Title

> **ivregress postestimation** — Postestimation tools for ivregress

Postestimation commands

The following postestimation commands are of special interest after `ivregress`:

Command	Description
estat endogenous	perform tests of endogeneity
estat firststage	report "first-stage" regression statistics
estat overid	perform tests of overidentifying restrictions
* estat sbknown	perform tests for a structural break with a known break date
* estat single	perform tests for a structural break with an unknown break date

These commands are not appropriate with svy estimation results.

*`estat sbknown` and `estat sbsingle` work only after `ivregress 2sls`.

The following postestimation commands are also available:

Command	Description
contrast	contrasts and ANOVA-style joint tests of estimates
estat summarize	summary statistics for the estimation sample
estat vce	variance–covariance matrix of the estimators (VCE)
estat (svy)	postestimation statistics for survey data
estimates	cataloging estimation results
etable	table of estimation results
† forecast	dynamic forecasts and simulations
† hausman	Hausman's specification test
lincom	point estimates, standard errors, testing, and inference for linear combinations of coefficients
margins	marginal means, predictive margins, marginal effects, and average marginal effects
marginsplot	graph the results from margins (profile plots, interaction plots, etc.)
nlcom	point estimates, standard errors, testing, and inference for nonlinear combinations of coefficients
predict	linear predictions and their SEs, probabilities, etc.
predictnl	point estimates, standard errors, testing, and inference for generalized predictions
pwcompare	pairwise comparisons of estimates
test	Wald tests of simple and composite linear hypotheses
testnl	Wald tests of nonlinear hypotheses

†`forecast` and `hausman` are not appropriate with svy estimation results.

predict

Description for predict

predict creates a new variable containing predictions such as linear predictions, residuals, standard errors, probabilities, and expected values.

Menu for predict

Statistics > Postestimation

Syntax for predict

> predict $\begin{bmatrix} type \end{bmatrix}$ *newvar* $\begin{bmatrix} if \end{bmatrix}$ $\begin{bmatrix} in \end{bmatrix}$ $\begin{bmatrix} , statistic \end{bmatrix}$

> predict $\begin{bmatrix} type \end{bmatrix}$ *stub** $\begin{bmatrix} if \end{bmatrix}$ $\begin{bmatrix} in \end{bmatrix}$, <u>sc</u>ores

statistic	Description
Main	
xb	linear prediction; the default
<u>residuals</u>	residuals
stdp	standard error of the prediction
stdf	standard error of the forecast
pr(*a*,*b*)	$\Pr(a < y_j < b)$ under exogeneity and normal errors
e(*a*,*b*)	$E(y_j \mid a < y_j < b)$ under exogeneity and normal errors
<u>ystar</u>(*a*,*b*)	$E(y_j^*)$, $y_j^* = \max\{a, \min(y_j, b)\}$ under exogeneity and normal errors

These statistics are available both in and out of sample; type predict ... if e(sample) ... if wanted only for the estimation sample.

stdf is not allowed with svy estimation results.

where *a* and *b* may be numbers or variables; *a* missing ($a \geq .$) means $-\infty$, and *b* missing ($b \geq .$) means $+\infty$; see [U] **12.2.1 Missing values**.

Options for predict

> Main

xb, the default, calculates the linear prediction.

residuals calculates the residuals, that is, $y_j - \mathbf{x}_j\mathbf{b}$. These are based on the estimated equation when the observed values of the endogenous variables are used—not the projections of the instruments onto the endogenous variables.

stdp calculates the standard error of the prediction, which can be thought of as the standard error of the predicted expected value or mean for the observation's covariate pattern. This is also referred to as the standard error of the fitted value.

stdf calculates the standard error of the forecast, which is the standard error of the point prediction for 1 observation. It is commonly referred to as the standard error of the future or forecast value. By construction, the standard errors produced by stdf are always larger than those produced by stdp; see *Methods and formulas* in [R] **regress postestimation**.

`pr(`*a*`,`*b*`)` calculates $\Pr(a < \mathbf{x}_j\mathbf{b} + u_j < b)$, the probability that $y_j|\mathbf{x}_j$ would be observed in the interval (a, b) under exogeneity and assuming errors are normally distributed.

> *a* and *b* may be specified as numbers or variable names; *lb* and *ub* are variable names;
> `pr(20,30)` calculates $\Pr(20 < \mathbf{x}_j\mathbf{b} + u_j < 30)$;
> `pr(`*lb*`,`*ub*`)` calculates $\Pr(lb < \mathbf{x}_j\mathbf{b} + u_j < ub)$; and
> `pr(20,`*ub*`)` calculates $\Pr(20 < \mathbf{x}_j\mathbf{b} + u_j < ub)$.

> *a* missing ($a \geq .$) means $-\infty$; `pr(.,30)` calculates $\Pr(-\infty < \mathbf{x}_j\mathbf{b} + u_j < 30)$;
> `pr(`*lb*`,30)` calculates $\Pr(-\infty < \mathbf{x}_j\mathbf{b} + u_j < 30)$ in observations for which $lb \geq .$
> and calculates $\Pr(lb < \mathbf{x}_j\mathbf{b} + u_j < 30)$ elsewhere.

> *b* missing ($b \geq .$) means $+\infty$; `pr(20,.)` calculates $\Pr(+\infty > \mathbf{x}_j\mathbf{b} + u_j > 20)$;
> `pr(20,`*ub*`)` calculates $\Pr(+\infty > \mathbf{x}_j\mathbf{b} + u_j > 20)$ in observations for which $ub \geq .$
> and calculates $\Pr(20 < \mathbf{x}_j\mathbf{b} + u_j < ub)$ elsewhere.

`e(`*a*`,`*b*`)` calculates $E(\mathbf{x}_j\mathbf{b} + u_j \mid a < \mathbf{x}_j\mathbf{b} + u_j < b)$, the expected value of $y_j|\mathbf{x}_j$ conditional on $y_j|\mathbf{x}_j$ being in the interval (a, b), meaning that $y_j|\mathbf{x}_j$ is truncated. *a* and *b* are specified as they are for `pr()`. Exogeneity and normally distributed errors are assumed.

`ystar(`*a*`,`*b*`)` calculates $E(y_j^*)$, where $y_j^* = a$ if $\mathbf{x}_j\mathbf{b} + u_j \leq a$, $y_j^* = b$ if $\mathbf{x}_j\mathbf{b} + u_j \geq b$, and $y_j^* = \mathbf{x}_j\mathbf{b} + u_j$ otherwise, meaning that y_j^* is censored. *a* and *b* are specified as they are for `pr()`. Exogeneity and normally distributed errors are assumed.

`scores` calculates the scores for the model. A new score variable is created for each endogenous regressor, as well as an equation-level score that applies to all exogenous variables and constant term (if present).

margins

Description for margins

margins estimates margins of response for linear predictions, probabilities, and expected values.

Menu for margins

Statistics > Postestimation

Syntax for margins

margins [*marginlist*] [, *options*]

margins [*marginlist*] , <u>predict</u>(*statistic* ...) [<u>predict</u>(*statistic* ...) ...] [*options*]

statistic	Description
xb	linear prediction; the default
pr(*a*,*b*)	$\Pr(a < y_j < b)$ under exogeneity and normal errors
e(*a*,*b*)	$E(y_j \mid a < y_j < b)$ under exogeneity and normal errors
<u>y</u>star(*a*,*b*)	$E(y_j^*)$, $y_j^* = \max\{a, \min(y_j, b)\}$ under exogeneity and normal errors
stdp	not allowed with margins
stdf	not allowed with margins
<u>res</u>iduals	not allowed with margins

Statistics not allowed with margins are functions of stochastic quantities other than e(b).

For the full syntax, see [R] **margins**.

estat

Description for estat

estat endogenous performs tests to determine whether endogenous regressors in the model are in fact exogenous. After GMM estimation, the C (difference-in-Sargan) statistic is reported. After 2SLS estimation with an unadjusted VCE, the Durbin (1954) and Wu–Hausman (Wu 1974; Hausman 1978) statistics are reported. After 2SLS estimation with a robust VCE, Wooldridge's (1995) robust score test and a robust regression-based test are reported. In all cases, if the test statistic is significant, then the variables being tested must be treated as endogenous. estat endogenous is not available after LIML estimation.

estat firststage reports various statistics that measure the relevance of the excluded exogenous variables. By default, whether the equation has one or more than one endogenous regressor determines what statistics are reported.

estat overid performs tests of overidentifying restrictions. If the 2SLS estimator was used, Sargan's (1958) and Basmann's (1960) χ^2 tests are reported, as is Wooldridge's (1995) robust score test; if the LIML estimator was used, Anderson and Rubin's (1950) χ^2 test and Basmann's F test are reported; and if the GMM estimator was used, Hansen's (1982) J statistic χ^2 test is reported. A statistically significant test statistic always indicates that the instruments may not be valid.

Menu for estat

Statistics > Postestimation

Syntax for estat

Perform tests of endogeneity

> estat <u>endog</u>enous [*varlist*] [, <u>l</u>ags(*#*) forceweights forcenonrobust]

Report "first-stage" regression statistics

> estat <u>first</u>stage [, all forcenonrobust]

Perform tests of overidentifying restrictions

> estat <u>over</u>id [, <u>l</u>ags(*#*) forceweights forcenonrobust]

collect is allowed with estat endogenous, estat firststage, and estat overid; see [U] **11.1.10 Prefix commands**.

Options for estat

Options for estat are presented under the following headings:

Options for estat endogenous

lags(#) specifies the number of lags to use for prewhitening when computing the heteroskedasticity-
and autocorrelation-consistent (HAC) version of the score test of endogeneity. Specifying lags(0)
requests no prewhitening. This option is valid only when the model was fit via 2SLS and an HAC
covariance matrix was requested when the model was fit. The default is lags(1).

forceweights requests that the tests of endogeneity be computed even though aweights, pweights,
or iweights were used in the previous estimation. By default, these tests are conducted only after
unweighted or frequency-weighted estimation. The reported critical values may be inappropriate
for weighted data, so the user must determine whether the critical values are appropriate for a
given application.

forcenonrobust requests that the Durbin and Wu–Hausman tests be performed after 2SLS estimation
even though a robust VCE was used at estimation time. This option is available only if the model
was fit by 2SLS.

Options for estat firststage

all requests that all first-stage goodness-of-fit statistics be reported regardless of whether the model
contains one or more endogenous regressors. By default, if the model contains one endogenous
regressor, then the first-stage R^2, adjusted R^2, partial R^2, and F statistics are reported, whereas
if the model contains multiple endogenous regressors, then Shea's partial R^2 and adjusted partial
R^2 are reported instead.

forcenonrobust requests that the minimum eigenvalue statistic and its critical values be reported
even though a robust VCE was used at estimation time. The reported critical values assume that
the errors are independent and identically distributed (i.i.d.) normal, so the user must determine
whether the critical values are appropriate for a given application.

Options for estat overid

lags(#) specifies the number of lags to use for prewhitening when computing the heteroskedasticity-
and autocorrelation-consistent (HAC) version of the score test of overidentifying restrictions.
Specifying lags(0) requests no prewhitening. This option is valid only when the model was fit
via 2SLS and an HAC covariance matrix was requested when the model was fit. The default is
lags(1).

forceweights requests that the tests of overidentifying restrictions be computed even though
aweights, pweights, or iweights were used in the previous estimation. By default, these tests
are conducted only after unweighted or frequency-weighted estimation. The reported critical values
may be inappropriate for weighted data, so the user must determine whether the critical values are
appropriate for a given application.

forcenonrobust requests that the Sargan and Basmann tests of overidentifying restrictions be
performed after 2SLS or LIML estimation even though a robust VCE was used at estimation time.
These tests assume that the errors are i.i.d. normal, so the user must determine whether the critical
values are appropriate for a given application.

Remarks and examples   stata.com

Remarks are presented under the following headings:

> estat endogenous
> estat firststage
> estat overid

estat endogenous

A natural question to ask is whether a variable presumed to be endogenous in the previously fit model could instead be treated as exogenous. If the endogenous regressors are in fact exogenous, then the OLS estimator is more efficient; and depending on the strength of the instruments and other factors, the sacrifice in efficiency by using an instrumental-variables estimator can be significant. Thus, unless an instrumental-variables estimator is really needed, OLS should be used instead. `estat endogenous` provides several tests of endogeneity after 2SLS and GMM estimation.

▷ Example 1

In example 1 of [R] **ivregress**, we fit a model of the average rental rate for housing in a state as a function of the percentage of the population living in urban areas and the average value of houses. We treated `hsngval` as endogenous because unanticipated shocks that affect rental rates probably affect house prices as well. We used family income and region dummies as additional instruments for `hsngval`. Here we test whether we could treat `hsngval` as exogenous.

```
. use https://www.stata-press.com/data/r18/hsng
(1980 Census housing data)
. ivregress 2sls rent pcturban (hsngval = faminc i.region)
 (output omitted )
. estat endogenous
 Tests of endogeneity
 H0: Variables are exogenous
 Durbin (score) chi2(1)         =  12.8473  (p = 0.0003)
 Wu-Hausman F(1,46)             =  15.9067  (p = 0.0002)
```

Because we did not specify any variable names after the `estat endogenous` command, Stata by default tested all the endogenous regressors (namely, `hsngval`) in our model. The null hypothesis of the Durbin and Wu–Hausman tests is that the variable under consideration can be treated as exogenous. Here both test statistics are highly significant, so we reject the null of exogeneity; we must continue to treat `hsngval` as endogenous.

◁

The difference between the Durbin and Wu–Hausman tests of endogeneity is that the former uses an estimate of the error term's variance based on the model assuming the variables being tested are exogenous, while the latter uses an estimate of the error variance based on the model assuming the variables being tested are endogenous. Under the null hypothesis that the variables being tested are exogenous, both estimates of the error variance are consistent. What we label the Wu–Hausman statistic is Wu's (1974) "T_2" statistic, which Hausman (1978) showed can be calculated very easily via linear regression. Baum, Schaffer, and Stillman (2003, 2007) provide a lucid discussion of these tests.

When you fit a model with multiple endogenous regressors, you can test the exogeneity of a subset of the regressors while continuing to treat the others as endogenous. For example, say you have three endogenous regressors, y1, y2, and y3, and you fit your model by typing

```
. ivregress depvar ... (y1 y2 y3 = ...)
```

Suppose you are confident that y1 must be treated as endogenous, but you are undecided about y2 and y3. To test whether y2 and y3 can be treated as exogenous, you would type

```
. estat endogenous y2 y3
```

The Durbin and Wu–Hausman tests assume that the error term is i.i.d. Therefore, if you requested a robust VCE at estimation time, `estat endogenous` will instead report Wooldridge's (1995) score test and a regression-based test of exogeneity. Both these tests can tolerate heteroskedastic and autocorrelated errors, while only the regression-based test is amenable to clustering.

▷ Example 2

We refit our housing model, requesting robust standard errors, and then test the exogeneity of hsngval:

```
. use https://www.stata-press.com/data/r18/hsng
(1980 Census housing data)

. ivregress 2sls rent pcturban (hsngval = faminc i.region), vce(robust)
  (output omitted)

. estat endogenous
  Tests of endogeneity
  H0: Variables are exogenous

  Robust score chi2(1)          =  2.10428  (p = 0.1469)
  Robust regression F(1,46)     =  4.31101  (p = 0.0435)
```

Wooldridge's score test does not reject the null hypothesis that hsngval is exogenous at conventional significance levels ($p = 0.1469$). However, the regression-based test does reject the null hypothesis at the 5% significance level ($p = 0.0435$). Typically, these two tests yield the same conclusion; the fact that our dataset has only 50 observations could be contributing to the discrepancy. Here we would be inclined to continue to treat hsngval as endogenous. Even if hsngval is exogenous, the 2SLS estimates are still consistent. On the other hand, if hsngval is in fact endogenous, the OLS estimates would not be consistent. Moreover, as we will see in our discussion of the `estat overid` command, our additional instruments may be invalid. To test whether an endogenous variable can be treated as exogenous, we must have a valid set of instruments to use to fit the model in the first place!

◁

Unlike the Durbin and Wu–Hausman tests, Wooldridge's score and the regression-based tests do not allow you to test a subset of the endogenous regressors in the model; you can test only whether all the endogenous regressors are in fact exogenous.

After GMM estimation, `estat endogenous` calculates what Hayashi (2000, 220) calls the C statistic, also known as the difference-in-Sargan statistic. The C statistic can be made robust to heteroskedasticity, autocorrelation, and clustering; and the version reported by `estat endogenous` is determined by the weight matrix requested via the `wmatrix()` option used when fitting the model with `ivregress`. Additionally, the test can be used to determine the exogeneity of a subset of the endogenous regressors, regardless of the type of weight matrix used.

If you fit your model using the LIML estimator, you can use the `hausman` command to carry out a traditional Hausman (1978) test between the OLS and LIML estimates.

estat firststage

For an excluded exogenous variable to be a valid instrument, it must be sufficiently correlated with the included endogenous regressors but uncorrelated with the error term. In recent decades, researchers have paid considerable attention to the issue of instruments that are only weakly correlated with the endogenous regressors. In such cases, the usual 2SLS, GMM, and LIML estimators are biased toward the OLS estimator, and inference based on the standard errors reported by, for example, ivregress can be severely misleading. For more information on the theory behind instrumental-variables estimation with weak instruments, see Nelson and Startz (1990); Staiger and Stock (1997); Hahn and Hausman (2003); the survey article by Stock, Wright, and Yogo (2002); and Angrist and Pischke (2009, chap. 4).

When the instruments are only weakly correlated with the endogenous regressors, some Monte Carlo evidence suggests that the LIML estimator performs better than the 2SLS and GMM estimators; see, for example, Poi (2006) and Stock, Wright, and Yogo (2002) (and the papers cited therein). On the other hand, the LIML estimator often results in confidence intervals that are somewhat larger than those from the 2SLS estimator.

Moreover, using more instruments is not a solution, because the biases of instrumental-variables estimators increase with the number of instruments. See Hahn and Hausman (2003).

estat firststage produces several statistics for judging the explanatory power of the instruments and is most easily explained with examples.

▷ Example 3

Again building on the model fit in example 1 of [R] **ivregress**, we now explore the degree of correlation between the additional instruments faminc, 2.region, 3.region, and 4.region and the endogenous regressor hsngval:

```
. use https://www.stata-press.com/data/r18/hsng
(1980 Census housing data)
. ivregress 2sls rent pcturban (hsngval = faminc i.region)
  (output omitted )
. estat firststage

  First-stage regression summary statistics
```

| | | Adjusted | Partial | | |
Variable	R-sq.	R-sq.	R-sq.	F(4,44)	Prob > F
hsngval	0.6908	0.6557	0.5473	13.2978	0.0000

```
Minimum eigenvalue statistic = 13.2978
Critical Values                   # of endogenous regressors:    1
H0: Instruments are weak          # of excluded instruments:     4
```

	5%	10%	20%	30%
2SLS relative bias	16.85	10.27	6.71	5.34

	10%	15%	20%	25%
2SLS size of nominal 5% Wald test	24.58	13.96	10.26	8.31
LIML size of nominal 5% Wald test	5.44	3.87	3.30	2.98

To understand these results, recall that the first-stage regression is

$$\text{hsngval}_i = \pi_0 + \pi_1\text{pcturban}_i + \pi_2\text{faminc} + \pi_3 2.\text{region} + \pi_4 3.\text{region} + \pi_5 4.\text{region} + v_i$$

where v_i is an error term. The column marked "R-sq." is the simple R^2 from fitting the first-stage regression by OLS, and the column marked "Adjusted R-sq." is the adjusted R^2 from that regression. Higher values purportedly indicate stronger instruments, and instrumental-variables estimators exhibit less bias when the instruments are strongly correlated with the endogenous variable.

Looking at just the R^2 and adjusted R^2 can be misleading, however. If `hsngval` were strongly correlated with the included exogenous variable `pcturban` but only weakly correlated with the additional instruments, then these statistics could be large even though a weak-instrument problem is present.

The partial R^2 statistic measures the correlation between `hsngval` and the additional instruments after *partialing out* the effect of `pcturban`. Unlike the R^2 and adjusted R^2 statistics, the partial R^2 statistic will not be inflated because of strong correlation between `hsngval` and `pcturban`. Bound, Jaeger, and Baker (1995) and others have promoted using this statistic.

The column marked "F(4, 44)" is an F statistic for the joint significance of π_2, π_3, π_4, and π_5, the coefficients on the additional instruments. Its p-value is listed in the column marked "Prob > F". If the F statistic is not significant, then the additional instruments have no significant explanatory power for `hsngval` after controlling for the effect of `pcturban`. However, Hall, Rudebusch, and Wilcox (1996) used Monte Carlo simulation to show that simply having an F statistic that is significant at the typical 5% or 10% level is not sufficient. Stock, Wright, and Yogo (2002) suggest that the F statistic should exceed 10 for inference based on the 2SLS estimator to be reliable when there is one endogenous regressor.

`estat firststage` also presents the Cragg and Donald (1993) minimum eigenvalue statistic as a further test of weak instruments. Stock and Yogo (2005) discuss two characterizations of weak instruments: first, weak instruments cause instrumental-variables estimators to be biased; second, hypothesis tests of parameters estimated by instrumental-variables estimators may suffer from severe size distortions. The test statistic in our example is 13.30, which is identical to the F statistic just discussed because our model contains one endogenous regressor.

The null hypothesis of each of Stock and Yogo's tests is that the set of instruments is weak. To perform these tests, we must first choose either the largest relative bias of the 2SLS estimator we are willing to tolerate or the largest rejection rate of a nominal 5% Wald test we are willing to tolerate. If the test statistic exceeds the critical value, we can conclude that our instruments are not weak.

The row marked "2SLS relative bias" contains critical values for the test that the instruments are weak based on the bias of the 2SLS estimator *relative to* the bias of the OLS estimator. For example, from past experience we might know that the OLS estimate of a parameter β may be 50% too high. Saying that we are willing to tolerate a 10% relative bias means that we are willing to tolerate a bias of the 2SLS estimator no greater than 5% (that is, 10% of 50%). In our rental rate model, if we are willing to tolerate a 10% relative bias, then we can conclude that our instruments are not weak because the test statistic of 13.30 exceeds the critical value of 10.27. However, if we were willing to tolerate only a relative bias of 5%, we would conclude that our instruments are weak because $13.30 < 16.85$.

The rows marked "2SLS Size of nominal 5% Wald test" and "LIML Size of nominal 5% Wald test" contain critical values pertaining to Stock and Yogo's (2005) second characterization of weak instruments. This characterization defines a set of instruments to be weak if a Wald test at the 5% level can have an actual rejection rate of no more than 10%, 15%, 20%, or 25%. Using the current example, suppose that we are willing to accept a rejection rate of at most 10%. Because $13.30 < 24.58$, we cannot reject the null hypothesis of weak instruments. On the other hand, if we use the LIML estimator instead, then we can reject the null hypothesis because $13.30 > 5.44$.

◁

❑ Technical note

Stock and Yogo (2005) tabulated critical values for 2SLS relative biases of 5%, 10%, 20%, and 30% for models with 1, 2, or 3 endogenous regressors and between 3 and 30 excluded exogenous variables (instruments). They also provide critical values for worst-case rejection rates of 5%, 10%, 20%, and 25% for nominal 5% Wald tests of the endogenous regressors with 1 or 2 endogenous regressors and between 1 and 30 instruments. If the model previously fit by ivregress has more instruments or endogenous regressors than these limits, the critical values are not shown. Stock and Yogo did not consider GMM estimators.

❑

When the model being fit contains more than one endogenous regressor, the R^2 and F statistics described above can overstate the relevance of the excluded instruments. Suppose that there are two endogenous regressors, Y_1 and Y_2, and that there are two additional instruments, z_1 and z_2. Say that z_1 is highly correlated with both Y_1 and Y_2 but z_2 is not correlated with either Y_1 or Y_2. Then, the first-stage regression of Y_1 on z_1 and z_2 (along with the included exogenous variables) will produce large R^2 and F statistics, as will the regression of Y_2 on z_1, z_2, and the included exogenous variables. Nevertheless, the lack of correlation between z_2 and Y_1 and Y_2 is problematic. Here, although the order condition indicates that the model is just identified (the number of excluded instruments equals the number of endogenous regressors), the irrelevance of z_2 implies that the model is in fact not identified. Even if the model is overidentified, including irrelevant instruments can adversely affect the properties of instrumental-variables estimators, because their biases increase as the number of instruments increases.

▷ Example 4

estat firststage presents different statistics when the model contains multiple endogenous regressors. For illustration, we refit our model of rental rates, assuming that both hsngval and faminc are endogenously determined. We use i.region along with popden, a measure of population density, as additional instruments.

```
. ivregress 2sls rent pcturban (hsngval faminc = i.region popden)
(output omitted)

. estat firststage
```

Shea's partial R-squared

Variable	Shea's partial R-sq.	Shea's adj. partial R-sq.
hsngval	0.3477	0.2735
faminc	0.1893	0.0972

Minimum eigenvalue statistic = 2.51666

| Critical Values
H0: Instruments are weak | # of endogenous regressors:
# of excluded instruments: | 2
4 |

	5%	10%	20%	30%
2SLS relative bias	11.04	7.56	5.57	4.73

	10%	15%	20%	25%
2SLS size of nominal 5% Wald test	16.87	9.93	7.54	6.28
LIML size of nominal 5% Wald test	4.72	3.39	2.99	2.79

Consider the endogenous regressor hsngval. Part of its variation is attributable to its correlation with the other regressors pcturban and faminc. The other component of hsngval's variation is peculiar to it and orthogonal to the variation in the other regressors. Similarly, we can think of the instruments as predicting the variation in hsngval in two ways, one stemming from the fact that the predicted values of hsngval are correlated with the predicted values of the other regressors and one from the variation in the predicted values of hsngval that is orthogonal to the variation in the predicted values of the other regressors.

What really matters for instrumental-variables estimation is whether the component of hsngval that is orthogonal to the other regressors can be explained by the component of the predicted value of hsngval that is orthogonal to the predicted values of the other regressors in the model. Shea's (1997) partial R^2 statistic measures this correlation. Because the bias of instrumental-variables estimators increases as more instruments are used, Shea's adjusted partial R^2 statistic is often used instead, as it makes a degrees-of-freedom adjustment for the number of instruments, analogous to the adjusted R^2 measure used in OLS regression. Although what constitutes a "low" value for Shea's partial R^2 depends on the specifics of the model being fit and the data used, these results, taken in isolation, do not strike us as being a particular cause for concern.

However, with this specification the minimum eigenvalue statistic is low. We cannot reject the null hypothesis of weak instruments for either of the characterizations we have discussed.

◁

By default, estat firststage determines which statistics to present based on the number of endogenous regressors in the model previously fit. However, you can specify the all option to obtain all the statistics.

❑ Technical note

If the previous estimation was conducted using aweights, pweights, or iweights, then the first-stage regression summary statistics are computed using those weights. However, in these cases the minimum eigenvalue statistic and its critical values are not available.

If the previous estimation included a robust VCE, then the first-stage F statistic is based on a robust VCE as well; for example, if you fit your model with an HAC VCE using the Bartlett kernel and four lags, then the F statistic reported is based on regression results using an HAC VCE using the Bartlett kernel and four lags. By default, the minimum eigenvalue statistic and its critical values are not displayed. You can use the forcenonrobust option to obtain them in these cases; the minimum eigenvalue statistic is computed using the weights, though the critical values reported may not be appropriate.

❑

estat overid

In addition to the requirement that instrumental variables be correlated with the endogenous regressors, the instruments must also be uncorrelated with the structural error term. If the model is overidentified, meaning that the number of additional instruments exceeds the number of endogenous regressors, then we can test whether the instruments are uncorrelated with the error term. If the model is just identified, then we cannot perform a test of overidentifying restrictions.

The estimator you used to fit the model determines which tests of overidentifying restrictions estat overid reports. If you used the 2SLS estimator without a robust VCE, estat overid reports Sargan's (1958) and Basmann's (1960) χ^2 tests. If you used the 2SLS estimator and requested a robust VCE, Wooldridge's robust score test of overidentifying restrictions is performed instead; without a robust VCE, Wooldridge's test statistic is identical to Sargan's test statistic. If you used the LIML estimator, estat overid reports the Anderson–Rubin (1950) likelihood-ratio test and Basmann's (1960) F test. estat overid reports Hansen's (1982) J statistic if you used the GMM estimator. Davidson and MacKinnon (1993, 235–236) give a particularly clear explanation of the intuition behind tests of overidentifying restrictions. Also see Judge et al. (1985, 614–616) for a summary of tests of overidentifying restrictions for the 2SLS and LIML estimators.

Tests of overidentifying restrictions actually test two different things simultaneously. One, as we have discussed, is whether the instruments are uncorrelated with the error term. The other is that the equation is misspecified and that one or more of the excluded exogenous variables should in fact be included in the structural equation. Thus, a significant test statistic could represent either an invalid instrument or an incorrectly specified structural equation.

▷ Example 5

Here we refit the model that treated just hsngval as endogenous using 2SLS, and then we perform tests of overidentifying restrictions:

```
. ivregress 2sls rent pcturban (hsngval = faminc i.region)
  (output omitted)
. estat overid

  Tests of overidentifying restrictions:
  Sargan (score) chi2(3) =  11.2877  (p = 0.0103)
  Basmann chi2(3)        =  12.8294  (p = 0.0050)
```

Both test statistics are significant at the 5% test level, which means that either one or more of our instruments are invalid or that our structural model is specified incorrectly.

One possibility is that the error term in our structural model is heteroskedastic. Both Sargan's and Basmann's tests assume that the errors are i.i.d.; if the errors are not i.i.d., then these tests are not valid. Here we refit the model by requesting heteroskedasticity-robust standard errors, and then we use estat overid to obtain Wooldridge's score test of overidentifying restrictions, which is robust to heteroskedasticity.

```
. ivregress 2sls rent pcturban (hsngval = faminc i.region), vce(robust)
(output omitted )

. estat overid

 Test of overidentifying restrictions:

 Score chi2(3)            =   6.8364  (p = 0.0773)
```

Here we no longer reject the null hypothesis that our instruments are valid at the 5% significance level, though we do reject the null at the 10% level. You can verify that the robust standard error on the coefficient for hsngval is more than twice as large as its nonrobust counterpart and that the robust standard error for pcturban is nearly 50% larger.

◁

❑ Technical note

The test statistic for the test of overidentifying restrictions performed after GMM estimation is simply the sample size times the value of the objective function $Q(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ defined in (5) of [R] **ivregress**, evaluated at the GMM parameter estimates. If the weighting matrix $\mathbf{W}$ is optimal, meaning that $\mathbf{W} = \text{Var}\left(\mathbf{z}_i u_i\right)$, then $Q(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) \overset{A}{\sim} \chi^2(q)$, where q is the number of overidentifying restrictions. However, if the estimated $\mathbf{W}$ is not optimal, then the test statistic will not have an asymptotic χ^2 distribution.

Like the Sargan and Basmann tests of overidentifying restrictions for the 2SLS estimator, the Anderson–Rubin and Basmann tests after LIML estimation are predicated on the errors' being i.i.d. If the previous LIML results were reported with robust standard errors, then estat overid by default issues an error message and refuses to report the Anderson–Rubin and Basmann test statistics. You can use the forcenonrobust option to override this behavior. You can also use forcenonrobust to obtain the Sargan and Basmann test statistics after 2SLS estimation with robust standard errors.

❑

By default, estat overid issues an error message if the previous estimation was conducted using aweights, pweights, or iweights. You can use the forceweights option to override this behavior, though the test statistics may no longer have the expected χ^2 distributions.

Stored results

After 2SLS estimation, estat endogenous stores the following in r():

Scalars

r(durbin)	Durbin χ^2 statistic
r(p_durbin)	p-value for Durbin χ^2 statistic
r(wu)	Wu–Hausman F statistic
r(p_wu)	p-value for Wu–Hausman F statistic
r(df)	degrees of freedom
r(wudf_r)	denominator degrees of freedom for Wu–Hausman F
r(r_score)	robust score statistic
r(p_r_score)	p-value for robust score statistic
r(hac_score)	HAC score statistic
r(p_hac_score)	p-value for HAC score statistic
r(lags)	lags used in prewhitening
r(regF)	regression-based F statistic
r(p_regF)	p-value for regression-based F statistic
r(regFdf_n)	regression-based F numerator degrees of freedom
r(regFdf_r)	regression-based F denominator degrees of freedom

After GMM estimation, `estat endogenous` stores the following in `r()`:

Scalars
`r(C)`	C χ^2 statistic
`r(p_C)`	p-value for C χ^2 statistic
`r(df)`	degrees of freedom

`estat firststage` stores the following in `r()`:

Scalars
`r(mineig)`	minimum eigenvalue statistic

Matrices
`r(mineigcv)`	critical values for minimum eigenvalue statistic
`r(multiresults)`	Shea's partial R^2 statistics
`r(singleresults)`	first-stage R^2 and F statistics

After 2SLS estimation, `estat overid` stores the following in `r()`:

Scalars
`r(lags)`	lags used in prewhitening
`r(df)`	χ^2 degrees of freedom
`r(score)`	score χ^2 statistic
`r(p_score)`	p-value for score χ^2 statistic
`r(basmann)`	Basmann χ^2 statistic
`r(p_basmann)`	p-value for Basmann χ^2 statistic
`r(sargan)`	Sargan χ^2 statistic
`r(p_sargan)`	p-value for Sargan χ^2 statistic

After LIML estimation, `estat overid` stores the following in `r()`:

Scalars
`r(ar)`	Anderson–Rubin χ^2 statistic
`r(p_ar)`	p-value for Anderson–Rubin χ^2 statistic
`r(ar_df)`	χ^2 degrees of freedom
`r(basmann)`	Basmann F statistic
`r(p_basmann)`	p-value for Basmann F statistic
`r(basmann_df_n)`	F numerator degrees of freedom
`r(basmann_df_d)`	F denominator degrees of freedom

After GMM estimation, `estat overid` stores the following in `r()`:

Scalars
`r(HansenJ)`	Hansen's J χ^2 statistic
`r(p_HansenJ)`	p-value for Hansen's J χ^2 statistic
`r(J_df)`	χ^2 degrees of freedom

Methods and formulas

Methods and formulas are presented under the following headings:

> *Notation*
> *estat endogenous*
> *estat firststage*
> *estat overid*

Notation

Recall from [R] **ivregress** that the model is

$$\mathbf{y} = \mathbf{Y}\boldsymbol{\beta}_1 + \mathbf{X}_1\boldsymbol{\beta}_2 + \mathbf{u} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\Pi}_1 + \mathbf{X}_2\boldsymbol{\Pi}_2 + \mathbf{V} = \mathbf{Z}\boldsymbol{\Pi} + \mathbf{V}$$

where $\mathbf{y}$ is an $N \times 1$ vector of the left-hand-side variable, N is the sample size, $\mathbf{Y}$ is an $N \times p$ matrix of p endogenous regressors, $\mathbf{X}_1$ is an $N \times k_1$ matrix of k_1 included exogenous regressors, $\mathbf{X}_2$ is an $N \times k_2$ matrix of k_2 excluded exogenous variables, $\mathbf{X} = [\mathbf{Y}\ \mathbf{X}_1]$, $\mathbf{Z} = [\mathbf{X}_1\ \mathbf{X}_2]$, $\mathbf{u}$ is an $N \times 1$ vector of errors, $\mathbf{V}$ is an $N \times p$ matrix of errors, $\boldsymbol{\beta} = [\boldsymbol{\beta}_1\ \boldsymbol{\beta}_2]$ is a $k = (p + k_1) \times 1$ vector of parameters, and $\boldsymbol{\Pi}$ is a $(k_1 + k_2) \times p$ vector of parameters. If a constant term is included in the model, then one column of $\mathbf{X}_1$ contains all ones.

estat endogenous

Partition $\mathbf{Y}$ as $\mathbf{Y} = [\mathbf{Y}_1\ \mathbf{Y}_2]$, where $\mathbf{Y}_1$ represents the p_1 endogenous regressors whose endogeneity is being tested and $\mathbf{Y}_2$ represents the p_2 endogenous regressors whose endogeneity is not being tested. If the endogeneity of all endogenous regressors is being tested, $\mathbf{Y} = \mathbf{Y_1}$ and $p_2 = 0$. After GMM estimation, estat endogenous refits the model treating $\mathbf{Y}_1$ as exogenous using the same type of weight matrix as requested at estimation time with the wmatrix() option; denote the Sargan statistic from this model by J_e and the estimated weight matrix by $\mathbf{W}_e$. Let $\mathbf{S}_e = \mathbf{W}_e^{-1}$. estat endogenous removes from $\mathbf{S}_e$ the rows and columns corresponding to the variables represented by $\mathbf{Y}_1$; denote the inverse of the resulting matrix by $\mathbf{W}'_e$. Next, estat endogenous fits the model treating both $\mathbf{Y}_1$ and $\mathbf{Y}_2$ as endogenous, using the weight matrix $\mathbf{W}'_e$; denote the Sargan statistic from this model by J_c. Then, $C = (J_e - J_c) \sim \chi^2(p_1)$. If one simply used the J statistic from the original model fit by ivregress in place of J_c, then in finite samples $J_e - J$ might be negative. The procedure used by estat endogenous is guaranteed to yield $C \geq 0$; see Hayashi (2000, 220).

Let $\widehat{\mathbf{u}}_c$ denote the residuals from the model treating both $\mathbf{Y}_1$ and $\mathbf{Y}_2$ as endogenous, and let $\widehat{\mathbf{u}}_e$ denote the residuals from the model treating only $\mathbf{Y}_2$ as endogenous. Then, Durbin's (1954) statistic is

$$D = \frac{\widehat{\mathbf{u}}'_e \mathbf{P}_{ZY_1} \widehat{\mathbf{u}}_e - \widehat{\mathbf{u}}'_c \mathbf{P}_Z \widehat{\mathbf{u}}_c}{\widehat{\mathbf{u}}'_e \widehat{\mathbf{u}}_e / N}$$

where $\mathbf{P}_Z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$ and $\mathbf{P}_{ZY_1} = [\mathbf{Z}\ \mathbf{Y}_1]([\mathbf{Z}\ \mathbf{Y}_1]'[\mathbf{Z}\ \mathbf{Y}_1])^{-1}[\mathbf{Z}\ \mathbf{Y}_1]'$ $D \sim \chi^2(p_1)$. The Wu–Hausman (Wu 1974; Hausman 1978) statistic is

$$WH = \frac{(\widehat{\mathbf{u}}'_e \mathbf{P}_{ZY_1} \widehat{\mathbf{u}}_e - \widehat{\mathbf{u}}'_c \mathbf{P}_Z \widehat{\mathbf{u}}_c)/p_1}{\{\widehat{\mathbf{u}}'_e \widehat{\mathbf{u}}_e - (\widehat{\mathbf{u}}'_e \mathbf{P}_{ZY_1} \widehat{\mathbf{u}}_e - \widehat{\mathbf{u}}'_c \mathbf{P}_Z \widehat{\mathbf{u}}_c)\}/(N - k_1 - p - p_1)}$$

$WH \sim F(p_1, N - k_1 - p - p_1)$. Baum, Schaffer, and Stillman (2003, 2007) discuss these tests in more detail.

Next, we describe Wooldridge's (1995) score test. The nonrobust version of Wooldridge's test is identical to Durbin's test. Suppose a robust covariance matrix was used at estimation time. Let $\widehat{\mathbf{e}}$ denote the sample residuals obtained by fitting the model via OLS, treating $\mathbf{Y}$ as exogenous. We then regress each variable represented in $\mathbf{Y}$ on $\mathbf{Z}$; call the residuals for the jth regression $\widehat{\mathbf{r}}_j$, $j = 1, \ldots, p$. Define $\widehat{k}_{ij} = \widehat{e}_i \widehat{r}_{ij}$, $i = 1, \ldots, N$. We then run the regression

$$\mathbf{1} = \theta_1 \widehat{\mathbf{k}}_1 + \cdots + \theta_p \widehat{\mathbf{k}}_p + \boldsymbol{\epsilon}$$

where $\mathbf{1}$ is an $N \times 1$ vector of ones and ϵ is a regression error term. $N - \text{RSS} \sim \chi^2(p)$, where RSS is the residual sum of squares from the regression just described. If instead an HAC VCE was used at estimation time, then before running the final regression we prewhiten the $\widehat{\mathbf{k}}_j$ series by using a VAR(q) model, where q is the number of lags specified with the `lags()` option.

The regression-based test proceeds as follows. Following Hausman (1978, 1259), we regress $\mathbf{Y}$ on $\mathbf{Z}$ and obtain the residuals $\widehat{\mathbf{V}}$. Next, we fit the augmented regression

$$\mathbf{y} = \mathbf{Y}\boldsymbol{\beta}_1 + \mathbf{X}_1\boldsymbol{\beta}_2 + \widehat{\mathbf{V}}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$$

by OLS regression, where ϵ is a regression error term. A test of the exogeneity of $\mathbf{Y}$ is equivalent to a test of $\boldsymbol{\gamma} = \mathbf{0}$. As Cameron and Trivedi (2005, 276) suggest, this test can be made robust to heteroskedasticity, autocorrelation, or clustering by using the appropriate robust VCE when testing $\boldsymbol{\gamma} = \mathbf{0}$. When a nonrobust VCE is used, this test is equivalent to the Wu–Hausman test described earlier. One cannot simply fit this augmented regression via 2SLS to test the endogeneity of a subset of the endogenous regressors; Davidson and MacKinnon (1993, 229–231) discuss a test of $\boldsymbol{\gamma} = \mathbf{0}$ for the homoskedastic version of the augmented regression fit by 2SLS, but an appropriate robust test is not apparent.

estat firststage

When the structural equation includes one endogenous regressor, `estat firststage` fits the regression

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\pi}_1 + \mathbf{X}_2\boldsymbol{\pi}_2 + \mathbf{v}$$

via OLS. The R^2 and adjusted R^2 from that regression are reported in the output, as well as the F statistic from the Wald test of $H_0: \boldsymbol{\pi}_2 = \mathbf{0}$. To obtain the partial R^2 statistic, `estat firststage` fits the regression

$$\mathbf{M}_{\mathbf{X}_1}\mathbf{y} = \mathbf{M}_{\mathbf{X}_1}\mathbf{X}_2\boldsymbol{\xi} + \boldsymbol{\epsilon}$$

by OLS, where ϵ is a regression error term, $\boldsymbol{\xi}$ is a $k_2 \times 1$ parameter vector, and $\mathbf{M}_{\mathbf{X}_1} = \mathbf{I} - \mathbf{X}_1(\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'$; that is, the partial R^2 is the R^2 between $\mathbf{y}$ and $\mathbf{X}_2$ after eliminating the effects of $\mathbf{X}_1$. If the model contains multiple endogenous regressors and the `all` option is specified, these statistics are calculated for each endogenous regressor in turn.

To calculate Shea's partial R^2, let $\mathbf{y}_1$ denote the endogenous regressor whose statistic is being calculated and $\mathbf{Y}_0$ denote the other endogenous regressors. Define $\widetilde{\mathbf{y}}_1$ as the residuals obtained from regressing $\mathbf{y}_1$ on $\mathbf{Y}_0$ and $\mathbf{X}_1$. Let $\widehat{\mathbf{y}}_1$ denote the fitted values obtained from regressing $\mathbf{y}_1$ on $\mathbf{X}_1$ and $\mathbf{X}_2$; that is, $\widehat{\mathbf{y}}_1$ are the fitted values from the first-stage regression for $\mathbf{y}_1$, and define the columns of $\widehat{\mathbf{Y}}_0$ analogously. Finally, let $\widetilde{\widehat{\mathbf{y}}}_1$ denote the residuals from regressing $\widehat{\mathbf{y}}_1$ on $\widehat{\mathbf{Y}}_0$ and $\mathbf{X}_1$. Shea's partial R^2 is the simple R^2 from the regression of $\widetilde{\mathbf{y}}_1$ on $\widetilde{\widehat{\mathbf{y}}}_1$; denote this as R_S^2. Shea's adjusted partial R^2 is equal to $1 - (1 - R_S^2)(N-1)/(N - k_Z + 1)$ if a constant term is included and $1 - (1 - R_S^2)(N-1)/(N - k_Z)$ if there is no constant term included in the model, where $k_Z = k_1 + k_2$. For one endogenous regressor, one instrument, no exogenous regressors, and a constant term, R_S^2 equals the adjusted R_S^2.

The Stock and Yogo minimum eigenvalue statistic, first proposed by Cragg and Donald (1993) as a test for underidentification, is the minimum eigenvalue of the matrix

$$\mathbf{G} = \frac{1}{k_Z}\widehat{\boldsymbol{\Sigma}}_{\mathbf{VV}}^{-1/2}\mathbf{Y}'\mathbf{M}_{\mathbf{X}_1}'\mathbf{X}_2(\mathbf{X}_2'\mathbf{M}_{\mathbf{X}_1}\mathbf{X}_2)^{-1}\mathbf{X}_2'\mathbf{M}_{\mathbf{X}_1}\mathbf{Y}\widehat{\boldsymbol{\Sigma}}_{\mathbf{VV}}^{-1/2}$$

where

$$\widehat{\boldsymbol{\Sigma}}_{\mathbf{VV}} = \frac{1}{N - k_Z} \mathbf{Y}' \mathbf{M}_{\mathbf{Z}} \mathbf{Y}$$

$\mathbf{M_Z} = \mathbf{I} - \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$, and $\mathbf{Z} = [\mathbf{X}_1 \ \mathbf{X}_2]$. Critical values are obtained from the tables in Stock and Yogo (2005).

estat overid

The Sargan (1958) and Basmann (1960) χ^2 statistics are calculated by running the auxiliary regression

$$\widehat{\mathbf{u}} = \mathbf{Z}\boldsymbol{\delta} + \mathbf{e}$$

where $\widehat{\mathbf{u}}$ are the sample residuals from the model and $\mathbf{e}$ is an error term. Then, Sargan's statistic is

$$S = N \left(1 - \frac{\widehat{\mathbf{e}}'\widehat{\mathbf{e}}}{\widehat{\mathbf{u}}'\widehat{\mathbf{u}}} \right)$$

where $\widehat{\mathbf{e}}$ are the residuals from that auxiliary regression. Basmann's statistic is calculated as

$$B = S\frac{N - k_Z}{N - S}$$

Both S and B are distributed $\chi^2(m)$, where m, the number of overidentifying restrictions, is equal to $k_Z - k$, where k is the number of endogenous regressors.

Wooldridge's (1995) score test of overidentifying restrictions is identical to Sargan's (1958) statistic under the assumption of i.i.d. and therefore is not recomputed unless a robust VCE was used at estimation time. If a heteroskedasticity-robust VCE was used, Wooldridge's test proceeds as follows. Let $\widehat{\mathbf{Y}}$ denote the $N \times k$ matrix of fitted values obtained by regressing the endogenous regressors on $\mathbf{X}_1$ and $\mathbf{X}_2$. Let $\mathbf{Q}$ denote an $N \times m$ matrix of excluded exogenous variables; the test statistic to be calculated is invariant to whichever m of the k_2 excluded exogenous variables is chosen. Define the ith element of $\widehat{\mathbf{k}}_j$, $i = 1, \dots, N$, $j = 1, \dots, m$, as

$$k_{ij} = \widehat{q}_{ij}\widehat{u}_i$$

where $\widehat{q}_{ij}$ is the ith element of $\widehat{\mathbf{q}}_j$, the residuals from regressing the jth column of $\mathbf{Q}$ on $\widehat{\mathbf{Y}}$ and $\mathbf{X}_1$. Finally, fit the regression

$$\mathbf{1} = \theta_1\widehat{\mathbf{k}}_1 + \dots + \theta_m\widehat{\mathbf{k}}_m + \boldsymbol{\epsilon}$$

where $\mathbf{1}$ is an $N \times 1$ vector of ones and $\boldsymbol{\epsilon}$ is a regression error term, and calculate the residual sum of squares, RSS. Then, the test statistic is $W = N - \text{RSS}$. $W \sim \chi^2(m)$. If an HAC VCE was used at estimation, then the $\widehat{\mathbf{k}}_j$ are prewhitened using a VAR(p) model, where p is specified using the `lags()` option.

The Anderson–Rubin (1950), AR, test of overidentifying restrictions for use after the LIML estimator is calculated as AR $= N(\kappa - 1)$, where κ is the minimal eigenvalue of a certain matrix defined in *Methods and formulas* of [R] **ivregress**. AR $\sim \chi^2(m)$. (Some texts define this statistic as $N\ln(\kappa)$ because $\ln(x) \approx (x - 1)$ for x near 1.) Basmann's F statistic for use after the LIML estimator is calculated as $B_F = (\kappa - 1)(N - k_Z)/m$. $B_F \sim F(m, N - k_Z)$.

Hansen's J statistic is simply the sample size times the value of the GMM objective function defined in (5) of [R] **ivregress**, evaluated at the estimated parameter values. Under the null hypothesis that the overidentifying restrictions are valid, $J \sim \chi^2(m)$.

John Denis Sargan (1924–1996) was born in Yorkshire, UK. He pioneered the theory of instrumental-variables (IV) estimation in an article published in 1958. In the article, he also developed overidentification tests, developed significance tests, and discussed possible instruments for applied work. A year later, he wrote an article extending the theory to models containing autoregressive errors. This extension was one of his many contributions to time-series econometric analysis. For example, in 1964 he published a paper in which he developed misspecification tests for dynamic equations, along with an IV estimator for models with nonlinear parameters, and a model with a long-run equilibrium. His paper laid the foundation for other econometric methods, such as cointegration analysis, and established what would be known as the London School of Economics (LSE) approach to econometric modeling. He spent twenty years at this institution, supervising the doctoral work of many econometricians who themselves made important contributions to econometrics. In addition to Sargan's many lasting contributions to econometrics, he also left a lasting impression on his students and colleagues through his generosity.

References

Anderson, T. W., and H. Rubin. 1950. The asymptotic properties of estimates of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics* 21: 570–582. https://doi.org/10.1214/aoms/1177729752.

Angrist, J. D., and J.-S. Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ: Princeton University Press.

Basmann, R. L. 1960. On finite sample distributions of generalized classical linear identifiability test statistics. *Journal of the American Statistical Association* 55: 650–659. https://doi.org/10.2307/2281588.

Baum, C. F., M. E. Schaffer, and S. Stillman. 2003. Instrumental variables and GMM: Estimation and testing. *Stata Journal* 3: 1–31.

———. 2007. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *Stata Journal* 7: 465–506.

Bound, J., D. A. Jaeger, and R. M. Baker. 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association* 90: 443–450. https://doi.org/10.2307/2291055.

Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications.* New York: Cambridge University Press.

Cragg, J. G., and S. G. Donald. 1993. Testing identifiability and specification in instrumental variable models. *Econometric Theory* 9: 222–240. https://doi.org/10.1017/S0266466600007519.

Davidson, R., and J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics.* New York: Oxford University Press.

Durbin, J. 1954. Errors in variables. *Review of the International Statistical Institute* 22: 23–32. https://doi.org/10.2307/1401917.

Hahn, J., and J. A. Hausman. 2003. Weak instruments: Diagnosis and cures in empirical econometrics. *American Economic Review Papers and Proceedings* 93: 118–125.

Hall, A. R., G. D. Rudebusch, and D. W. Wilcox. 1996. Judging instrument relevance in instrumental variables estimation. *International Economic Review* 37: 283–298. https://doi.org/10.2307/2527324.

Hansen, L. P. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50: 1029–1054. https://doi.org/10.2307/1912775.

Hausman, J. A. 1978. Specification tests in econometrics. *Econometrica* 46: 1251–1271. https://doi.org/10.2307/1913827.

Hayashi, F. 2000. *Econometrics.* Princeton, NJ: Princeton University Press.

Judge, G. G., W. E. Griffiths, R. C. Hill, H. Lütkepohl, and T.-C. Lee. 1985. *The Theory and Practice of Econometrics.* 2nd ed. New York: Wiley.

Nelson, C. R., and R. Startz. 1990. The distribution of the instrumental variable estimator and its t ratio when the instrument is a poor one. *Journal of Business* 63: S125–S140.

Pflueger, C. E., and S. Wang. 2015. A robust test for weak instruments in Stata. *Stata Journal* 15: 216–225.

Poi, B. P. 2006. Jackknife instrumental variables estimation in Stata. *Stata Journal* 6: 364–376.

Sargan, J. D. 1958. The estimation of economic relationships using instrumental variables. *Econometrica* 26: 393–415. https://doi.org/10.2307/1907619.

Shea, J. S. 1997. Instrument relevance in multivariate linear models: A simple measure. *Review of Economics and Statistics* 79: 348–352. https://doi.org/10.1162/rest.1997.79.2.348.

Staiger, D. O., and J. H. Stock. 1997. Instrumental variables regression with weak instruments. *Econometrica* 65: 557–586. https://doi.org/10.2307/2171753.

Stock, J. H., J. H. Wright, and M. Yogo. 2002. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics* 20: 518–529. https://doi.org/10.1198/073500102288618658.

Stock, J. H., and M. Yogo. 2005. Testing for weak instruments in linear IV regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. D. W. K. Andrews and J. H. Stock, 80–108. New York: Cambridge University Press.

Sun, L. 2018. Implementing valid two-step identification-robust confidence sets for linear instrumental-variables models. *Stata Journal* 18: 803–825.

Wooldridge, J. M. 1995. Score diagnostics for linear models estimated by two stage least squares. In *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C. R. Rao*, ed. G. S. Maddala, P. C. B. Phillips, and T. N. Srinivasan, 66–87. Oxford: Blackwell.

Wu, D.-M. 1974. Alternative tests of independence between stochastic regressors and disturbances: Finite sample results. *Econometrica* 42: 529–546. https://doi.org/10.2307/1911789.

Also see

[R] **ivregress** — Single-equation instrumental-variables regression

[U] **20 Estimation and postestimation commands**