The Stata Journal (2003)
**3**, Number 1, pp. 57–70

# Implementing tests with correct size in the simultaneous equations model

Marcelo J. Moreira
Department of Economics
Harvard University

Brian P. Poi
Stata Corporation

**Abstract.** In this paper, we propose a fix to the size distortions of tests for structural parameters in the simultaneous equations model by computing critical value functions based on the conditional distribution of test statistics. The conditional tests can then be used to construct informative confidence regions for the structural parameter with correct coverage probability. Commands to implement these tests in Stata are also introduced.

**Keywords:** instrumental variables, weak instruments, similar tests, score test, Wald test, likelihood-ratio test, confidence regions, 2SLS estimator, LIML estimator

## 1 Introduction

When making inferences about coefficients of endogenous variables in a structural equation, applied researchers often rely on asymptotic approximations. However, as emphasized in recent work by Bound et al. (1995) and Staiger and Stock (1997), these approximations are not satisfactory when instruments are weakly correlated with the regressors. In particular, if identification can be arbitrarily weak, Dufour (1997) showed that Wald-type confidence intervals have a zero confidence level. The problem arises because inference is based on nonpivotal statistics whose exact distributions depart substantially from their asymptotic approximations when identification is weak.

Based on the methods developed by Moreira (2001a,b) and explained thoroughly by Moreira (2002), we construct valid tests of structural coefficients based on the conditional distribution of nonpivotal statistics. The conditional approach is then employed to find critical value functions for Wald and likelihood-ratio tests yielding correct rejection probabilities, no matter how weak the instruments.

Together with the Anderson–Rubin (1949) and score tests, the conditional Wald and likelihood-ratio tests can be used to construct confidence intervals that have correct coverage probability even when instruments may be weak and that are informative when instruments are good. The regions based on the conditional Wald test necessarily contain the 2SLS estimator, while the ones based on the conditional likelihood-ratio and score tests are centered around the limited-information maximum likelihood (LIML) estimator. Therefore, confidence regions based on these tests can be used as reliable evidence of the accuracy of commonly used estimators.

In Section 2, exact results are developed for the two-equation model under the assumption that the reduced-form disturbances are normally distributed with a known covariance matrix. Sections 3 extends the results to more realistic cases and explains how to construct confidence regions based on the conditional test. Section 4 considers the syntax of the Stata commands and demonstrates their usage.

# 2 The conditional approach

## 2.1 The model

To simplify exposition, consider a simple model in which the structural equation of interest is

$$\mathbf{y}_1 = \mathbf{y}_2\beta + \mathbf{u} \tag{1}$$

where $\mathbf{y}_1$ and $\mathbf{y}_2$ are $n \times 1$ vectors of observations on two endogenous variables, $\mathbf{u}$ is an $n \times 1$ unobserved disturbance vector, and $\beta$ is an unknown scalar parameter. This equation is assumed to be part of a larger linear simultaneous equations model, which implies that $\mathbf{y}_2$ is correlated with $\mathbf{u}$. The complete system contains exogenous variables that can be used as instruments for conducting inference on $\beta$. Specifically, it is assumed that the reduced form for $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2]$ can be written as

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{Z}\boldsymbol{\pi}\beta + \mathbf{v}_1 \\ \mathbf{y}_2 &= \mathbf{Z}\boldsymbol{\pi} + \mathbf{v}_2 \end{aligned} \tag{2}$$

where $\mathbf{Z}$ is an $n \times k$ matrix of exogenous variables having full column rank $k$ and $\boldsymbol{\pi}$ is a $k \times 1$ vector; the $n$ rows of the $n \times 2$ matrix of reduced form errors $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2]$ are i.i.d. normally distributed with mean zero and $2 \times 2$ nonsingular covariance matrix $\boldsymbol{\Omega} = [\omega_{i,j}]$. The goal here is to test the null hypothesis $H_0\colon \beta = \beta_0$ against the alternative $H_1\colon \beta \neq \beta_0$.

Commonly used tests reject the null hypothesis when a test statistic $\mathcal{J}$ takes on a value greater than a specified critical value $c$. The test is said to have size $\alpha$ if, when the null hypothesis is true,

$$\Pr(\mathcal{J} > c) \leq \alpha$$

for all admissible values of the nuisance parameters $\boldsymbol{\pi}$ and $\boldsymbol{\Omega}$. Since $\boldsymbol{\pi}$ and $\boldsymbol{\Omega}$ are unknown, finding a test with correct size is nontrivial. Of course, if the null distribution of $\mathcal{J}$ does not depend on the nuisance parameters, the $1-\alpha$ quantile of $\mathcal{J}$ can be used for $c$ and the null rejection probability will be identically equal to $\alpha$. In that case, $\mathcal{J}$ is said to be *pivotal* and the test is said to be *similar*.

In practice, one often uses test statistics that are only asymptotically pivotal:

$$\lim_{n\to\infty} \Pr\left(\mathcal{J} > c_\alpha\right) = \alpha$$

However, the actual size of the test may differ substantially from the size based on the asymptotic distribution of $\mathcal{J}$. In fact, based on earlier work by Gleser and Hwang (1987), Dufour (1997) shows that the true levels of the usual Wald tests deviate arbitrarily from their nominal levels if $\boldsymbol{\pi}$ cannot be bounded away from the origin. Since weak instruments are common in empirical research, it would be desirable to find tests with approximately correct size $\alpha$ even when $\boldsymbol{\pi}$ cannot be bounded away from the origin.

## 2.2 Known covariance matrix

To ease exposition, suppose for now that besides the assumption of normality $\mathbf{\Omega}$ is known. In this case, the $k \times 2$ matrix $\mathbf{Z}'\mathbf{Y}$ is a sufficient statistic for the unknown parameters $(\beta, \boldsymbol{\pi})$. Hence, without loss of generality, any test depends on the data only through $\mathbf{Z}'\mathbf{Y}$. However, for any known nonsingular, nonrandom $2 \times 2$ matrix $\mathbf{D}$, the $k \times 2$ matrix $\mathbf{Z}'\mathbf{Y}\mathbf{D}$ is also sufficient. A convenient choice is the matrix $\mathbf{D} = [\mathbf{b}, \mathbf{\Omega}^{-1}\mathbf{a}]$, where $\mathbf{b} = (1, -\beta_0)'$ and $\mathbf{a} = (\beta_0, 1)'$. Then the sufficient statistic can be represented by the pair of $k \times 1$ vectors

$$\mathbf{S} = \mathbf{Z}'\mathbf{Y}\mathbf{b} = \mathbf{Z}'(\mathbf{y}_1 - \beta_0 \mathbf{y}_2) \qquad \text{and} \qquad \mathbf{T} = \mathbf{Z}'\mathbf{Y}\mathbf{\Omega}^{-1}\mathbf{a}$$

which are two independent, normally distributed vectors, $\mathbf{T}$ having a null distribution depending on $\boldsymbol{\pi}$ and $\mathbf{S}$ having a null distribution not depending on $\boldsymbol{\pi}$.

The goal is to find a similar test at level $\alpha$ based on a test statistic $\psi(\mathbf{S}, \mathbf{T}, \mathbf{\Omega}, \beta_0)$. The following approach is suggested by the analysis in Lehmann (1986, Chapter 4). Although the marginal distribution of $\psi$ may depend on $\boldsymbol{\pi}$, the conditional null distribution of $\psi$ *given that* $\mathbf{T}$ *takes on the value* $\mathbf{t}$ does not depend on $\boldsymbol{\pi}$ at all. As long as the conditional distribution is continuous, its $(1-\alpha)$-quantile $c(\mathbf{t}, \mathbf{\Omega}, \beta_0, \alpha)$ can be computed and used to construct the similar test that rejects $H_0\colon \beta = \beta_0$ if

$$\psi(\mathbf{S}, \mathbf{T}, \mathbf{\Omega}, \beta_0) > c_\psi(\mathbf{T}, \mathbf{\Omega}, \beta_0, \alpha)$$

Furthermore, Moreira (2001a) shows that $\mathbf{T} = \mathbf{a}'\mathbf{\Omega}^{-1}\mathbf{a} \cdot \mathbf{Z}'\mathbf{Z}\widehat{\boldsymbol{\pi}}$, where $\widehat{\boldsymbol{\pi}}$ is the maximum likelihood estimator of $\boldsymbol{\pi}$ when $\beta$ is constrained to take the null value $\beta_0$ and $\mathbf{\Omega}$ is known. Therefore, this method of finding similar tests can be interpreted as adjusting the critical value based on a preliminary estimate of $\boldsymbol{\pi}$. We illustrate below the conditional method to the four test statistics included in the commands accompanying this paper.

**Example 1** *The Anderson–Rubin statistic for known* $\mathbf{\Omega}$ *is*

$$AR = \mathbf{S}'\left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{S}/\sigma_0^2$$

*where* $\sigma_0^2 = \mathbf{b}'\mathbf{\Omega}\mathbf{b}$. *The distribution of* $AR$ *is* $\chi^2(k)$ *under the null hypothesis and its critical value function collapses to a constant*

$$c_{AR}(\mathbf{t}, \mathbf{\Omega}, \beta_0, \alpha) = q_\alpha(k)$$

*where* $q_\alpha(df)$ *is the* $1-\alpha$ *quantile of a* $\chi^2$ *distribution with* $df$ *degrees of freedom.*

**Example 2** *A score statistic is given by*

$$LM = \frac{(\mathbf{S}'\widehat{\boldsymbol{\pi}})^2}{\sigma_0^2 \widehat{\boldsymbol{\pi}}'\mathbf{Z}'\mathbf{Z}\widehat{\boldsymbol{\pi}}}$$

*The null distribution of* $LM$ *is* $\chi^2(1)$ *and its critical value function collapses to a constant*

$$c_{LM}(\mathbf{t}, \mathbf{\Omega}, \beta_0, \alpha) = q_\alpha(1)$$

**Example 3** *The Wald statistic centered around the 2SLS estimator is given by*

$$W = (b_{2SLS} - \beta_0)' \mathbf{y}_2' \mathbf{N}_Z \mathbf{y}_2 (b_{2SLS} - \beta_0) / \hat{\sigma}^2$$

*where* $\mathbf{N}_Z = \mathbf{Z}(\mathbf{Z}'\mathbf{Z}^{-1})\mathbf{Z}'$, $b_{2SLS} = (\mathbf{y}_2'\mathbf{N}_Z\mathbf{y}_2)^{-1}\mathbf{y}_2'\mathbf{N}_Z\mathbf{y}_1$, *and* $\hat{\sigma}^2 = [1 \; -b_{2SLS}]\mathbf{\Omega}[1 \; -b_{2SLS}]'$. *Here, the nonstandard structural error variance estimate exploits the fact that* $\mathbf{\Omega}$ *is known. The critical value function for* $W$ *can be simplified to*

$$c_W (\mathbf{T}, \mathbf{\Omega}, \beta_0, \alpha) = \bar{c}_W (\tau, \mathbf{\Omega}, \beta_0, \alpha)$$

*where* $\tau \equiv \mathbf{t}'(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{t} / \left(\mathbf{a}'\mathbf{\Omega}^{-1}\mathbf{a}\right)$.

**Example 4** *The likelihood-ratio statistic, for a known* $\mathbf{\Omega}$, *is given by*

$$LR = \frac{1}{2}\left(\overline{\mathbf{S}}'\overline{\mathbf{S}} - \overline{\mathbf{T}}'\overline{\mathbf{T}} + \left[\left(\overline{\mathbf{S}}'\overline{\mathbf{S}} + \overline{\mathbf{T}}'\overline{\mathbf{T}}\right)^2 - 4\left\{\overline{\mathbf{S}}'\overline{\mathbf{S}} \cdot \overline{\mathbf{T}}'\overline{\mathbf{T}} - (\overline{\mathbf{S}}'\overline{\mathbf{T}})^2\right\}\right]^{\frac{1}{2}}\right)$$

*where* $\overline{\mathbf{S}} = \left(\mathbf{b}'\mathbf{\Omega}\mathbf{b} \cdot \mathbf{Z}'\mathbf{Z}\right)^{-1/2}\mathbf{S}$ *and* $\overline{\mathbf{T}} = \left(\mathbf{a}'\mathbf{\Omega}^{-1}\mathbf{a} \cdot \mathbf{Z}'\mathbf{Z}\right)^{-1/2}\mathbf{T}$. *The critical value function for the likelihood-ratio test has the form*

$$c_{LR} (\mathbf{T}, \mathbf{\Omega}, \beta_0, \alpha) = \bar{c}_{LR} (\tau, \alpha)$$

*Note that it does not depend directly on* $\mathbf{\Omega}$ *and* $\beta_0$.

To implement the conditional procedure based on a statistic $\psi$, this paper's commands compute the conditional quantile $c_\psi(\mathbf{t}, \mathbf{\Omega}, \beta_0, \alpha)$ using Monte Carlo simulation from the *known* null distribution of $\mathbf{S}$. Indeed, the commands need only do a simulation for the actual value $\mathbf{t}$ observed in the sample and for the particular $\beta_0$ being tested; there is no need to derive the whole critical value function $c_\psi(\mathbf{t}, \mathbf{\Omega}, \beta_0, \alpha)$.

# 3 Extensions

The commands also extend the previous theory to a structural equation with additional exogenous variables. Consider the structural equation

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

where the "underlying" equation that relates the endogenous explanatory variable and the instruments is given by

$$\mathbf{y}_2 = \mathbf{Z}\boldsymbol{\pi} + \mathbf{X}\boldsymbol{\delta} + \mathbf{v}_2$$

The unknown parameters associated with $\mathbf{X}$ can be eliminated by taking orthogonal projections and considering the statistics

$$\boldsymbol{\mathcal{S}} = \mathbf{Z}'\mathbf{M}_X\mathbf{Y}\mathbf{b} \qquad \text{and} \qquad \boldsymbol{\mathcal{T}} = \mathbf{Z}'\mathbf{M}_X\mathbf{Y}\mathbf{\Omega}^{-1}\mathbf{a}$$

where $\mathbf{M}_X = \mathbf{I} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}$. For a nonpivotal statistic $\psi(\boldsymbol{\mathcal{S}}, \boldsymbol{\mathcal{T}}, \boldsymbol{\Omega}, \beta_0)$, the commands find the critical value by computing the $1-\alpha$ quantile of the distribution of $\psi$ conditioned on $\boldsymbol{\mathcal{T}} = \mathbf{t}$.

For the case in which the error distribution is unknown, Moreira (2002) shows that the conditional approach can be modified by replacing $\boldsymbol{\Omega}$ by the consistent OLS estimator. Weak-instrument asymptotics developed by Staiger and Stock (1997) and Monte Carlo evidence show that this modification does not affect significantly the size and power of the resulting test.

The commands also compute confidence regions by inverting the conditional tests; that is, the whole region in which each statistic is below its critical value curve. Unlike Wald-type confidence intervals, the confidence regions based on the conditional tests have correct coverage probability even when the instruments are weak and are also informative when instruments are good.

# 4 Stata implementation

## 4.1 Syntax

`condivreg` *depvar* [ *varlist*$_1$ ] `(`*endogvar* = *varlist*$_{\texttt{IV}}$`)` [`if` *exp*] [`in` *range*]
 [ , [`2sls`|`liml`] `nocons` `noinstcons` `level(`#`)` ]

`condtest` [ , `beta(`#`)` `reps(`#`)` `level(`#`)` ]

`condgraph` , `stats(`*string*`)` [`reps(`#`)` `points(`#`)` `range(`# #`)` `level(`#`)`
 `saving(`*filename*`)` `replace` `text` `comma`]

## 4.2 Options

### Options for condivreg

`2sls` requests that the 2SLS estimator be used. `2sls` is the default.

`liml` requests that the LIML estimator instead be used. `2sls` and `liml` are mutually exclusive.

`nocons` indicates that no constant term is to be included in the regression equation. The default is to include a constant term.

`noinstcons` indicates that no constant term is to be included in the first-stage regression of the endogenous variable on the instruments and the exogenous variables. Stata's `ivreg` command excludes a constant from both equations if its `noconstant` option is specified. Usually, one will not want to specify `noinstcons` unless `nocons` is also specified, but we give the user the option to experiment. By default, a constant term is included.

`level(#)` specifies the nominal significance level to be used when displaying the results. The default is to use the value stored in the global macro `$S_level`.

### Options for condtest

`beta(#)` contains the hypothesized value $\beta_0$ of the parameter on the endogenous variable. If this option is not specified, a default value of $\beta_0 = 0$ is used.

`reps(#)` specifies the number of simulations to perform to compute the critical values of the test statistics. The default is 200.

`level(#)` specifies the nominal significance level to be used when displaying the results. The default is to use the value stored in the global macro `$S_level`.

### Options for condgraph

`stats(`*string*`)` lists the test statistics and critical values to be included in the graph. Any one or two of the following may be specified: `AR`, `LM`, `LR`, and `Wald`, for the Anderson–Rubin, Lagrange multiplier (score), likelihood-ratio, and Wald statistics, respectively. `stats(`*string*`)` is not optional.

`reps(#)` specifies the number of simulations to perform for each value of $\beta_0$ plotted on the graph. The default is 200.

`points(#)` specifies the number of equally spaced values of $\beta_0$ to include on the graph. The default is 20.

`range(# #)` takes two numbers representing the minimum and maximum values of $\beta_0$ to include in the graph. The default is to use an interval centered at the 2SLS or LIML estimate from the previous `condivreg` result, with a radius twice that of a confidence interval based on the confidence level specified by `level(#)`. That is, if $\hat{\beta}$ is the estimated value of the parameter on the endogenous variable, $\hat{\sigma}_\beta$ is its standard error, and $1-\alpha$ is the confidence level, then the default endpoints for the graph are $\hat{\beta} \pm 2z_{\alpha/2}\hat{\sigma}_\beta$.

`level(#)` specifies the nominal significance level to be used when presenting the results. The default is to use the value stored in the global macro `$S_level`.

`saving(`*filename*`)` requests that a file be saved that contains, for each value of $\beta_0$, the four test statistics along with their critical values. The file is saved in Stata's `.dta` dataset format with the name *filename*`.dta` unless the `text` option is specified.

`replace` instructs the program to replace any existing version of the file when saving to disk. The default is to print out an error message and not change the file on disk. `replace` can only be specified if `saving(`*filename*`)` is used.

`text` requests that the test statistics be saved as a text file instead of a Stata dataset. The filename will be *filename*`.out`, and the columns will be tab-delimited. If `text` is requested without the `saving(`*filename*`)` option, an error message is printed.

`comma` requests that commas be used as column delimiters instead of tabs. If `comma` is requested without the `saving(`*filename*`)` and `text` options, an error message is printed.

## 4.3 Remarks

Three commands are available to implement the test procedures described in the previous sections. The first command, `condivreg`, estimates a regression equation with an endogenous regressor using either 2SLS or LIML. We are unaware of any other widely available programs for LIML estimation within Stata. Two commands, `condtest` and `condgraph`, allow for post-estimation testing and construction of confidence intervals.

The first step in conducting hypothesis tests involving the endogenous variable is to fit the regression model using the `condivreg` command The syntax is similar but not identical to that for Stata's built-in `ivreg` command. *depvar* is the dependent variable in the model, and the optional $varlist_1$ contains the exogenous variables. *endogvar* denotes the endogenous variable in the equation; currently, both the theory and the command are limited to the case of a single endogenous variable. $varlist_{\texttt{IV}}$ includes the instruments to be used. The `if` and `in` modifiers work in the usual way.

After the equation has been estimated, the user can perform conditional tests involving the parameter on *endogvar* using the `condtest` command. Additionally, the `condgraph` command can be used to create graphs showing the test statistics and critical values for a range of null hypotheses. Using these graphs, the user can then determine confidence intervals.

`condivreg` and `condtest` save results in `e()` and `r()`, respectively. Details are relegated to *Saved Results*. The following examples illustrate the commands' usage.

## 4.4 A simple example

To illustrate how one uses these commands, we use the same dataset and regression specification that is used in [R] **ivreg**. The first step is to perform the regression:

(*Continued on next page*)

```
. use http://www.stata-press.com/data/r7/hsng2.dta, clear
(1980 Census housing data)

. condivreg rent pcturban (hsngval = faminc reg2-reg4)

Instrumental variables (2SLS) regression

First-stage results                                     Number of obs =      50
-----------------------                                 F(  2,    47) =   42.66
F(  5,    44) =   19.66                                 Prob > F      =  0.0000
Prob > F      =  0.0000                                 R-squared     =  0.5989
R-squared     =  0.6908                                 Adj R-squared =  0.5818
Adj R-squared =  0.6557                                 Root MSE      =  22.862

------------------------------------------------------------------------------
        rent |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     hsngval |   .0022398   .0003388     6.61   0.000     .0015583    .0029213
    pcturban |    .081516   .3081528     0.26   0.793    -.5384074    .7014394
       _cons |   120.7065   15.70688     7.68   0.000     89.10834    152.3047
------------------------------------------------------------------------------
Instrumented:  hsngval
Instruments:   pcturban faminc reg2 reg3 reg4
------------------------------------------------------------------------------
For hypothesis testing and confidence regions for
the parameter on hsngval use condtest and condgraph.
------------------------------------------------------------------------------
```

Notice that the coefficient estimates and their standard errors, as well as the summary statistics in the upper right-hand corner of the output, are identical to those shown in the first example of [R] **ivreg** since we (by default) selected the 2SLS estimator. Instead of showing an ANOVA table, however, `condivreg` displays statistics from the first stage regression of the endogenous variable `hsngval` on the instruments `pcturban`, `faminc`, `reg2`, `reg3`, and `reg4`. If the first-stage $R^2$ was low, then traditional tests involving the parameter on `hsngval` could be very misleading; of course, the tests implemented by the commands accompanying this article possess good statistical properties even in that case.

Suppose that we wanted to test the null hypothesis that the parameter on `hsngval` equaled 0.002. To do this, we use the `condtest` command. Because the critical values are based on simulation, we will set the random number seed so that the results can be reproduced.

(*Continued on next page*)

```
. set seed 123
. condtest, beta(0.002) reps(1000)
Size-adjusted tests based on
Moreira's (2002) conditional approach.
H0: b[hsngval] = 0.0020
Critical values based on 1000 simulations.
```

| Statistic | Value | 95% C.V. | Asy. C.V.* |
|---|---|---|---|
| Anderson-Rubin | 15.9699 | 9.4877 | |
| Likelihood Ratio | 4.4091 | 3.7422 | 3.8415 |
| Lagrange Multiplier (Score) | 3.9811 | 3.8415 | |
| Wald | 0.6198 | 4.7326 | 3.8415 |

```
*Asy. C.V. denotes the usual asymptotic chi-square-one critical
 values for the Wald and likelihood ratio test statistics.
```

Here, we asked that 1,000 simulations be performed in computing the critical values for the likelihood-ratio and Wald test statistics. At the 5% significance level, based on the Anderson–Rubin, score, and conditional likelihood-ratio tests, we can reject the null hypothesis that $\beta = 0.002$ because each one exceeds its corresponding 95% critical value. However, the conditional Wald test does not allow us to reject that null hypothesis. Also shown in the output are the standard $\chi^2(1)$ critical values associated with the traditional unmodified likelihood-ratio and Wald tests. The size-correct Lagrange multiplier statistic always has the same critical value as its asymptotic $\chi^2(1)$ counterpart. The Anderson–Rubin statistic has a $\chi^2(k)$ distribution, where $k$ is the number of exogenous variables excluded from the main equation; here $k = 4$.

Now, suppose that we would like to make a graph of the conditional likelihood-ratio test statistic for a range of values of $\beta_0$ so that we can determine an approximate 95% confidence interval. As shown by Moreira (2001b, 2002), the conditional likelihood-ratio test has better overall power properties than the other three tests. Since our last estimation results are still those from `condivreg`, we do not need to refit the model. Making confidence intervals is easy when we save the data on the graph, so we do that here.

```
. set matsize 110
. condgraph, stats(lr) points(50) reps(1000) range(0.0018 0.0040) saving(grpoints)
```

`condgraph` requires that the `matsize` be set at least as high as the number of points in the graph plus the number of instrumental and exogenous variables in the equation being estimated. Also, note that `condgraph` is computationally intensive. Figure 1 shows the graph from the prior command.

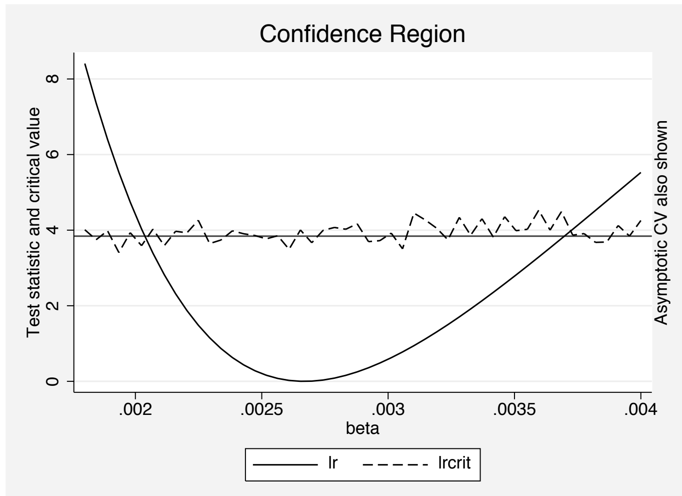


Figure 1: Conditional likelihood-ratio test statistics and critical values

Whenever the Wald or likelihood-ratio statistics are requested, the asymptotic $\chi^2(1)$ critical value is also plotted on the graph for comparison purposes. In this example, the simulated critical values are similar to their asymptotic counterparts. However, the discrepancy can be large if the instruments are weak.

As discussed in Section 3, the confidence region is the region of the graph where the observed statistic lies below its critical value. Looking at the graph, the 95% confidence region based on the likelihood-ratio statistics appears to be roughly $[0.020, 0.039]$. Since we saved the data plotted on the graph with the filename `grpoints.dta`, we can get a more accurate confidence interval:

```
. use grpoints, clear
. sort beta
. list beta lr lrcrit if lr[_n-1] > lrcrit[_n-1] & lr[_n] <= lrcrit[_n] & _n > 1
```

| | beta | lr | lrcrit |
|---|---|---|---|
| 7. | .0020694 | 3.39149 | 4.027182 |

```
. list beta lr lrcrit if lr[_n-1] <= lrcrit[_n-1] & lr[_n] > lrcrit[_n] & _n > 1
```

| | beta | lr | lrcrit |
|---|---|---|---|
| 44. | .0037306 | 4.023512 | 3.867011 |

The first `list` command finds the value $\beta^*$ such that for all $\beta < \beta^*$, the likelihood-ratio statistic is greater than the corresponding critical value, and for all $\beta \geq \beta^*$, the statistic is less than its critical value. Analogously, the second `list` command finds the right endpoint of the confidence interval. Hence, the 95% confidence interval is approximately $[\,0.00207, 0.00373\,]$. Notice that this 95% confidence interval based on the conditional likelihood-ratio test statistic differs quite substantially from the traditional one reported by `condivreg`. In this example, the conditional likelihood-ratio confidence interval lies to the right of the usual one, and it is slightly wider.

The `test` command can also be used to compare the size-correct Wald statistic with the traditional Wald statistic. Since we just loaded in the graph data, we first have to reload the original dataset; we do not need to refit the model since the estimation results remain current.

```
. use http://www.stata-press.com/data/r7/hsng2.dta, clear
(1980 Census housing data)
. test hsngval = 0.002
 ( 1)  hsngval = .002
       F(  1,    47) =    0.50
            Prob > F =    0.4825
```

Notice that the Wald test statistic for the hypothesis $H_0 : \beta_0 = 0.002$ computed by `condtest` was 0.62, while here the Wald statistic is 0.50. The difference arises because different estimators for the variance of the disturbances in (1) are being used.

## 4.5 A weak-instrument example

In the previous example, the instruments were quite highly correlated with the endogenous variable: the first-stage $R^2$ was 0.69. As Figure 1 showed, the simulated critical value of the size-correct likelihood ratio statistic was close to its asymptotic counterpart. However, in many applications, the correlation is often low. This example illustrates the practical ramifications.

To begin, we load in an example dataset and fit a model using `condivreg`:

```
. use example, clear
. condivreg y1 x1 (y2 = z1-z3)
Instrumental variables (2SLS) regression
First-stage results                                         Number of obs =     400
-----------------------                                     F(  2,   397) =   24.76
F(  4,   395) =    4.10                                     Prob > F      =  0.0000
Prob > F      =  0.0029                                     R-squared     =  0.6696
R-squared     =  0.0399                                     Adj R-squared =  0.6679
Adj R-squared =  0.0302                                     Root MSE      =  32.338
------------------------------------------------------------------------------
          y1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          y2 |   1.011903   .1947045     5.20   0.000     .6291221    1.394684
          x1 |    4.15736   1.702533     2.44   0.015     .8102529    7.504466
       _cons |   .7493957   1.655456     0.45   0.651    -2.505161    4.003952
------------------------------------------------------------------------------
Instrumented:  y2
Instruments:   x1 z1 z2 z3
------------------------------------------------------------------------------
For hypothesis testing and confidence regions for
the parameter on y2 use condtest and condgraph.
------------------------------------------------------------------------------
```

Notice that in this example, the first-stage $R^2$ is only 0.04. Next, we use `condgraph` to plot the likelihood-ratio statistic and its critical values for various levels of $\beta_0$:

```
. condgraph, stats(lr) points(50) reps(1000) range(0 2)
```

The results are shown in Figure 2.

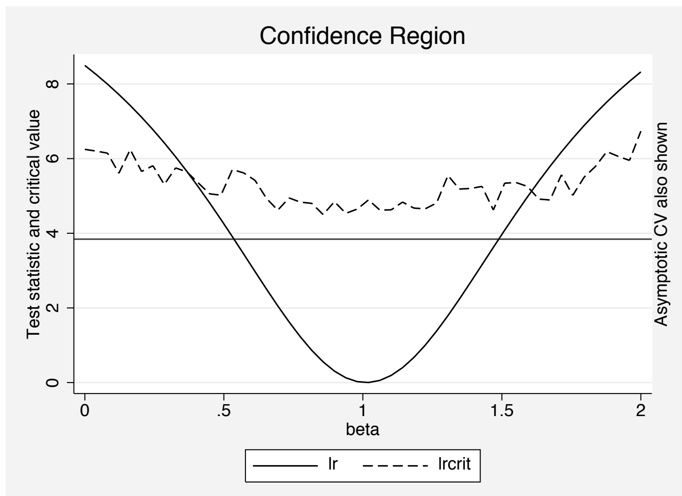


Figure 2: Conditional likelihood-ratio test statistics and critical values—weak instruments

The simulated critical value function lies above the asymptotic $\chi^2(1)$ counterpart and highlights the fact that inference based on the usual tests may be misleading. Moreover, as the graph illustrates, traditional confidence intervals would be too narrow. In this example, the size-correct confidence interval is approximately $[\,0.4082, 1.6327\,]$, while the confidence interval using the likelihood-ratio test statistic and the asymptotic critical value is $[\,0.5714, 1.5102\,]$. The asymptotic confidence interval is over 30% too narrow here.

(*Continued on next page*)

# 5 Saved Results

`condivreg` saves in `e()`:

| | | | |
|---|---|---|---|
| Scalars | | | |
| `e(N)` | number of observations | `e(rss)` | residual sum of squares |
| `e(df_m)` | model degrees of freedom | `e(F_first)` | first-stage $F$ statistic |
| `e(df_r)` | residual degrees of freedom | `e(df_m_first)` | first-stage model degrees of freedom |
| `e(F)` | $F$ statistic | `e(df_r_first)` | first-stage residual degrees of freedom |
| `e(r2)` | $R$-squared | `e(r2_first)` | first-stage $R$-squared |
| `e(r2_a)` | adjusted $R$-squared | `e(r2_a_first)` | first-stage adjusted $R$-squared |
| `e(rmse)` | root mean-square error | | |
| `e(mss)` | model sum of squares | | |
| Macros | | | |
| `e(cmd)` | `condivreg` | `e(instd)` | instrumented variable |
| `e(cons)` | `yes` or `no`—constant in model | `e(insts)` | instruments |
| `e(instcons)` | `yes` or `no`—constant in instruments list | `e(inst)` | excluded exogenous variables |
| `e(model)` | `2sls` or `liml` | `e(exog)` | included exogenous variables |
| | | `e(depvar)` | dependent variable |
| Matrices | | | |
| `e(b)` | coefficient vector | `e(V)` | variance–covariance matrix |
| Functions | | | |
| `e(sample)` | marks estimation sample | | |

`condtest` saves in `r()`:

| | | | |
|---|---|---|---|
| Scalars | | | |
| `r(beta)` | hypothesized value of beta | `r(lmcrit)` | critical value of Lagrange multiplier statistic |
| `r(ar)` | Anderson–Rubin statistic | `r(lrcrit)` | critical value of likelihood-ratio statistic |
| `r(lm)` | Lagrange multiplier statistic | `r(waldcrit)` | critical value of Wald statistic |
| `r(lr)` | likelihood-ratio statistic | | |
| `r(wald)` | Wald statistic | | |
| `r(arcrit)` | critical value of Anderson–Rubin statistic | | |

# 6 References


Anderson, T. and H. Rubin. 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics* 20: 46–63.

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

Dufour, J.-M. 1997. Some impossibility theorems in econometrics with applications to structural and dynamic models. *Econometrica* 65: 1365–1388.

Gleser, L. and J. Hwang. 1987. The non-existence of $100(1-\alpha)\%$ confidence sets of finite expected diameter in errors-in-variables and related models. *Annals of Statistics* 15: 1351–1362.

Lehmann, E. 1986. *Testing Statistical Hypotheses*. 2d ed. New York: John Wiley & Sons.

Moreira, M. 2001a. Tests with correct size when instruments can be arbitrarily weak. Center for Labor Economics Working Paper 37, UC Berkeley.

—. 2001b. A conditional likelihood ratio test for structural models. Manuscript, UC Berkeley.

—. 2002. *Tests with correct size in the simultaneous equations model.* Ph.D. thesis, UC Berkeley.

Staiger, D. and J. H. Stock. 1997. Instrumental variables regression with weak instruments. *Econometrica* 65(3): 557–586.

### About the Authors

Marcelo J. Moreira received his Ph.D. in economics and M.A. in statistics from the University of California at Berkeley and is currently an assistant professor in the economics department at Harvard University.

Brian P. Poi received his Ph.D. in economics from the University of Michigan and joined Stata Corporation as a staff statistician.