Package: drlate (via r-universe)

October 23, 2026

Type Package

Title Doubly Robust Estimation of Local Average Treatment Effects

Version 0.3.1

Description Estimates the local average treatment effect (LATE) and the local average treatment effect on the treated (LATT) using observational data with a binary instrument, implementing the complete estimator suite of Sloczynski, Uysal, and Wooldridge: the doubly robust estimators of Sloczynski, Uysal, and Wooldridge (2022) <doi:10.48550/arXiv.2208.01300> -- inverse probability weighted regression adjustment (IPWRA), inverse probability weighting (IPW), augmented inverse probability weighting (AIPW), and regression adjustment (RA) -- and the Abadie-kappa weighting estimators of Sloczynski, Uysal, and Wooldridge (2025) <doi:10.1080/07350015.2024.2332763>. Supports linear, logistic, probit, Poisson, and fractional (fractional-logit and fractional-probit) outcome and treatment models, and instrument propensity scores estimated by maximum likelihood, covariate balancing (CBPS), or inverse probability tilting (IPT). Standard errors are computed jointly for all estimation stages by stacking the moment conditions of every model into a single M-estimation system; weak-instrument-robust Fieller confidence sets, cluster-aware bootstrap inference, design diagnostics, and a doubly robust Hausman-type test of unconfoundedness are included. Estimates and standard errors are validated against the authors' Stata commands 'drlate' (Statistical Software Components S459708) and 'kappalate' (S459257).

License MIT + file LICENSE

Encoding UTF-8

Language en-US

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Depends R (>= 4.1)

Imports stats, numDeriv

Suggests ggplot2, testthat (>= 3.0.0), knitr, rmarkdown, haven

Config/testthat/edition 3

Config/Needs/website pkgdown

VignetteBuilder knitr

URL https://github.com/kvenkita/drlate,
https://kvenkita.github.io/drlate/

BugReports https://github.com/kvenkita/drlate/issues

LazyData true

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

Date/Publication 2026-06-20 14:15:19 UTC

RemoteUrl https://github.com/kvenkita/drlate

RemoteRef HEAD

RemoteSha fb6475e4f836b837e7d5924872888921c8ee8b16

Contents

balance . . . 2
balance_test . . . 3
complier_means . . . 4
confint.drlate . . . 5
dr_hausman . . . 6
drlate . . . 8
drlate_compare . . . 11
drlate_sim . . . 12
kappa_weights . . . 13
plot.drlate . . . 14

Index **15**

`balance` *Covariate balance across instrument arms*

Description

Computes standardized mean differences (SMDs) of the model covariates between the two instrument arms, before and after weighting by the inverse of the estimated instrument propensity score. Well-balanced weighted covariates (conventionally, absolute SMD below 0.1) indicate that the propensity score model is doing its job.

Usage

```
balance(object, ...)

## S3 method for class 'drlate'
balance(object, detail = FALSE, ...)
```

Arguments

object	A fitted drlate() object (with keep_data = TRUE).
...	Currently unused.
detail	Logical. If TRUE, append the IPW-weighted arm means (mean_weighted_z1, mean_weighted_z0) and the unweighted and weighted variance ratios (vratio_unweighted, vratio_weighted, each s_1^2/s_0^2), mirroring the Stata latebalance summarize report. Defaults to FALSE.

Details

The covariate set is the union of the columns of the instrument, outcome, and treatment model matrices (the intercept is dropped). The SMD denominator is the unweighted pooled standard deviation $\sqrt{(s_1^2 + s_0^2)/2}$ in both columns, so the two columns are directly comparable. Weighted arm means are Hájek means using the inverse-propensity weights implied by the fit (for estimand = "latt", the Z=0 arm uses the ATT odds weights $p/(1-p)$, matching the estimator).

Value

A data frame with one row per covariate and columns variable, smd_unweighted, and smd_weighted; with detail = TRUE, the four additional columns described above.

See Also

plot.drlate() with type = "balance" for the love plot.

---

balance_test *Imai-Ratkovic covariate-balance test*

Description

Tests whether the estimated instrument propensity score balances the covariates, using the overidentification test of Imai and Ratkovic (2014). The propensity-score MLE score equations identify the coefficients; the covariate-balancing (CBPS) moments are the overidentifying restrictions. A large statistic is evidence that the propensity-score model does not balance the covariates — a misspecification diagnostic. This is the Stata latebalance overid postestimation feature.

Usage

```
balance_test(object)
```

Arguments

object	A fitted drlate() object (with keep_data = TRUE) using a logistic or probit instrument propensity score.

Value

An object of class drlate_balance_test: a list with statistic (Hansen's J), df, p.value, ivmodel, and n, with a print method.

References

Imai, K. and Ratkovic, M. (2014). Covariate Balancing Propensity Score. *Journal of the Royal Statistical Society B* 76(1), 243–263.

See Also

balance() for the standardized-mean-difference diagnostics.

Examples

```
fit <- drlate(lwage ~ age + educ, nvstat ~ age + educ,
              rsncode ~ age + educ, data = drlate_sim)
balance_test(fit)
```

---

complier_means *Complier covariate means*

Description

Compares the average of each covariate in the full estimation sample with its average in the complier subpopulation, the latter computed with the normalized Abadie kappa weights of kappa_weights(). Because the local average treatment effect is a causal effect for compliers, knowing how compliers differ from the population aids interpretation. This is the Stata estat compliers postestimation feature.

Usage

```
complier_means(object, vars = NULL)
```

Arguments

object	A fitted drlate() object (with keep_data = TRUE) using an instrument propensity score (any method except "ra").
vars	Optional character vector selecting a subset of the model covariates. Defaults to all covariates across the three model formulas.

Details

Covariate values are reported on their original scale.

Value

A data frame with one row per covariate and columns `variable`, `population_mean`, `complier_mean`, and `difference` (`complier_mean - population_mean`).

See Also

`kappa_weights()`

Examples

```
fit <- drlate(lwage ~ age + educ, nvstat ~ age + educ,
              rsncode ~ age + educ, data = drlate_sim)
complier_means(fit)
```

---

confint.drlate *Confidence intervals for drlate fits*

Description

Confidence intervals for drlate fits

Usage

```
## S3 method for class 'drlate'
confint(object, parm, level = 0.95, method = c("default", "fieller"), ...)
```

Arguments

`object`	A fitted `drlate()` object.
`parm`	Coefficients to include (names or indices); defaults to all three reported quantities.
`level`	Confidence level.
`method`	`"default"` gives Wald intervals from the joint sandwich (or bootstrap percentile intervals when the fit used `vcov = "bootstrap"`). `"fieller"` inverts the test of `num - t * denom = 0` using the joint covariance of the numerator and denominator, giving a confidence set for the LATE/LATT ratio that remains valid when the first stage is weak; the set may be an interval, the complement of an interval, or the whole line, and is returned as a `"drlate_fieller"` object with its own print method.
`...`	Currently unused.

Value

For `method = "default"`, a numeric matrix with one row per requested coefficient (`parm`) and two columns holding the lower and upper confidence limits. The columns are labelled with the corresponding percentiles (for the default 95% level, `"2.5 %"` and `"97.5 %"`). The limits are Wald intervals from the joint sandwich covariance, or percentile intervals from the resampling draws when the fit was computed with `vcov = "bootstrap"`.

For `method = "fieller"`, an object of class `"drlate_fieller"`: a list describing the weak-instrument-robust confidence set for the LATE/LATT ratio (its endpoints and shape, the estimand name, and the confidence level), with its own `print` method. Because a Fieller set need not be a bounded interval, it is returned in this form rather than as a matrix of endpoints.

---

dr_hausman *Doubly robust Hausman test of unconfoundedness*

Description

Tests whether the treatment is unconfounded given the covariates, using the comparison proposed by Słoczyński, Uysal, and Wooldridge (2022, Section 5), building on Donald, Hsu, and Lieli (2014). Under **one-sided noncompliance** (nobody takes the treatment without the instrument: $\Pr(D = 1 \mid Z = 0) = 0$), the LATT identified through the instrument equals the ATT identified through unconfoundedness of the treatment — so a significant difference between the doubly robust LATT estimate (which uses the instrument) and the doubly robust ATT estimate (which does not) is evidence against unconfoundedness. Unlike the textbook OLS-vs-IV Hausman test, this comparison is robust to treatment effect heterogeneity.

Usage

```
dr_hausman(
  outcome,
  treatment,
  instrument,
  data,
  omodel = c("linear", "logit", "poisson"),
  tmodel = c("logit", "linear", "poisson"),
  ivmodel = c("logit", "ipt"),
  weights = NULL,
  cluster = NULL,
  pstolerance = 1e-05,
  subset = NULL
)
```

Arguments

`outcome`	A formula `y ~ covariates` for the outcome model. Use `y ~ 1` for no covariates (required when `method = "ipw"`).
`treatment`	A formula `d ~ covariates` for the treatment model.

`instrument`	A formula `z ~ covariates` for the instrument propensity score model; `z` must be binary 0/1. Use `z ~ 1` when `method = "ra"`.
`data`	A data frame containing all variables.
`omodel`	Outcome model family: `"linear"` (default; continuous), `"logit"` or `"probit"` (outcome must be 0/1), `"poisson"` (outcome must be non-negative), or `"flogit"` / `"fprobit"` (fractional outcome in `[0, 1]`, e.g. a proportion). The `f`-prefixed families share all estimation with `"logit"` / `"probit"` and only relax the response to the unit interval, matching the **Stata** `lateffects omodel` options.
`tmodel`	Treatment model family: `"logit"` (default; treatment must be 0/1), `"probit"`, `"linear"`, or `"poisson"`.
`ivmodel`	Instrument propensity score model for the LATT half: `"logit"` (default) or `"ipt"`.
`weights`	Optional sampling weights (a numeric vector, or a column name in `data` given as a string).
`cluster`	Optional cluster identifier for clustered standard errors (a vector, or a column name in `data` given as a string).
`pstolerance`	Overlap tolerance: estimation stops with an error if any estimated instrument propensity score is below `pstolerance` or above `1 - pstolerance`. Default `1e-5`.
`subset`	Optional logical or integer vector selecting rows of `data`.

Details

The DR ATT estimator follows the paper's equation (33): a treatment propensity score $\Pr(D = 1 \mid X)$ is fitted by logit QMLE on the treatment-equation covariates; the outcome model is fitted on the untreated sample weighted by the odds $\hat{p}/(1-\hat{p})$; and $\hat{\tau}_{ATT}$ is the treated-sample mean outcome minus the mean imputed counterfactual. The standard error of the difference comes from stacking the moment conditions of *both* estimators (and the difference) into one M-estimation system, so the covariance between them is accounted for analytically — the analytic option suggested in the paper.

Note that the two halves adjust on their respective formulas: the LATT half's propensity score uses the *instrument*-equation covariates, while the ATT half's uses the *treatment*-equation covariates (both share the outcome model). Supply the same covariate set to all three formulas unless you intend them to differ.

Value

An object of class `"htest"` with the z statistic, p-value, and the DR LATT, DR ATT, and difference estimates.

References

Słoczyński, T., S. D. Uysal, and J. M. Wooldridge (2022). "Doubly Robust Estimation of Local Average Treatment Effects Using Inverse Probability Weighted Regression Adjustment." doi:10.48550/arXiv.2208.01300

Donald, S. G., Y.-C. Hsu, and R. P. Lieli (2014). "Testing the Unconfoundedness Assumption via Inverse Probability Weighted Estimators of (L)ATT." *Journal of Business & Economic Statistics* 32(3), 395-415.

Examples

```
d <- drlate_sim
d$nvstat[d$rsncode == 0] <- 0L   # impose one-sided noncompliance
dr_hausman(lwage ~ age + educ, nvstat ~ age + educ,
           rsncode ~ age + educ, data = d)
```

`drlate` *Doubly robust estimation of the LATE and LATT*

Description

Estimates the local average treatment effect (LATE) or the local average treatment effect on the treated (LATT) with a binary instrument, following Słoczyński, Uysal, and Wooldridge (2022). A faithful R port of the Stata package `drlate` (SSC S459708): point estimates come from sequential weighted regressions, and standard errors are computed jointly for the instrument propensity score, the outcome regression, the treatment regression, and the causal estimand by stacking all moment conditions into a single M-estimation system.

Usage

```
drlate(
  outcome,
  treatment,
  instrument,
  data,
  omodel = c("linear", "logit", "probit", "poisson", "flogit", "fprobit"),
  tmodel = c("logit", "probit", "linear", "poisson"),
  ivmodel = c("logit", "cbps", "ipt", "probit"),
  method = c("ipwra", "ipw", "aipw", "ra", "kappa", "kappa0", "kappa10"),
  estimand = c("late", "latt"),
  normalized = TRUE,
  weights = NULL,
  cluster = NULL,
  pstolerance = 1e-05,
  osample = FALSE,
  subset = NULL,
  keep_data = TRUE,
  vcov = c("analytic", "bootstrap"),
  boot_reps = 999L,
  boot_seed = NULL,
  cores = 1L
)
```

Arguments

outcome	A formula y ~ covariates for the outcome model. Use y ~ 1 for no covariates (required when method = "ipw").
treatment	A formula d ~ covariates for the treatment model.
instrument	A formula z ~ covariates for the instrument propensity score model; z must be binary 0/1. Use z ~ 1 when method = "ra".
data	A data frame containing all variables.
omodel	Outcome model family: "linear" (default; continuous), "logit" or "probit" (outcome must be 0/1), "poisson" (outcome must be non-negative), or "flogit" / "fprobit" (fractional outcome in [0, 1], e.g. a proportion). The f-prefixed families share all estimation with "logit" / "probit" and only relax the response to the unit interval, matching the **Stata** lateffects omodel options.
tmodel	Treatment model family: "logit" (default; treatment must be 0/1), "probit", "linear", or "poisson".
ivmodel	Instrument propensity score model: "logit" (maximum likelihood; default), "cbps" (covariate balancing, Imai and Ratkovic 2014; not available with estimand = "latt"), "ipt" (inverse probability tilting, Graham, Pinto, and Egel 2012), or "probit" (maximum likelihood; mirrors kappalate's zmodel(probit) and is available only for the weighting estimators that command covers — "ipw", "kappa", "kappa0", "kappa10" — with estimand = "late").
method	Estimator: "ipwra" (inverse-probability-weighted regression adjustment; default), "ipw", "aipw", "ra", or one of the kappa-weighting estimators of Słoczyński, Uysal, and Wooldridge (2025): "kappa" (unnormalized Abadie kappa; kappalate's tau_a), "kappa0" (untreated-arm kappa; tau_a,0), or "kappa10" (normalized kappa; tau_a,10). The kappa estimators require intercept-only outcome and treatment formulas, a binary treatment, and estimand = "late"; ivmodel = "cbps" is available for "kappa" only, and "ipt" for none of them. drlate's normalized and unnormalized "ipw" coincide with kappalate's tau_u and tau_a,1.
estimand	"late" (default) or "latt".
normalized	Logical; use normalized moment conditions (default TRUE). Only relevant for method = "ipw" and method = "aipw".
weights	Optional sampling weights (a numeric vector, or a column name in data given as a string).
cluster	Optional cluster identifier for clustered standard errors (a vector, or a column name in data given as a string).
pstolerance	Overlap tolerance: estimation stops with an error if any estimated instrument propensity score is below pstolerance or above 1 - pstolerance. Default 1e-5.
osample	Logical; if TRUE, overlap violations do not stop estimation with an error. Instead drlate() returns (invisibly) a logical vector marking the violating observations.
subset	Optional logical or integer vector selecting rows of data.
keep_data	Logical; retain the internal estimation context (model matrices, fitted propensity scores, weights) on the returned object (default TRUE). Required by plot.drlate(), balance(), and the bootstrap; set to FALSE for a leaner object.

vcov	"analytic" (default) for the joint M-estimation sandwich replicating the Stata package, or "bootstrap" for nonparametric bootstrap standard errors and percentile confidence intervals (whole clusters are resampled when cluster is supplied). The analytic variance is always computed and stored either way. Draws that fail (degenerate resamples, non-convergence, overlap violations) are dropped and counted; because such failures concentrate where identification is weak, a non-trivial failure rate is itself a sign that percentile intervals are unreliable and the Fieller set (confint(., method = "fieller")) should be preferred.
boot_reps	Number of bootstrap replications (default 999).
boot_seed	Optional seed for reproducible bootstrap draws. Results are reproducible for a fixed number of cores; serial and parallel runs use different (both valid) random streams.
cores	Number of CPU cores for the bootstrap (default 1). Values above 1 use a PSOCK cluster and require the package to be installed (not merely loaded with devtools::load_all()).

Value

An object of class "drlate", a list with components including coefficients (the causal estimate, the numerator effect of Z on Y, and the denominator effect of Z on D), vcov3 (their variance matrix, diagonal by construction, as in the Stata package), vcov_full (the joint variance matrix of all stacked parameters), theta (all stacked parameter estimates), N, dmeanz1, dmeanz0, and the call. For method = "kappa10" only the causal estimate is reported (the estimator is a difference of two ratios, so no single numerator/denominator pair exists). For "kappa" and "kappa0" the third coefficient is the mean of the corresponding kappa weight: under the LATE assumptions it estimates the same complier share as the IPW first-stage contrast (the population ATE of Z on D), but it is a different sample statistic and the two can diverge under propensity score misspecification.

References

Słoczyński, T., S. D. Uysal, and J. M. Wooldridge (2022). "Doubly Robust Estimation of Local Average Treatment Effects Using Inverse Probability Weighted Regression Adjustment." doi:10.48550/arXiv.2208.01300

Słoczyński, T., S. D. Uysal, and J. M. Wooldridge (2025). "Abadie's Kappa and Weighting Estimators of the Local Average Treatment Effect." *Journal of Business & Economic Statistics* 43(1), 164–177. doi:10.1080/07350015.2024.2332763

Examples

```
data(drlate_sim)
fit <- drlate(lwage ~ age + educ, nvstat ~ age + educ,
              rsncode ~ age + educ, data = drlate_sim)
summary(fit)
```

drlate_compare *Compare drlate estimators in one call*

Description

Runs several estimators on the same specification and collects the causal estimates with their confidence intervals — the sensitivity comparison applied papers routinely report. Formula restrictions are handled automatically: `method = "ipw"` drops the outcome/treatment covariates and `method = "ra"` drops the instrument covariates (each with a message), matching the requirements of those estimators.

Usage

```
drlate_compare(
  outcome,
  treatment,
  instrument,
  data,
  methods = c("ipwra", "ipw", "aipw", "ra"),
  both_norms = FALSE,
  ...
)
```

Arguments

`outcome`	A formula `y ~ covariates` for the outcome model. Use `y ~ 1` for no covariates (required when `method = "ipw"`).
`treatment`	A formula `d ~ covariates` for the treatment model.
`instrument`	A formula `z ~ covariates` for the instrument propensity score model; `z` must be binary 0/1. Use `z ~ 1` when `method = "ra"`.
`data`	A data frame containing all variables.
`methods`	Estimators to run (any of the `method` values accepted by `drlate()`).
`both_norms`	Logical; also run the unnormalized variants of `"ipw"` and `"aipw"` (default `FALSE`).
`...`	Passed on to `drlate()` (e.g. `omodel`, `tmodel`, `ivmodel`, `estimand`, `weights`, `cluster`).

Details

Because IPW carries no outcome/treatment regressions and RA carries no instrument propensity score, the automatic formula adjustment means the rows do not share a single adjustment specification: differences between the IPW or RA row and the doubly robust rows reflect both the estimator *and* the reduced specification. Read the comparison as a robustness display, not as a test that isolates estimator choice; the doubly robust rows (IPWRA, AIPW) are the like-for-like pair.

Value

An object of class "drlate_compare": a data frame with columns method, normalized, estimate, se, ci_lo, ci_hi, with a print method and a dot-whisker plot method.

Examples

```
cmp <- drlate_compare(lwage ~ age + educ, nvstat ~ age + educ,
                      rsncode ~ age + educ, data = drlate_sim)
cmp
```

---

drlate_sim *Simulated example data for drlate*

Description

A simulated dataset with a binary instrument, a binary treatment with two-sided noncompliance, and continuous, positive, and binary outcome variables, designed to exercise every model family supported by drlate(). The complier average treatment effect (LATE) used in the data-generating process is 0.5. The treatment is genuinely endogenous (compliance type shifts the baseline outcome, so naive OLS is biased upward) and the instrument is only conditionally valid (its propensity depends on age and educ, so the raw Wald ratio is biased too).

Usage

```
drlate_sim
```

Format

A data frame with 2,000 rows and 7 variables:

lwage continuous outcome

kwage positive outcome (for Poisson models), exp(lwage / 2)

hijob binary outcome (for logit models)

nvstat binary treatment

rsncode binary instrument

age continuous covariate

educ factor covariate with levels hs, college, graduate

Source

Simulated; see data-raw/drlate_sim.R in the package sources.

kappa_weights *Abadie's kappa weights*

Description

Returns the per-observation Abadie kappa weight implied by a fitted `drlate()` object,

$$\kappa = 1 - \frac{D(1-Z)}{1-p(X)} - \frac{(1-D)Z}{p(X)},$$

where $p(X)$ is the estimated instrument propensity score. The kappa weights identify the complier subpopulation: for any function g of the data, $E[g \mid \text{complier}] = E[\kappa g]/E[\kappa]$ (Abadie 2003). They are the weights used by `complier_means()` and are the Stata `estat compliers, genkappa()` object.

Usage

```
kappa_weights(object, normalize = TRUE)
```

Arguments

`object`	A fitted `drlate()` object (with `keep_data = TRUE`) using an instrument propensity score (any `method` except `"ra"`).
`normalize`	Logical. If `TRUE` (default), the returned weights are the sampling-weighted, normalized weights $w\kappa/\sum w\kappa$ that sum to one (the form used to compute complier averages). If `FALSE`, the raw kappa values are returned.

Value

A numeric vector with one entry per estimation-sample observation.

See Also

`complier_means()`

Examples

```
fit <- drlate(lwage ~ age + educ, nvstat ~ age + educ,
              rsncode ~ age + educ, data = drlate_sim)
head(kappa_weights(fit))
```

`plot.drlate` *Diagnostic plots for drlate fits*

Description

Diagnostic plots for drlate fits

Usage

```
## S3 method for class 'drlate'
plot(
  x,
  type = c("overlap", "balance", "balance_density", "weights"),
  bins = 30,
  geom = c("histogram", "density"),
  var = NULL,
  ...
)
```

Arguments

`x`	A fitted `drlate()` object (with `keep_data = TRUE`).
`type`	One of: • `"overlap"`: histograms (or kernel densities, see `geom`) of the estimated instrument propensity score by instrument arm, with the `pstolerance` bounds marked. Mass piling up near 0 or 1 signals overlap problems. • `"balance"`: a love plot of standardized mean differences from `balance()`, unweighted vs IPW-weighted, with the conventional |SMD| = 0.1 reference lines. • `"balance_density"`: kernel densities of the covariates by instrument arm, raw versus IPW-weighted (the Stata `latebalance density` display). Weighting that balances a covariate brings the two arm densities together in the weighted panel. • `"weights"`: distributions of the implied IPW weights by arm; a long right tail means a few observations dominate the estimate.
`bins`	Number of histogram bins for `"overlap"` and `"weights"`.
`geom`	For `type = "overlap"`, either `"histogram"` (default) or `"density"` (a kernel-density overlap matching Stata `lateoverlap`).
`var`	For `type = "balance_density"`, an optional character vector selecting covariates to plot; defaults to all model covariates.
`...`	Currently unused.

Value

A `ggplot` object.

Index

∗ **datasets**
drlate_sim, 12

balance, 2
balance(), *4*, *9*, *14*
balance_test, 3

complier_means, 4
complier_means(), *13*
confint.drlate, 5

dr_hausman, 6
drlate, 8
drlate(), *3–5*, *11–14*
drlate_compare, 11
drlate_sim, 12

kappa_weights, 13
kappa_weights(), *4, 5*

plot.drlate, 14
plot.drlate(), *3*, *9*