

# Package ‘ManyIV’

February 17, 2021

**Title** Inference in Instrumental Variables Models with Many Instruments

**Version** 0.0.2

**Description** Calculate estimators and standard errors in a linear instrumental variables model with possibly many instruments and many exogenous regressors under homoskedasticity of the reduced form. Also computes tests of overidentifying restrictions.

**Depends** R (>= 4.0.0)

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**Imports** Matrix,  
stats,  
Formula,  
numDeriv

**RoxygenNote** 7.1.1

**Suggests** spelling,  
testthat,  
knitr,  
rmarkdown

**VignetteBuilder** knitr

**LazyLoad** yes

**Language** en-US

**BugReports** <https://github.com/kolesarm/ManyIV/issues>

**URL** <https://github.com/kolesarm/ManyIV>

## R topics documented:

ak80 . . . . .	2
IVData . . . . .	3
IVoverid . . . . .	3
IVreg . . . . .	4
IVreg.fit . . . . .	5

---

ak80*Angrist-Krueger (1991) census data*

---

**Description**

Sample of males born in the US in 1930-39 from 5

**Usage**

ak80

**Format**

A data frame with 329,509 observations on 10 variables:

**age** Age, measured at quarterly precision

**education** Years of education

**lwage** Log of weekly wage

**married** Indicator for being married

**qob** Quarter of birth

**sob** State of birth (FIPS code)

**black** Indicator for being black

**smsa** SMSA indicator

**yob** Year of birth

**division** Factor variable indicating division

**Source**

Josh Angrist's website, <http://economics.mit.edu/faculty/angrist/data1/data/angkru1991>

---

IVData	<i>Class constructor for IVData</i>
--------	-------------------------------------

---

### Description

Convert data to standardized format for use with low-level functions. Uses *Matrix* package, which speeds up calculations.

### Usage

```
IVData(Y, X, Z, W, moments = TRUE, approx = TRUE)
```

### Arguments

Y	n-vector
X	n-vector
Z	[n x k] Matrix of instruments, class <i>Matrix</i>
W	[n x ell] Matrix of covariates, class <i>Matrix</i>
moments	if TRUE, compute estimates of third and fourth moments of the reduced-form errors based on least squares residuals
approx	if TRUE, then estimates of third and fourth moments use an approximation to speed up the calculations.

---

IVoverid	<i>Test of overidentifying restrictions</i>
----------	---

---

### Description

Report test statistic and p-value for testing of overidentifying restrictions. The Sargan test is valid under few instruments. The Modified Cragg-Donald test (Modified-CD) corresponds to a test due to Cragg and Donald (1993), with a critical value modified to make it robust to many instruments and many exogenous regressors.

### Usage

```
IVoverid(r)
```

### Arguments

r	object of class <i>RDResults</i>
---	----------------------------------

IVreg

*Fit instrumental-variable regression***Description**

Fit instrumental-variable regression

**Usage**

```
IVreg(formula, data, subset, na.action, inference = "standard", approx = TRUE)
```

**Arguments**

formula	specification of the regression relationship and the instruments of the form $y \sim x + w_1 + w_2 \mid z_1 + z_2 + z_3$ , where $y$ is the outcome variable, $x$ is a scalar endogenous variable, $w_1, w_2$ are exogenous regressors, and $z_1, z_2, z_3$ are excluded instruments.
data	optional data frame, list or environment (or object coercible by <code>as.data.frame</code> to a data frame) containing the outcome and running variables in the model. If not found in data, the variables are taken from <code>environment(formula)</code> , typically the environment from which the function is called.
subset	optional vector specifying a subset of observations to be used in the fitting process.
na.action	function which indicates what should happen when the data contain NAs. The default is set by the <code>na.action</code> setting of options (usually <code>na.omit</code> ).
inference	Vector specifying inference method(s). The elements of the vector can consist of the following methods: <b>"standard"</b> Report inference based on TSLS, LIML, and MBTSLS, along with homoskedastic and heteroskedasticity-robust standard errors valid under standard asymptotic sequence, as well as standard errors that are valid under heteroskedasticity and treatment effect heterogeneity. <b>"re"</b> Inference based on Hessian of random effects likelihood <b>"il"</b> Inference based on Hessian of invariant likelihood, evaluated numerically <b>"lil"</b> Inference based on information matrix of limited information likelihood <b>"md"</b> Inference based on the minimum distance objective function
approx	if TRUE, then estimates of third and fourth moments used in inference based on the minimum distance objective function ( <code>inference="md"</code> ) use an approximation to speed up the calculations.

**Examples**

```
## Specification as in Table V, columns (1) and (2) in Angrist and Krueger
IVreg(lwage~education+as.factor(yob)|as.factor(qob)*as.factor(yob),
      data=ak80, inference=c("standard", "re", "il", "lil"))
## Only quarter of birth as instrument, add married, black and smsa as
```

```
## exogenous regressors
IVreg(lwage~education+as.factor(yob)+black+smsa+married|as.factor(qob),
      data=ak80, inference=c("standard", "re", "il", "lil"))
```

IVreg.fit

*Low-level computing engine called by IVreg***Description**

Low-level computing engine called by IVreg

**Usage**

```
IVreg.fit(d, inference)
```

**Arguments**

d	Object of class "IVData"
inference	<p>Vector specifying inference method(s). The elements of the vector can consist of the following methods:</p> <p><b>"standard"</b> Report inference based on TSLS, LIML, and MBTSLS, along with homoskedastic and heteroskedasticity-robust standard errors valid under standard asymptotic sequence, as well as standard errors that are valid under heteroskedasticity and treatment effect heterogeneity.</p> <p><b>"re"</b> Inference based on Hessian of random effects likelihood</p> <p><b>"il"</b> Inference based on Hessian of invariant likelihood, evaluated numerically</p> <p><b>"lil"</b> Inference based on information matrix of limited information likelihood</p> <p><b>"md"</b> Inference based on the minimum distance objective function</p>

# Index

## \* **datasets**

ak80, [2](#)

ak80, [2](#)

IVData, [3](#)

IOverid, [3](#)

IVreg, [4](#)

IVreg.fit, [5](#)