

# Package: clusterIV (via r-universe)

July 1, 2026

**Title** Clustered Jackknife Instrumental Variables Estimation

**Version** 0.1.0

**Description** Tools for instrumental variables estimation and inference under clustered errors with many instruments. The current release provides the cluster-jackknife IV estimator (CJIVE) of Frandsen, Leslie and McIntyre (2025) <[doi:10.1162/rest.a.263](https://doi.org/10.1162/rest.a.263)> for a single endogenous regressor in a just-identified design, with cluster-robust inference: each observation's first-stage value is fitted leaving out its entire cluster, which removes the many-instrument bias that survives clustering. The leave-cluster-out fits use an exact Woodbury block update -- one factorisation of the instrument Gram matrix plus a small solve per cluster -- so the estimator scales to large samples. A companion 'iv\_compare()' reports ordinary least squares, two-stage least squares, the observation-level jackknife and CJIVE on a common cluster-robust standard error.

**License** MIT + file LICENSE

**Encoding** UTF-8

**Imports** stats

**RoxygenNote** 7.3.1

**URL** <https://github.com/atal-kat/Clustered-Estimation-and-Inference>

**BugReports** <https://github.com/atal-kat/Clustered-Estimation-and-Inference/issues>

**Repository** <https://atal-kat.r-universe.dev>

**Date/Publication** 2026-06-24 15:41:25 UTC

**RemoteUrl** <https://github.com/atal-kat/clustered-estimation-and-inference>

**RemoteRef** HEAD

**RemoteSha** 8747547f84ee1fb876d9415d0384f07b4c58edea

## Contents

cjive . . . . .	2
iv_compare . . . . .	5

---

cjive	<i>Cluster-jackknife instrumental variables estimation (CJIVE)</i>
-------	--

---

### Description

Computes the cluster-jackknife IV estimator of Frandsen, Leslie and McIntyre (2025) for a single endogenous regressor in a just-identified design, with cluster-robust inference. The first-stage value for each observation is fitted from a regression that leaves out the observation's entire cluster, which removes the many-instrument bias that survives clustering.

### Usage

```

cjive(y, ...)

## Default S3 method:
cjive(
  y,
  x,
  z,
  cluster,
  controls = NULL,
  weights = NULL,
  level = 0.95,
  intercept = TRUE,
  method = c("auto", "dense", "leaveout_mean"),
  ...
)

## S3 method for class 'formula'
cjive(
  formula,
  data,
  cluster,
  controls = NULL,
  weights = NULL,
  level = 0.95,
  intercept = TRUE,
  method = c("auto", "dense", "leaveout_mean"),
  ...
)

## S3 method for class 'cjive'
print(x, digits = max(3L, getOption("digits") - 3L), ...)

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

```

```
## S3 method for class 'summary.cjive'
print(x, digits = max(3L, getOption("digits") - 3L), ...)

## S3 method for class 'cjive'
coef(object, ...)

## S3 method for class 'cjive'
vcov(object, ...)

## S3 method for class 'cjive'
confint(object, parm, level = 0.95, ...)
```

### Arguments

y	Outcome (numeric vector), or a two-sided formula $y \sim x \mid z$ for the formula method (the bar separates the endogenous regressor from the instruments).
...	Unused.
x	Single endogenous regressor (numeric vector); or, for the print methods, a fitted "cjive" object.
z	Instruments: a numeric vector/matrix, or a factor/character grouping vector for a judge design (expanded internally to a dummy design with one reference level dropped, the intercept supplying the rest).
cluster	Cluster identifiers (length n). For the formula method a one-sided formula ( $\sim g$ ) or a column name is also accepted.
controls	Optional covariates (FLM's $X$ ): a matrix or data frame, or a one-sided formula in the formula method. May be rank deficient (fixed effects are allowed). An intercept is added unless <code>intercept = FALSE</code> .
weights	Optional strictly positive precision weights.
level	Confidence level for the reported interval.
intercept	Logical; partial out an intercept (default TRUE).
method	One of "auto", "dense", "leaveout_mean". "auto" and "dense" both use the dense Frisch-Waugh-Lovell block-jackknife and are the default. "leaveout_mean" evaluates FLM's printed leave-cluster-out group-mean form and is available only for a grouping-factor z with intercept-only controls; it differs from the default by an $O(1/n_g)$ intercept term and is never selected automatically.
formula	A formula $y \sim x \mid z_1 + z_2$ .
data	A data frame in which to evaluate the formula.
object	A fitted "cjive" object.
digits	Number of significant digits to print.
parm	Ignored (a single coefficient is estimated).

## Details

The estimator is the covariance ratio  $\hat{\delta} = \widehat{Cov}(Y, \hat{p}) / \widehat{Cov}(D, \hat{p})$  with the cluster-jackknife constructed instrument  $\hat{p}$ . Covariates are handled by Frisch-Waugh-Lovell:  $Y$ ,  $D$  and each instrument are residualised on the covariates (with an intercept) once, up front, then the estimator runs on the residuals. This dense route is the single convention everywhere, so `cjive()` and `iv_compare` return the identical CJIVE for any design. The leave-cluster-out fits are computed by a Woodbury block update (one Cholesky of  $Z'Z$  plus a small solve per cluster), exact against the brute-force definition, and collapsing to observation-level JIVE when every cluster is a singleton.

## Value

An object of class "cjive": a list with `coefficient`, `se`, `statistic`, `p.value`, `conf.low`, `conf.high`, `level`, the diagnostics `n`, `G`, `p`, `path` ("dense" or "leaveout\_mean") and `maxlev` (the maximum within-cluster leverage  $\max_g \lambda_{\max}(H_g)$ , a conditioning diagnostic; NA on the mean path), and the `call`.

## Methods (by class)

- `cjive(default)`: matrix/vector interface.
- `cjive(formula)`: formula interface  $y \sim x \mid z$ .
- `print(cjive)`: Print a concise summary of the fit.
- `summary(cjive)`: Build a summary object.
- `print(summary.cjive)`: Print method for the summary object.
- `coef(cjive)`: Extract the point estimate.
- `vcov(cjive)`: Extract the cluster-robust (co)variance.
- `confint(cjive)`: Confidence interval for the coefficient.

## References

Frandsen, B., Leslie, E. and McIntyre, S. (2025). Cluster Jackknife Instrumental Variables Estimation. *Review of Economics and Statistics*.

## Examples

```
set.seed(1)
G <- 40; ng <- 6; n <- G * ng
cl <- rep(seq_len(G), each = ng)
j <- factor(rep(rep(1:4, length.out = ng), G)) # judge identity
u <- rnorm(G)[cl]
x <- as.numeric(j) + u + rnorm(n)
y <- 1.5 * x + u + rnorm(n)
fit <- cjive(y, x, j, cluster = cl)
print(fit)

## formula interface
dat <- data.frame(y = y, x = x, j = j, cl = cl)
cjive(y ~ x | j, data = dat, cluster = ~cl)
```

---

iv_compare	<i>Compare IV estimators on a common cluster-robust SE (FLM Table 1)</i>
------------	--

---

### Description

Returns OLS, 2SLS, JIVE and CJIVE for the same design, each reported with the *identical* just-identified cluster-robust IV sandwich standard error; only the constructed instrument differs between rows. This reproduces the shape of Table 1 in Frandsen, Leslie and McIntyre (2025).

### Usage

```
iv_compare(
  y,
  x,
  z,
  cluster,
  controls = NULL,
  weights = NULL,
  level = 0.95,
  intercept = TRUE
)
```

### Arguments

<code>y</code>	Outcome (numeric vector).
<code>x</code>	Single endogenous regressor (numeric vector).
<code>z</code>	Instruments (numeric matrix/vector or a grouping factor).
<code>cluster</code>	Cluster identifiers (length $n$ ). For the formula method a one-sided formula ( $\sim g$ ) or a column name is also accepted.
<code>controls</code>	Optional covariates (FLM's $X$ ): a matrix or data frame, or a one-sided formula in the formula method. May be rank deficient (fixed effects are allowed). An intercept is added unless <code>intercept = FALSE</code> .
<code>weights</code>	Optional strictly positive precision weights.
<code>level</code>	Confidence level for the reported interval.
<code>intercept</code>	Logical; partial out an intercept (default TRUE).

### Details

The constructed instruments are: OLS, the residualised  $x$  itself; 2SLS, the full-sample fit  $Z\hat{\pi}$ ; JIVE, the leave-one-out fit  $(\hat{x} - hx)/(1 - h)$ ; CJIVE, the leave-cluster-out block fit. The CJIVE row equals `cjive(..., method = "dense")` on the same design.

### Value

A data frame with one row per estimator (in the order OLS, 2SLS, JIVE, CJIVE) and columns `estimator`, `coefficient`, `se`, `statistic`, `p.value`, `conf.low`, `conf.high`.

**References**

Frandsen, B., Leslie, E. and McIntyre, S. (2025). Cluster Jackknife Instrumental Variables Estimation. *Review of Economics and Statistics*.

**Examples**

```
set.seed(2)
G <- 50; ng <- 5; n <- G * ng
cl <- rep(seq_len(G), each = ng)
z <- matrix(rnorm(n * 3), n, 3)
u <- rnorm(G)[cl]
x <- z %>% c(1, -1, 0.5) + u + rnorm(n)
y <- 2 * x + u + rnorm(n)
iv_compare(y, x, z, cluster = cl)
```

# Index

`cjive`, [2](#)  
`coef.cjive (cjive)`, [2](#)  
`confint.cjive (cjive)`, [2](#)  
  
`iv_compare`, [4, 5](#)  
  
`print.cjive (cjive)`, [2](#)  
`print.summary.cjive (cjive)`, [2](#)  
  
`summary.cjive (cjive)`, [2](#)  
  
`vcov.cjive (cjive)`, [2](#)