

# Package ‘civ’

May 8, 2026

**Title** Categorical Instrumental Variables

**Version** 0.1.0

**Date** 2023-12-07

**Description** Implementation of the categorical instrumental variable (CIV) estimator proposed by Wiemann (2023) <[doi:10.48550/arXiv.2311.17021](https://doi.org/10.48550/arXiv.2311.17021)>. CIV allows for optimal instrumental variable estimation in settings with relatively few observations per category. To obtain valid inference in these challenging settings, CIV leverages a regularization assumption that implies existence of a latent categorical variable with fixed finite support achieving the same first stage fit as the observed instrument.

**License** GPL (>= 3)

**URL** <https://github.com/thomaswiemann/civ>

**BugReports** <https://github.com/thomaswiemann/civ/issues>

**Encoding** UTF-8

**RoxygenNote** 7.2.3

**Depends** R (>= 3.6)

**Imports** stats, AER, kcmeans

**Suggests** testthat (>= 3.0.0), covr, knitr, rmarkdown

**Config/testthat/edition** 3

**VignetteBuilder** knitr

**NeedsCompilation** no

**Author** Thomas Wiemann [aut, cre]

**Maintainer** Thomas Wiemann <[wiemann@uchicago.edu](mailto:wiemann@uchicago.edu)>

**Repository** CRAN

**Date/Publication** 2023-12-08 11:00:06 UTC

## Contents

civ . . . . .	2
summary.civ . . . . .	3

civ

*Categorical Instrumental Variable Estimator.***Description**

Implementation of the categorical instrumental variable estimator.

**Usage**

```
civ(y, D, Z, X = NULL, K = 2)
```

**Arguments**

y	The outcome variable, a numerical vector.
D	A matrix of endogenous variables.
Z	A matrix of instruments, where the first column corresponds to the categorical instrument.
X	An optional matrix of control variables.
K	The number of support points of the estimated instrument $\hat{m}_K$ , an integer greater than 2.

**Value**

civ returns an object of S3 class `civ`. An object of class `civ` is a list containing the following components:

`coef` A vector of second-stage coefficient estimates.

`iv_fit` Object of class `ivreg` from the IV regression of `y` on `D` and `X` using the the estimated  $\hat{F}_K$  as an instrument for `D`. See also [AER::ivreg\(\)](#) for details.

`kcmeans_fit` Object of class `kcmeans` from the K-Conditional-Means regression of `D` on `Z` and `X`. See also [kcmeans::kcmeans\(\)](#) for details.

**K** Pass-through of selected user-provided arguments. See above.

**References**

Fox J, Kleiber C, Zeileis A (2023). "ivreg: Instrumental-Variables Regression by '2SLS', '2SM', or '2SMM', with Diagnostics". R package.

Wiemann T (2023). "Optimal Categorical Instruments."

**Examples**

```
# Simulate data from a simple IV model with 800 observations
nobs = 800 # sample size
Z <- sample(1:20, nobs, replace = TRUE) # observed instrument
Z0 <- Z %% 2 # underlying latent instrument
U_V <- matrix(rnorm(2 * nobs, 0, 1), nobs, 2) %*%
  chol(matrix(c(1, 0.6, 0.6, 1), 2, 2)) # first and second stage errors
D <- Z0 + U_V[, 2] # endogenous variable
y <- D + U_V[, 1] # outcome variable
# Estimate categorical instrument variable estimator with K = 2
civ_fit <- civ(y, D, Z, K = 3)
summary(civ_fit)
```

summary.civ

*Inference Methods for the Categorical Instrumental Variable Estimator.*

**Description**

Inference methods for the categorical instrumental variable estimators. Simple wrapper for [AER::summary.ivreg\(\)](#).

**Usage**

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

**Arguments**

`object` An object of class `civ` as fitted by [civ\(\)](#).

`...` Additional arguments passed to [summary.ivreg](#). See [AER::summary.ivreg\(\)](#) for a complete list of arguments.

**Value**

An object of class `summary.ivreg` with inference results.

**References**

Fox J, Kleiber C, Zeileis A (2023). "ivreg: Instrumental-Variables Regression by '2SLS', '2SM', or '2SMM', with Diagnostics". R package.

Wiemann T (2023). "Optimal Categorical Instruments."

**See Also**

[AER::summary.ivreg\(\)](#)

**Examples**

```
# Simulate data from a simple IV model with 800 observations
nobs = 800 # sample size
Z <- sample(1:20, nobs, replace = TRUE) # observed instrument
Z0 <- Z %% 2 # underlying latent instrument
U_V <- matrix(rnorm(2 * nobs, 0, 1), nobs, 2) %*%
  chol(matrix(c(1, 0.6, 0.6, 1), 2, 2)) # first and second stage errors
D <- Z0 + U_V[, 2] # endogenous variable
y <- D + U_V[, 1] # outcome variable
# Estimate categorical instrument variable estimator with K = 2
civ_fit <- civ(y, D, Z, K = 3)
summary(civ_fit)
```

# Index

AER::ivreg(), 2  
AER::summary.ivreg(), 3  
  
civ, 2  
civ(), 3  
  
kcmeans::kcmeans(), 2  
  
summary.civ, 3