

# Package ‘fejiv’

October 14, 2025

**Type** Package

**Title** Fixed Effect Jackknife Instrumental Variables Estimation

**Version** 0.1.1

**Description** Implements the Fixed Effect Jackknife Instrumental Variables ('FEJIV') estimator of Chao, Swanson, and Woutersen (2023) <[doi:10.1016/j.jeconom.2022.12.011](https://doi.org/10.1016/j.jeconom.2022.12.011)>, allowing consistent IV estimation with many (possibly weak) instruments, cluster fixed effects, heteroskedastic errors, and many exogenous covariates. The estimator is recommended by Słoczyński (2024) <[doi:10.48550/arXiv.2011.06695](https://doi.org/10.48550/arXiv.2011.06695)> as an alternative to two-stage least squares when estimating the interacted specification of Angrist and Imbens (1995) <[doi:10.1080/01621459.1995.10476535](https://doi.org/10.1080/01621459.1995.10476535)>.

**License** GPL (>= 2)

**Encoding** UTF-8

**Depends** R (>= 4.0)

**Imports** MASS, Matrix, stats

**Suggests** haven

**RoxygenNote** 7.3.3

**NeedsCompilation** no

**Author** Qihui Lei [aut, cre],  
Tymon Słoczyński [aut]

**Maintainer** Qihui Lei <qlei9@wisc.edu>

**Repository** CRAN

**Date/Publication** 2025-10-14 18:20:15 UTC

## Contents

|       |   |
|-------|---|
| fejiv | 2 |
|-------|---|

|       |   |
|-------|---|
| Index | 5 |
|-------|---|

`fejiv`*Fixed effect jackknife IV ('FEJIV') estimation*

## Description

`fejiv` implements the fixed effect jackknife IV (FEJIV) estimator of Chao, Swanson, and Woutersen (2023), which enables consistent IV estimation with many (possibly weak) instruments, cluster fixed effects, heteroskedastic errors, and possibly many exogenous explanatory variables.

## Usage

```
fejiv(Y, D, Z, X = NULL, absorb = NULL)
```

## Arguments

|                     |  |
|---------------------|--|
| <code>Y</code>      | The dependent variable (numeric vector).   |
| <code>D</code>      | The endogenous explanatory variable (numeric vector or matrix with one column).  |
| <code>Z</code>      | The instrumental variables (numeric vector, matrix, or data frame).  |
| <code>X</code>      | Optional exogenous explanatory variables (numeric vector, matrix, or data frame; Default <code>NULL</code> .).   |
| <code>absorb</code> | Optional categorical variable to be absorbed (vector or factor; Default <code>NULL</code> .). All categories should have a frequency of three or more. |

## Details

Consistency of the FEJIV estimator requires that instrument strength satisfies a key growth condition: the concentration parameter must grow faster than the square root of the number of instruments. Mikusheva and Sun (2022) show that this condition is necessary for the existence of a consistent test and also propose a test of this condition, implemented in the Stata command `manyweakivpretest`, available at Liyang Sun's GitHub.

Słoczyński (2024) recommends the FEJIV estimator as an alternative to two-stage least squares (2SLS) when estimating the fully interacted specification of Angrist and Imbens (1995). Within the local average treatment effect (LATE) framework, when strong monotonicity is doubtful but weak monotonicity is plausible, the fully interacted specification eliminates the problem of "negative weights."

This is a companion software package for Słoczyński (2024). If you use it, please cite both Słoczyński (2024) and Chao, Swanson, and Woutersen (2023).

## Value

A list of class "fejiv" with elements:

|                          |  |
|--------------------------|--|
| <code>coefficient</code> | Coefficient on the endogenous regressor. |
| <code>vcov</code>        | Estimated variance.                      |

|                    |                                  |
|--------------------|----------------------------------|
| <code>se</code>    | Standard error.                  |
| <code>N</code>     | Number of observations.          |
| <code>treat</code> | Name of the endogenous variable. |
| <code>call</code>  | Matched function call.           |
| <code>title</code> | Character string for printing.   |

## Acknowledgments

This command is based on MATLAB code for the estimators of Chao, Swanson, and Woutersen (2023), generously shared by Tiemen Woutersen.

## License

This package is licensed under the MIT License. See the LICENSE file included with the distribution.

## Author(s)

Qihui Lei, University of Wisconsin, Email: <qlei9@wisc.edu> Tymon Słoczyński, Brandeis University, Email: <tslocz@brandeis.edu>

## References

- Angrist, Joshua D., and Guido W. Imbens (1995). Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity. *Journal of the American Statistical Association*, 90(430), 431–442.
- Chao, John C., Norman R. Swanson, and Tiemen Woutersen (2023). Jackknife Estimation of a Cluster-Sample IV Regression Model with Many Weak Instruments. *Journal of Econometrics*, 235(2), 1747–1769.
- Mikusheva, Anna, and Liyang Sun (2022). Inference with Many Weak Instruments. *Review of Economic Studies*, 89(5), 2663–2686.
- Słoczyński, Tymon (2024). When Should We (Not) Interpret Linear IV Estimands as LATE? arXiv:2011.06695. <https://arxiv.org/abs/2011.06695>.

## Examples

```
# one fast example for demonstration
set.seed(2025)
n <- 100
Y <- rnorm(n)
D <- rnorm(n)
Z <- matrix(rnorm(n * 5), n, 5)

# Basic usage - no fixed effects
result <- fejiv(Y, D, Z)
print(result)

# Example with fixed effects
```

```
absorb_var <- rep(1:10, each = 10)
result_fe <- fejiv(Y, D, Z, absorb = absorb_var)
print(result_fe)

# -----
# Example: Revisiting Card (1995) using fejiv
# -----
# Realistic example with a larger sample and fixed effects
# This takes longer due to the computational complexity of fejiv

if (requireNamespace("haven", quietly = TRUE)) {
  library(haven)

  # Load data directly from Tymon Słoczyński's GitHub
  data <- read_dta("https://tslocz.github.io/card.dta")

  # Create a college dummy
  data$college <- as.numeric(data$educ > 12)

  # Construct cluster groups following Słoczyński (2024)
  data$group <- interaction(
    data$black, data$smsa, data$smsa66, data$south, data$south66
  )

  # Drop clusters with fewer than 3 observations
  data$gsize <- ave(rep(1, nrow(data)), data$group, FUN = length)
  data <- data[data$gsize >= 3, ]

  # Run Fixed Effect Jackknife IV (FEJIV) regression
  # Instruments: nearc4 interacted with cluster group (no main effects)
  model <- fejiv(
    Y      = data$lwage,
    D      = data$college,
    Z      = model.matrix(~ nearc4:factor(group) - 1, data = data),
    absorb = data$group
  )

  print(model)
}
```

# **Index**

[fejiv, 2](#)