.nt.model`) and the unbiased (UG) Du-Bentler gamma are defined for the classical continuous, complete-data, random-x ML case. Other families use the standard statistic and biased gamma. Fixed or conditional observed exogenous predictors are currently refused. Full WLS/ADF is refused outright: its weight is already the inverse moment covariance, so the correction is exactly the identity and every test would equal the ordinary chi-square.

GLS, ULS, categorical DWLS/ULS, FIML missing data, and nested FIML or categorical comparison are validated to numerical tolerance against an independent implementation and are marked stable; see the Stability section in [semTests-support](#).

Both entry points require a converged fit and warn if lavaan reports an inadmissible solution. Nested comparisons require the same requested estimator, fitting conventions, variables, groups, sample sizes, raw data, and missingness mask. semTests checks comparability. The substantive nesting argument remains part of the analysis.

For FIML, `fiml.convention = "observed"` (the default) uses observed saturated and model information throughout. An independent implementation validates this construction. The "lavaan" option reproduces lavaan 0.7-2's robust-test construction. For a single model it uses lavaan's inspected UGamma. For nested models it uses lavaan's H1 weight convention and selected model

information. The returned object records the convention, which keeps this inferential choice visible in saved results.

For a single FIML model, "lavaan" means the eigenvalue spectrum returned by `lavInspect(fit, "UGamma")`. The scalar Yuan–Bentler–Mplus test stored by a default MLR fit is a different correction and may give a different result. For nested FIML models, "lavaan" reproduces `lavTestLRT(..., method = "satorra.2000")`.

Value

A named numeric vector of p-values, of class `semTests_pvalues`, carrying an "semtests" attribute that records the requested tests, requested and base estimator, base-statistic and gamma choices, information type, data type, parameterization where applicable, nested method and restriction map where applicable, and degrees of freedom.

References

- Foldnes, N., Moss, J., & Grønneberg, S. (2025). Improved goodness of fit procedures for structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 32(1), 1–13. doi:10.1080/10705511.2024.2372028
- Foldnes, N., Grønneberg, S., & Moss, J. (2026). Penalized eigenvalue block averaging: Extension to nested model comparison and Monte Carlo evaluations. *Behavior Research Methods*, 58, article 107. doi:10.3758/s13428026029684
- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in Multivariate Statistical Analysis* (pp. 233–247). Kluwer Academic. doi:10.1007/9781461546030_17
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507–514. doi:10.1007/BF02296192
- Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent Variables Analysis: Applications for Developmental Research* (pp. 399–419). Sage.
- Asparouhov, T., & Muthén, B. O. (2010). *Simple second order chi-square correction*. Mplus Technical Appendix. https://www.statmodel.com/download/WLSMV_new_chi21.pdf
- Wu, H., & Lin, J. (2016). A Scaled F Distribution as an Approximation to the Distribution of Test Statistics in Covariance Structure Analysis. *Structural Equation Modeling*. doi:10.1080/10705511.2015.1057733
- Foldnes, N., & Grønneberg, S. (2018). Approximating Test Statistics Using Eigenvalue Block Averaging. *Structural Equation Modeling*, 25(1), 101–114. doi:10.1080/10705511.2017.1373021
- Du, H., & Bentler, P. M. (2022). 40-Year Old Unbiased Distribution Free Estimator Reliably Improves SEM Statistics for Nonnormal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, 29(6), 872–887. doi:10.1080/10705511.2022.2063870
- Kenward, M. G., & Molenberghs, G. (1998). Likelihood based frequentist inference when data are missing at random. *Statistical Science*, 13(3), 236–247. doi:10.1214/ss/1028905886
- Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons. doi:10.1002/9781118619179
- Browne, M. W. (1974). Generalized least squares estimators in the analysis of covariance structures. *South African Statistical Journal*, 8, 1–24.

See Also

[semTests-support](#) for the full list of supported configurations.

Examples

```
library("semTests")
library("lavaan")
model <- "visual =~ x1 + x2 + x3
          textual =~ x4 + x5 + x6
          speed  =~ x7 + x8 + x9"
object <- cfa(model, HolzingerSwineford1939, estimator = "MLM")
pvalues(object)

# For the pEBA6 method with biased gamma and ML chisq statistic:
pvalues(object, "pEBA6_ML")

# Nested model comparison (constrain the textual loadings to be equal):
constrained <- "visual =~ x1 + x2 + x3
                textual =~ a*x4 + a*x5 + a*x6
                speed  =~ x7 + x8 + x9"
m1 <- cfa(model, HolzingerSwineford1939, estimator = "MLM")
m0 <- cfa(constrained, HolzingerSwineford1939, estimator = "MLM")
pvalues_nested(m0, m1)
```

semTests-support	<i>Supported estimators, data types, and configurations</i>
------------------	---

Description

Use this page to check what [pvalues\(\)](#) and [pvalues_nested\(\)](#) support. The eigenvalue-based p-values target a limiting null law that is a weighted sum of chi-squares. This result holds for any minimum-discrepancy estimator and therefore includes estimators beyond normal-theory ML. The tables below give the supported combinations. Configurations outside them are refused at the entry point (see [check_supported\(\)](#)), and the test suite follows the same boundary.

Details

Single-model ([pvalues\(\)](#)):

Estimator	Data	Groups	Missing	Availability	Stability
ML / MLM / MLR	continuous	single or multi	complete	available	stable
GLS	continuous	single or multi	complete	available	stable
ULS	continuous	single or multi	complete	available	stable
ML / MLR (FIML)	continuous	single or multi	FIML	available	stable
DWLS family	ordered/mixed	single or multi	listwise/pairwise	available	stable
ULS family	ordered/mixed	single or multi	listwise/pairwise	available	stable
WLS (ADF)	continuous	single or multi	complete	rejected	–

WLS (ADF) ordered/mixed single or multi listwise/pairwise rejected –

Nested (pvalues_nested()):

Estimators	Data	Groups	Missing	Method	A.method	Availability	S
ML/MLM/MLR	continuous	single or multi	complete	2000	–	available	s
GLS, ULS	continuous	single or multi	complete	2000	–	available	s
ML / MLR (FIML)	continuous	single or multi	FIML (both fits)	2000 only	exact or delta	available	s
DWLS/ULS families	ordered/mixed	single or multi	listwise/pairwise	2000 only	delta only	available	s
any	continuous	–	mixed / non-FIML	–	–	rejected	–

Stability:

The classical normal-theory ML path (continuous, complete data) is the mature, paper-backed core. The other supported estimators – GLS, ULS, categorical DWLS/ULS fits, FIML missing-data fits, and their nested comparisons – are also marked stable: each is validated to numerical tolerance against an independent implementation (magmaan) across a simulation battery, in addition to the deterministic implementation tests every available row carries. The `fiml.convention = "lavaan"` option reproduces lavaan's own spectrum for compatibility rather than as an independently validated construction. Configurations outside the tables are refused at the entry point.

Observed exogenous covariates:

Observed exogenous predictors are supported when they are modeled jointly with the other variables. In lavaan, request this random-x analysis with `fixed.x = FALSE` and `conditional.x = FALSE`. The independent validation suite covers continuous ML, GLS, ULS, FIML, and mixed categorical DWLS in this setting, for single models and nested comparisons.

Fixed or conditional observed exogenous predictors are currently refused. Their reference distribution conditions on the realized covariate design and needs a separate saturated-model projection. If a covariate is intended to be fixed, keep that scientific choice and do not silently change it to random merely to pass the software check.

Fit quality and nested comparability:

Both entry points require converged lavaan fits and warn when lavaan's post-estimation check reports an inadmissible solution. Nested comparisons additionally require the same requested estimator, fitting conventions, observed variables, groups, sample sizes, raw data, and missingness mask. These checks establish comparability. The scientific argument that one model is genuinely nested within the other still belongs in the analysis.

The constrained model belongs in `m0`. If the models are supplied in reverse df order, `semTests` warns and swaps them. Only the Satorra-2000 reduction is available. `method = "2001"` has been withdrawn because it performs poorly.

Statistic and gamma options:

Test names use one of `TEST`, `TEST_UG`, `TEST_ML`, `TEST_RLS`, `TEST_UG_ML`, or `TEST_UG_RLS` (e.g. "SB_UG_RLS"). See `pvalues()` for the test families. Two of the options are defined only for the classical case:

- The **RLS** statistic (`browne.residual.nt.model`, Browne 1974) and the UG (Du-Bentler) unbiased gamma are available **only** for classical normal-theory ML with continuous, complete

data, estimator = "ML", fixed.x = FALSE, and conditional.x = FALSE. lavaan degrades RLS to ADF elsewhere, and the Du-Bentler correction has no derivation for the other families or a fixed covariate design. semTests refuses either request and points you to the standard statistic and biased gamma.

- **Categorical and mixed-indicator fits** use lavaan's biased inspected UGamma spectrum and unscaled estimator statistic directly. Supported estimator families are DWLS (including WLSMV/WLSM/WLSMVS) and ULS (including ULSMV). Nested comparison is restricted to method = "2000" and A.method = "delta".
- **Pairwise categorical missingness** means support for lavaan's pairwise sample-statistic and UGamma calculation. Pairwise deletion has its own missingness assumptions and does not provide FIML or generally MAR-valid inference.
- **FIML** (missing data) uses the biased gamma and the standard statistic only. UG and RLS are refused. FIML supports one or several groups with continuous data. Observed exogenous predictors are supported under joint random-x inference, using fixed.x = FALSE and conditional.x = FALSE. fiml.convention = "observed" (the default) uses observed saturated and model information, following the recommendation of Kenward and Molenberghs (1998), and is independently validated against magmaan. "lavaan" reproduces lavaan 0.7-2's inspected robust-test spectrum, including its expected H1 weight.

Why full WLS/ADF is refused:

For full WLS (ADF), with continuous or categorical sample statistics, the fitting weight is already the inverse of the asymptotic moment covariance, so UGamma is a projector: its nonzero eigenvalues are all exactly one and the eigenvalue correction collapses to the identity. Every robust p-value would therefore equal the ordinary $1 - \text{pchisq}(\text{chisq}, \text{df})$. Because the correction adds nothing, full WLS is refused at the entry point. Fit a DWLS/ULS family for a non-degenerate robust test, or read the standard chi-square directly.

References

- Foldnes, N., Moss, J., & Grønneberg, S. (2025). Improved goodness of fit procedures for structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 32(1), 1–13. doi:10.1080/10705511.2024.2372028
- Foldnes, N., Grønneberg, S., & Moss, J. (2026). Penalized eigenvalue block averaging: Extension to nested model comparison and Monte Carlo evaluations. *Behavior Research Methods*, 58, article 107. doi:10.3758/s13428026029684
- Kenward, M. G., & Molenberghs, G. (1998). Likelihood based frequentist inference when data are missing at random. *Statistical Science*, 13(3), 236–247. doi:10.1214/ss/1028905886
- Satorra, A. (2000). Scaled and adjusted restricted tests in multi-sample analysis of moment structures. In R. D. H. Heijmans, D. S. G. Pollock, & A. Satorra (Eds.), *Innovations in Multivariate Statistical Analysis* (pp. 233–247). Kluwer Academic. doi:10.1007/9781461546030_17

See Also

`pvalues()`, `pvalues_nested()`

Index

* **models**

semTests-support, [6](#)

check_supported(), [6](#)

print.semTests_pvalues, [2](#)

pvalues, [2](#)

pvalues(), [2](#), [6–8](#)

pvalues_nested(pvalues), [2](#)

pvalues_nested(), [2](#), [6](#), [8](#)

semTests-support, [4](#), [6](#), [6](#)