Implementing Conditional Tests with Correct Size in the Simultaneous Equations Model

Marcelo J. Moreira
Department of Economics
Harvard University

Brian P. Poi
Stata Corporation

Abstract. This paper fixes 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, Jaeger and Baker (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 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 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 known

covariance matrix. Sections 3 extends the results to more realistic cases and explains how to construct confidence regions based on the conditional test. Finally, Section 4 considers the syntax of the program and illustrates a few examples.

2   The Conditional Approach

2.1   The Model

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

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

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

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

where $\boldsymbol{Z}$ 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 $\boldsymbol{V} = [\boldsymbol{v}_1, \boldsymbol{v}_2]$ are i.i.d. normally distributed with mean zero and 2×2 nonsingular covariance matrix $\boldsymbol{\Omega} = [\omega_{i,j}]$. The goal here is to test the null hypothesis $H_0 : \beta = \beta_0$ against the alternative $H_1 : \beta \neq \beta_0$.

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 α if, when the null hypothesis is true,

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

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 α. 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:

$$\lim_{n \to \infty} \mathrm{prob}\,(\mathcal{J} > c_\alpha) = \alpha$$

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

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

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

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

$$\psi\left(\boldsymbol{S}, \boldsymbol{T}, \boldsymbol{\Omega}, \beta_0\right) > c_\psi\left(\boldsymbol{T}, \boldsymbol{\Omega}, \beta_0, \alpha\right)$$

Furthermore, Moreira (2001a) shows that $\boldsymbol{T} = \boldsymbol{a'\Omega}^{-1}\boldsymbol{a} \cdot \boldsymbol{Z'Z}\widehat{\boldsymbol{\pi}}$, where $\widehat{\boldsymbol{\pi}}$ is the maximum likelihood estimator of $\boldsymbol{\pi}$ when β is constrained to take the null value β_0 and $\boldsymbol{\Omega}$ 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 package.

Example 1 *The Anderson-Rubin statistic for known $\boldsymbol{\Omega}$ is*

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

The distribution of AR is $\chi^2(k)$ under the null hypothesis and its critical value function collapses to a constant

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

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

Example 2 *A score statistic is given by:*

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

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

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

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

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

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

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

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

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

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

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

$$c_{LR} \left(\boldsymbol{T}, \boldsymbol{\Omega}, \beta_0, \alpha \right) = \bar{c}_{LR} \left(\tau, \alpha \right).$$

Note that it does not dependent directly on $\boldsymbol{\Omega}$ and β_0.

To implement the conditional procedure based on a statistic ψ, the package computes the conditional quantile $c_\psi(\boldsymbol{t}, \boldsymbol{\Omega}, \beta_0, \alpha)$ using Monte Carlo simulation from the *known* null distribution of $\boldsymbol{S}$. Indeed, the package need only do a simulation for the actual value $\boldsymbol{t}$ observed in the sample and for the particular β_0 being tested; there is no need to derive the whole critical value function $c_\psi(\boldsymbol{t}, \boldsymbol{\Omega}, \beta_0, \alpha)$.

3   Extensions

The package also extends the previous theory to a structural equation with additional exogenous variables. Consider the structural equation

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

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

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

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

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

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

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 shows that this modification does not affect significantly the size and power of the resulting test.

The package also computes 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 Command syntax

The software package accompanying this article contains three Stata commands to implement the tests discussed above. 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. First we discuss the syntax of the commands, and then we illustrate their use with an example.

The first step in conducting hypothesis tests involving the endogenous variable is to fit the regression model using the `condivreg` command:

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

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*$_{\text{IV}}$ includes the instruments to be used. The `if` and `in` modifiers work in the usual way. There are five options that may be specified:

> `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.

The 2SLS estimator is implemented using the formulas shown in [R] **ivreg**, and the LIML estimator follows the derivation in Davidson and MacKinnon (1993, pp. 644–651).

After the equation has been estimated, the user can perform conditional tests involving the parameter on *endogvar* using the condtest command. The syntax is

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

This command has three options:

beta(#) contains the hypothesized value β_0 of the parameter on the endogenous variable. If this option is not specified, a default value of $\beta_0 = 0$ is used.

The reps(#) option 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.

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. The syntax is

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

The supported options are

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 β_0 plotted on the graph. The default is 200.

points(#) specifies the number of equally spaced values of β_0 to include on the graph. The default is 20.

range(*numlist*) takes two numbers representing the minimum and maximum values of β_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 which contains, for each value of β_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.

condivreg and condtest save results as e-class and r-class macros, respectively. Details are relegated to the appendix. condgraph does not save results as macros because it has the saving() option. The commands are easier to use than the syntax diagrams may lead one to believe, and the following examples clarify their usage.

4.2   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 estimate the regression:

(Continued on next page)

```
. use http://www.stata-press.com/data/r7/hsng2.dta, clear
. 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 tests on hsngval use condtest and condgraph.
------------------------------------------------------------------------------
```

Notice that the coefficient estimates and their standard errors as well as the summary statistics in the upper righthand corner of the output are identical to those shown on page 133 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 were low, then traditional tests involving the parameter on `hsngval` could be very misleading; of course, the tests implemented by this software package possess good statistical properties even in that case.

Suppose that we wanted to test the null hypothesis that the parameter on `hsngval` equalled 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-correct test statistics 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. For the size-correct Lagrange multiplier statistic, the critical value is the same 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 $k = 4$.

Suppose now we would like to make a graph of the conditional likelihood ratio test statistic for a range of values of β_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 reestimate the model. Making confidence intervals is easy when we save the data on the graph, so we will do that here.

```
. set matsize 110
. condgraph, stats(lr) points(50) 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. Because condgraph is computationally intensive, be prepared to wait awhile when requesting a large number of points on the graph and a large number of simulations. Figure 1 shows the graph from the prior command.
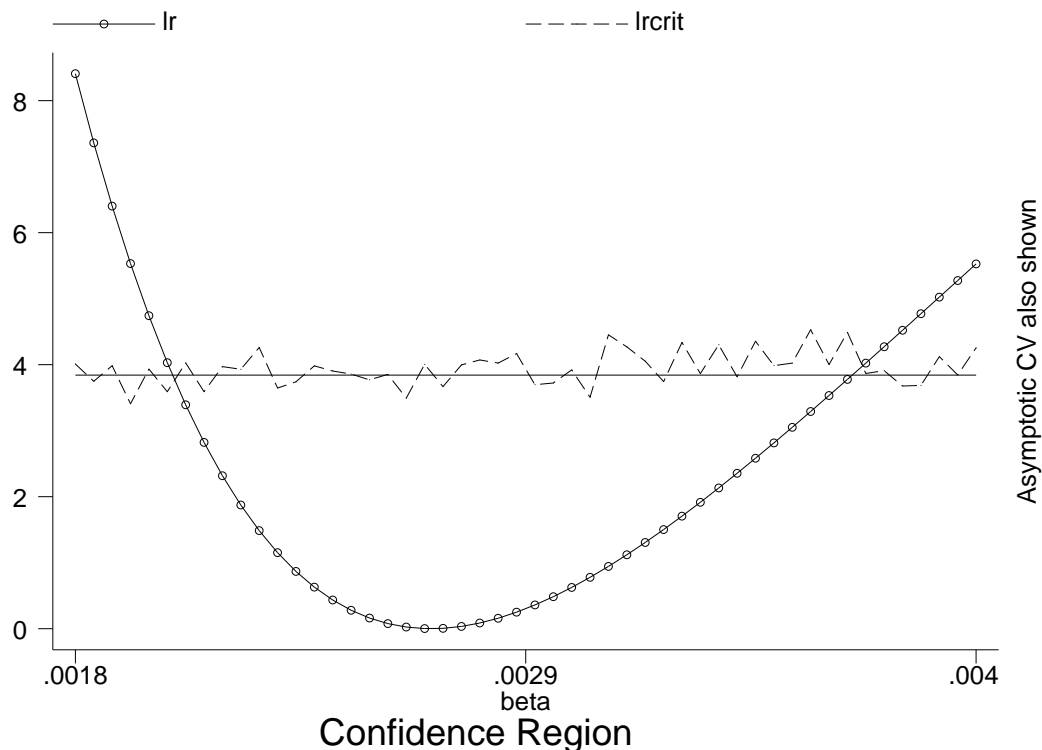
Figure 1: Conditional Likelihood Ratio Test Statistics and Critical Values

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 approximately $[0.020, 0.039]$. 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 β^* 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 reestimate the model since it is still in memory.

```
. 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 equation (1) are being used.

4.3   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; and this example illustrates the practical ramifications.

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

(Continued on next page)

```
. 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 tests 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 β_0:

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

The results are shown in Figure 2. 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.
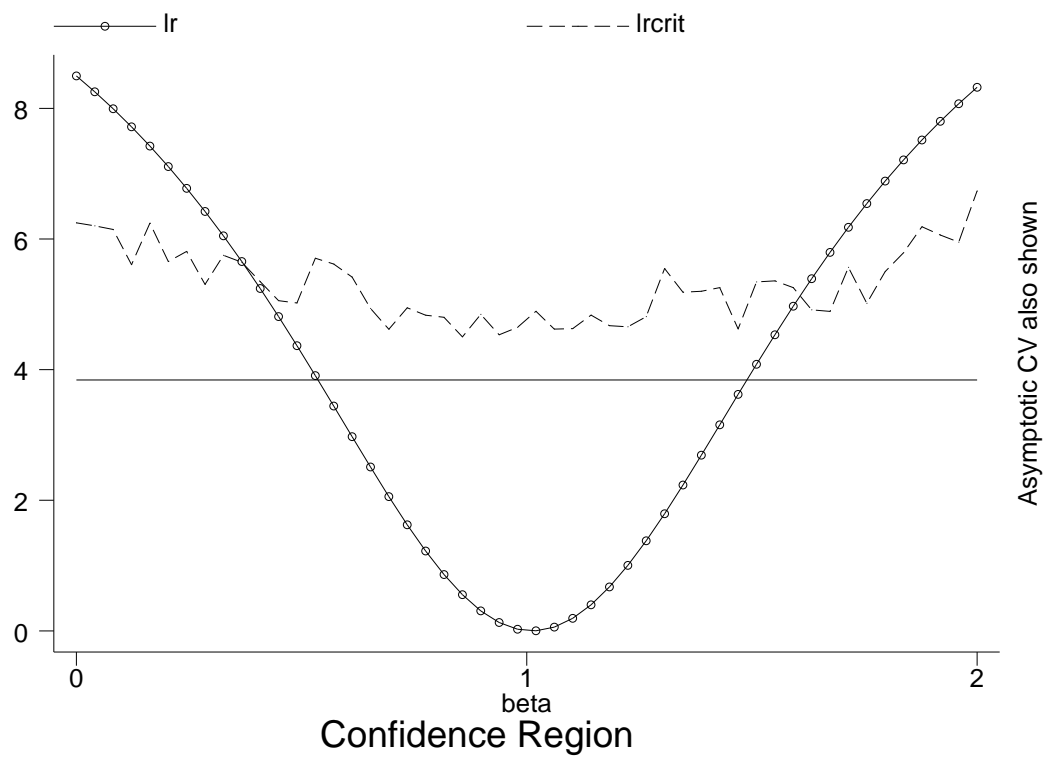
(Continued on next page)

Figure 2: Conditional Likelihood Ratio Test Statistics and Critical Values – Weak Instruments

5 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 variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak," *Journal of American Statistical Association*, 90, 443-50.

DAVIDSON, R., AND J. G. MACKINNON (1993): *Estimation and Inference in Econometrics.* New York, NY: Oxford University Press.

DUFOUR, J-M. (1997): "Some impossibility theorems in econometrics with applications to structural and dynamic models," *Econometrica*, 65, 1365-88.

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, 1351-62.

KLEIBERGEN, F.(2000): "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression," forthcoming *Econometrica.*

LEHMANN, E. (1986): *Testing Statistical Hypothesis.* 2nd edition, Wiley Series in Probability and Mathematical Statistics.

MOREIRA, M. (2001a): "Tests with Correct Size when Instruments Can Be Arbitrarily Weak,"*Center for Labor Economics Working Paper*, 37, UC Berkeley.

MOREIRA, M. (2001b): *A Conditional Likelihood Ratio Test for Structural Models.* Manuscript, UC Berkeley.

MOREIRA, M. (2002): *Tests with Correct Size in the Simultaneous Equations Model.* PhD Thesis, UC Berkeley.

STAIGER, D., AND J. STOCK (1997): "Instrumental variables regression with weak instruments," *Econometrica*, 65, 557-86.

6 Appendix – Saved Results

6.1 condivreg

`condivreg` saves in `e()`:

Scalars

e(N)	number of observations
e(df_m)	model degrees of freedom
e(df_r)	residual degrees of freedom
e(F)	F statistic

e(r2)	R-squared
e(r2_a)	adjusted R-squared
e(rmse)	root mean square error
e(mss)	model sum of squares
e(rss)	residual sum of squares
e(F_first)	first-stage F statistic
e(df_m_first)	first-stage model degrees of freedom
e(df_r_first)	first-stage residual degrees of freedom
e(r2_first)	first-stage R-squared
e(r2_a_first)	first-stage adjusted R-squared

Macros

e(cmd)	`condivreg`
e(cons)	`yes` or `no` — constant in model
e(instcons)	`yes` or `no` — constant in instruments list
e(model)	`2sls` or `liml`
e(instd)	instrumented variable
e(insts)	instruments
e(inst)	excluded exogenous variables
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

6.2 condtest

`condtest` saves in `r()`:

Scalars

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

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.