Package 'ivDiag'

November 2, 2026

Title Estimation and Diagnostic Tools for Instrumental Variables Designs

Version 1.0.8

Date 2026-11-01

Description Estimation and diagnostic tools for instrumental variables designs with one endogenous treatment, implementing the guidelines in Lal, Lockhart, Xu, and Zu (2024) <doi:10.1017/pan.2024.2>. The package reports ordinary least squares and two-stage least squares estimates with analytic and bootstrapped confidence intervals, several first-stage F statistics including the effective F statistic, the Anderson-Rubin test and its confidence set, the tF procedure of Lee, McCrary, Moreira, and Porter (2022) <doi:10.1257/aer.20211063>, and local-to-zero sensitivity tests. It also supports fuzzy regression discontinuity designs.

URL https://yiqingxu.org/packages/ivDiag/

BugReports https://github.com/apoorvalal/ivDiag/issues

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Depends R (>= 3.5.0)

Imports foreach, doRNG, parallelly, doParallel, parallel, lfe, ggplot2, ggfortify, wCorr, haven, patchwork, stats, graphics, utils

Suggests testthat (>= 3.1.5)

NeedsCompilation no

Author Apoorva Lal [aut] (ORCID: <https://orcid.org/0000-0002-3697-614X>), Yiqing Xu [aut, cre] (ORCID: <https://orcid.org/0000-0003-2041-6671>)

Maintainer Yiqing Xu <yiqingxu@stanford.edu>

Repository CRAN

Date/Publication 2026-11-02 09:10:02 UTC

Contents

ivDiag-package . . . 2
AR_test . . . 3
eff_F . . . 4
gsz . . . 5
gsz_south . . . 6
ivDiag . . . 7
ltz . . . 9
mcclendon . . . 10
plot_coef . . . 11
plot_ltz . . . 12
rueda . . . 13
tF . . . 13

Index **15**

`ivDiag-package` *IV Estimation and Diagnostics*

Description

Conducts various estimation and diagnostic procedure for instrumental variable designs in one shot.

Details

Provides estimation and diagnostic tools for instrumental variables designs with one endogenous treatment, implementing the guidelines in Lal et al. (2024): bootstrapped confidence intervals, several first-stage F statistics including the effective F statistic, the Anderson-Rubin test and its confidence set, the tF procedure, local-to-zero sensitivity tests, and an option for fuzzy regression discontinuity designs.

See `ivDiag` for details.

Author(s)

Apoorva Lal; Yiqing Xu

References

Lal, Apoorva, Mackenzie Lockhart, Yiqing Xu, and Ziwen Zu. 2024. "How Much Should We Trust Instrumental Variable Estimates in Political Science? Practical Advice Based on 67 Replicated Studies." Political Analysis 32 (4): 521–540. doi:10.1017/pan.2024.2

AR_test *Anderson Rubin Test*

Description

Performs the Anderson Rubin test, which is robust to weak instruments.

Usage

```
AR_test(data, Y, D, Z, controls, FE = NULL, cl = NULL, weights = NULL,
  prec = 4, CI = TRUE, alpha = 0.05, parallel = NULL, cores = NULL)
```

Arguments

data	name of a dataframe.
Y	a string indicating the outcome variable.
D	a string indicating the treatment variable.
Z	a vector of strings indicating the instrumental variables. A factor or character instrument enters as dummy variables for all its levels but the first (base) level, named <instrument>_<level>, so a factor with K levels counts as K - 1 instruments.
controls	a vector of strings indicating the control variables.
FE	a vector of strings indicating the fixed effects variables.
cl	a string indicating the clustering variable.
weights	a string indicating the variable that stores weights.
CI	a logical flag controlling whether to calcualte the confidence interval using the inversion method.
prec	precision of results (4 by default).
alpha	level of statitical significance; the default is 0.05.
parallel	no longer used (the confidence set no longer needs a grid of regressions); kept for backward compatibility.
cores	no longer used; kept for backward compatibility.

Details

The confidence set collects every value β_0 that the AR test does not reject, on the whole real line. The reduced-form coefficients on the instruments are linear in β_0, and their robust or cluster-robust variance is quadratic in β_0; both are recovered exactly from three regressions and checked at a fourth. With one instrument the set solves a quadratic inequality: a bounded interval, two unbounded rays, the whole real line, or (rarely) empty. With several instruments the AR statistic is evaluated over the whole real line (writing $\beta_0 = \hat{\beta}_{2SLS} + SE \cdot \tan\theta$) and every boundary is refined by root finding; the set can then also be a union of several intervals. If the variance is not quadratic in β_0 (for example, a multiway cluster-robust variance repaired to be positive definite), the regression is run at each point instead.

Versions up to 1.0.6 inverted the test on a grid of $\hat{\beta}_{2SLS} \pm 10$ SEs and could miss pieces of the set outside the grid.

Value

Fstat	F statistic, degrees of freedoms, and p-value.
ci.print	Confidence interval via intervsion (printed version).
ci	Confidence set via inversion (numeric version): the end points in increasing order, in pairs, with -Inf/Inf for unbounded pieces; e.g. c(-Inf, a, b, Inf) is $(-\infty, a] \cup [b, \infty)$. NA if empty.
bounded	If the confidence set is bounded and not empty.

References

Chernozhukov, Victor, and Christian Hansen. 2008. "The Reduced Form: A Simple Approach to Inference with Weak Instruments." Economics Letters 100 (1): 68–71.

See Also

ivDiag

Examples

```
data(ivDiag)
AR.out <- AR_test(data = mcclendon, Y = "intended", D = "openedesteem",
    Z = "esteem", controls = NULL)
AR.out$ci.print
```

---

eff_F *Effective F*

Description

Computes the effective F statistic.

Usage

```
eff_F(data, Y, D, Z, controls = NULL, FE = NULL, cl = NULL,
  weights = NULL, prec = 4)
```

Arguments

data	name of a dataframe.
Y	a string indicating the outcome variable.
D	a string indicating the treatment variable.
Z	a vector of strings indicating the instrumental variables. A factor or character instrument enters as dummy variables for all its levels but the first (base) level, named <instrument>_<level>, so a factor with K levels counts as K - 1 instruments.

controls	a vector of strings indicating the control variables.
FE	a vector of strings indicating the fixed effects variables.
cl	a string indicating the clustering variable.
weights	a string indicating the variable that stores weights.
prec	precision of results (4 by default).

Details

The instruments must be identified: after the intercept, controls and fixed effects are partialled out, they must be linearly independent of each other and not (numerically) zero. Otherwise `eff_F()` stops with an error. Collinear controls are harmless and are ignored. The instruments are centred and scaled internally; the effective F does not depend on their scale or mean. This is not done with two or more clustering variables: lfe then repairs the multiway cluster variance in a way that depends on the scale of the regressors, and the instruments are used as they are, so that with one instrument the effective F equals the cluster-robust F of `ivDiag()`.

Value

the effective F statistic.

References

Olea, José Luis Montiel, and Carolin Pflueger. 2013. "A Robust Test for Weak Instruments."" Journal of Business & Economic Statistics 31 (3): 358–69.

See Also

ivDiag

Examples

```
data(ivDiag)
effF <- eff_F(data = mcclendon, Y = "intended", D = "openedesteem",
    Z = "esteem", controls = NULL)
```

---

gsz *Data from GSZ (2016)*

Description

Data from Guiso, Sapienza, and Zingales (2016)

Format

A data frame with 5357 rows and 11 columns.

Details

The authors revisit Putnam, Leonardi, and Nanetti (1992)'s celebrated conjecture that Italian cities that achieved self-government in the Middle Ages have higher modern-day levels of social capital. More specifically, they study the effects of free city-state status on social capital as measured by the number of nonprofit organizations and organ donations per capita, and a measure of whether students cheat in mathematics. We focus on the first outcome, the number of nonprofit organizations.

References

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2016. "Long-Term Persistence." Journal of the European Economic Association 14 (6): 1401–36.

---

gsz_south *Data from GSZ (2016): Subsample*

Description

Data from Guiso, Sapienza, and Zingales (2016); southern Italian cities

Format

A data frame with 2175 rows and 11 columns.

Details

The authors revisit Putnam, Leonardi, and Nanetti (1992)'s celebrated conjecture that Italian cities that achieved self-government in the Middle Ages have higher modern-day levels of social capital. More specifically, they study the effects of free city-state status on social capital as measured by the number of nonprofit organizations and organ donations per capita, and a measure of whether students cheat in mathematics. We focus on the first outcome, the number of nonprofit organizations.

This dataset is a subsample of southern Italian cities, which is used as a zero-first-stage sample.

References

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2016. "Long-Term Persistence." Journal of the European Economic Association 14 (6): 1401–36.

ivDiag *Omnibus Function for IV Estimation and Diagnostics*

Description

Conducts various estimation and diagnostic procedure for instrumental variable designs in one shot.

Usage

```
ivDiag(data, Y, D, Z = NULL, controls = NULL, FE = NULL, cl = NULL, weights = NULL,
  bootstrap = TRUE, run.AR = TRUE,
  nboots = 1000, parallel = TRUE, cores = NULL,
  seed = 94305, prec = 4, debug = FALSE,
  running = NULL, cutoff = 0, rd.poly = 1, bandwidth = NULL,
  kernel = "triangular")
```

Arguments

data	name of a dataframe.
Y	a string indicating the outcome variable.
D	a string indicating the treatment variable.
Z	a vector of strings indicating the instrumental variables. May be NULL when running is given; the instrument is then the indicator for crossing the cutoff. A factor or character instrument enters as dummy variables for all its levels but the first (base) level, named <instrument>_<level>, so a factor with K levels counts as K - 1 instruments. The tF procedure is then reported only for a factor with two levels.
controls	a vector of strings indicating the control variables.
FE	a vector of strings indicating the fixed effects variables.
cl	a string indicating the clustering variable.
weights	a string indicating the variable that stores weights.
bootstrap	whether to turn on bootstrap (TRUE by default).
run.AR	whether to run AR test (TRUE by default).
nboots	a numeric value indicating the number of bootstrap runs.
parallel	a logical flag controlling parallel computing.
cores	setting the number of cores.
prec	precision of CI in string (4 by default).
seed	seed for the bootstrap. Each bootstrap draw gets its own random-number stream derived from seed (via **doRNG**), so results are the same with or without parallel computing and for any number of cores.
debug	for debugging purposes.

running	a string indicating the running variable of a fuzzy regression discontinuity (RD) design. When it is given, the instrument is the indicator for crossing the cutoff (built when Z = NULL), and a polynomial in the distance to the cutoff, separate on each side, is added to controls, so that only the jump at the cutoff identifies the effect. A treatment predicted from the running variable (for example, from an assignment rule) is not a valid instrument on its own: it also varies smoothly with the running variable, and that variation overstates the first stage. If such a Z is supplied, it is used with the RD controls and a message is printed.
cutoff	the cutoff, or several cutoffs (each unit is then compared with its nearest cutoff, and cutoff fixed effects are added). 0 by default.
rd.poly	order of the polynomial in the running variable, 0 to 4 (1 by default).
bandwidth	keep units within this distance of the cutoff; NULL (the whole sample) by default. See rdrobust::rdbwselect for a data-driven choice.
kernel	"triangular" (default) or "uniform"; used only with a bandwidth. Kernel weights multiply weights.

Value

est_ols	results from an OLS regression.
est_2sls	results from a 2SLS regression.
AR	results from an Anderson-Rubin test
F_stat	various F statistics.
rho	Pearson correlation coefficient between the treatment and predicted treatment from the first stage regression (all covariates are partialled out).
tF	results from the tF procedure based on Lee et al. (2022)
est_rf	results from the reduced form regression.
est_fs	results from the first stage regression.
p_iv	the number of instruments.
N	the number of observations.
N_cl	the number of clusters.
df	the degree of freedom left from the 2SLS regression
nvalues	the unique values the outcome Y, the treatment D, and each instrument in Z in the 2SLS regression.
rd	with running only: the RD setup (running variable, cutoffs, polynomial order, bandwidth, kernel) and the number of units on each side of the cutoff.

Author(s)

Apoorva Lal; Yiqing Xu

References

Lal, Apoorva, Mackenzie Lockhart, Yiqing Xu, and Ziwen Zu. 2024. "How Much Should We Trust Instrumental Variable Estimates in Political Science? Practical Advice Based on 67 Replicated Studies." Political Analysis 32 (4): 521–540. doi:10.1017/pan.2024.2

Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack Porter. 2022. "Valid t-Ratio Inference for IV." American Economic Review 112 (10): 3260–90.

See Also

plot_coef eff_F AR_test tF

Examples

```
data(ivDiag)
g <- ivDiag(data = mcclendon, Y = "intended", D = "openedesteem",
    Z = "esteem", bootstrap = FALSE)
plot_coef(g)
```

---

ltz *Local-to-Zero Test*

Description

Estimates Local-to-Zero IV coefficients and SEs for a single instrument.

Usage

```
ltz(data, Y, D, Z, controls, FE = NULL, cl = NULL, weights = NULL, prior, prec = 4)
```

Arguments

data	name of a dataframe.
Y	a string indicating the outcome variable.
D	a string indicating the treatment variable.
Z	a vector of strings indicating the instrumental variables. A factor or character instrument enters as dummy variables for all its levels but the first (base) level, named <instrument>_<level>, so a factor with K levels counts as K - 1 instruments.
controls	a vector of strings indicating the control variables.
FE	a vector of strings indicating the fixed effects variables.
cl	a string indicating the clustering variable.
weights	a string indicating the variable that stores weights.
prior	prior mean and standard deviation of the direct effect of instrument on outcome.
prec	precision of results (4 by default).

Value

`iv`	results from a 2SLS regression.
`ltz`	results after local-to-zerio adjustment.
`prior`	prior mean and standard deviation

References

Conley, Timothy G, Christian B Hansen, and Peter E Rossi. 2012. "Plausibly Exogenous." Review of Economics and Statistics 94 (1): 260–72.

See Also

`plot_coef`

Examples

```
data(ivDiag)
controls <- c('altitudine', 'escursione', 'costal', 'nearsea', 'population',
    'pop2', 'gini_land', 'gini_income')
ltz_out <- ltz(data = gsz, Y = "totassoc_p", D = "libero_comune_allnord",
    Z = "bishopcity", controls = controls, weights = "population",
    prior = c(0.178, 0.137))
plot_ltz(ltz_out)
```

---

mcclendon *Data from McClendon (2014)*

Description

Data from a field experiment in McClendon (2014) AJPS.

Format

A data frame with 3647 rows and 4 columns:

intended 1 if the person said they would attend the rally (RSVP), 0 otherwise.

attended 1 if the person attended the rally, 0 otherwise.

openedesteem 1 if the person opened an email promising public recognition for attending (the treatment), 0 otherwise.

esteem 1 if the person was randomly assigned to receive such an email (the instrument), 0 if assigned to the control email.

Details

McClendon (2014) worked with an LGBT advocacy organization to invite its members to a Pride rally. Members were randomly assigned to a control email or to one of two emails that promised to recognize attendees publicly, in the organization's newsletter or on its Facebook page. Random assignment to a recognition email (`esteem`) is an instrument for opening such an email (`openedesteem`); only people assigned to a recognition email could open one. The data keep the 3,647 members with complete records. In Lal et al. (2024) the outcome is `intended`, as in Table 2, column 1, of the original article.

The replication data are in the public domain (CC0 1.0), Harvard Dataverse, doi:10.7910/DVN/LVOSAF.

Source

Replication data for McClendon (2014), Harvard Dataverse, doi:10.7910/DVN/LVOSAF (CC0 1.0).

References

McClendon, Gwyneth H. 2014. "Social Esteem and Participation in Contentious Politics: A Field Experiment at an LGBT Pride Rally." American Journal of Political Science 58 (2): 279–290.

`plot_coef` *Plot OLS and IV Coefficents*

Description

Visualise point estimates and confidence intervals of OLS and IV estimates.

Usage

```
plot_coef(out,
  ols.methods = c("analy","bootc","boott"),
  iv.methods = c("analy","bootc","boott","ar","tf"),
  main = NULL, ylab = "Coefficient", grid = TRUE,
  stats = TRUE, ylim = NULL)
```

Arguments

`out`	output from `ivDiag::ivDiag`.
`ols.methods`	a vector specifying the inferential methods for OLS to be shown. The default is `c("analy","bootc","boott")`.
`iv.methods`	a vector specifying inferential methods for 2SLS to be shown. The default is `c("analy","bootc","boott","ar","tf")`.
`main`	a string specifying the title of the plot.
`ylab`	a string specifying ylab of the plot.
`grid`	a logical flag indicating whether to show the grids.

stats	a logical flag indicating whether to show the statistics, including the effective F, the number of observations, and the number of clusters (if applicable).
ylim	a two-element vector specifying the range of the y-axis.

Value

A base R plot object.

See Also

ivDiag

`plot_ltz` *Visualizing Local-to-Zero Adjustment*

Description

Visualise approximate sampling distributions for scalar IV coefficient with local-to-zero adjustment.

Usage

```
plot_ltz(out = NULL, iv_est = NULL, ltz_est = NULL, prior = NULL, xlim = NULL)
```

Arguments

out	output from `ivDiag::ltz`.
iv_est	a two-element vector of IV estimate and standard error.
ltz_est	a two-element vector of local-to-zero estimate and standard error.
prior	a two-element vector of prior mean and standard deviation.
xlim	a two-element vector specifying the range of the x-axis.

Value

A ggplot2 object.

References

Conley, Timothy G, Christian B Hansen, and Peter E Rossi. 2012. "Plausibly Exogenous." Review of Economics and Statistics 94 (1): 260–72.

See Also

ltz

rueda *Data from Rueda (2017)*

Description

Data from Rueda (2017) AJPS.

Format

A data frame with 4352 rows and 6 columns.

Details

Rueda (2017) studies the persistence of vote buying in developing democracies despite the use of secret ballots and argues that brokers condition future payments on published electoral results to enforce these transactions and that this is effective only when the results of small voting groups are available. The study examines the relationship between polling station size and vote buying using three different measures of the incidence of vote buying, two at the municipality level and one at the individual level.

The size of the polling station, predicted by the rules limiting the number of voters per polling station, is used as an instrument of the actual polling place size. The institutional rule predicts sharp reductions in the size of the average polling station of a municipality every time the number of registered voters reaches a multiple of the maximum number of voters allowed to vote in a polling station. Such sharp reductions are used as a source of exogenous variation in polling place size to estimate the causal effect of this variable on vote buying.

Source

Replication data for Rueda (2017), Harvard Dataverse, doi:10.7910/DVN/K6ZOOW (CC0 1.0).

References

Rueda, Miguel R. 2017. "Small Aggregates, Big Manipulation: Vote Buying Enforcement and Collective Monitoring." American Journal of Political Science 61 (1): 163–177.

tF *Valid t-Ratio Procedure*

Description

Performs the valid t-ratio procedure.

Usage

```
tF(coef, se, Fstat, prec = 4)
```

Arguments

coef	a 2SLS coefficient.
se	a standard error estimate for the estimated 2SLS coefficient.
Fstat	a first-stage partial F statistic.
prec	precision of results (4 by default).

Details

The critical value cF is interpolated from the table in Lee et al. (2022) by the square root of Fstat. It is 1.96 for Fstat of 106.09 (10.3 squared) and above. Below Fstat = 4, where the table starts, cF is Inf: the valid critical value is infinite for Fstat up to 3.84 (1.96 squared) and above 18.66 between 3.84 and 4. The confidence interval is then the whole real line and the p-value is 1.

Value

A named vector: the F statistic (F), the critical value (cF), the coefficient, its SE, the t statistic, the confidence interval (Coef plus or minus cF times SE) and the adjusted p-value.

References

Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack Porter. 2022. "Valid t-Ratio Inference for IV." American Economic Review 112 (10): 3260–90.

Examples

```
tF(coef = 0.3149, se = 0.0890, Fstat = 207.18)
```

Index

∗ **datasets**
gsz, 5
gsz_south, 6
mcclendon, 10
rueda, 13

AR_test, 3, *9*

eff_F, 4, *9*

gsz, 5
gsz_south, 6

ivDiag, *2*, *4, 5*, 7, *12*
ivDiag-package, 2

ltz, 9, *12*

mcclendon, 10

plot_coef, *9*, *10*, 11
plot_ltz, 12

rueda, 13

tF, *9*, 13