



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search

<http://ageconsearch.umn.edu>

aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

THE STATA JOURNAL

Editor

H. Joseph Newton
Department of Statistics
Texas A & M University
College Station, Texas 77843
979-845-3142; FAX 979-845-3144
jnewton@stata-journal.com

Associate Editors

Christopher Baum
Boston College

Rino Bellocco
Karolinska Institutet, Sweden and
Univ. degli Studi di Milano-Bicocca, Italy

A. Colin Cameron
University of California–Davis

David Clayton
Cambridge Inst. for Medical Research

Mario A. Cleves
Univ. of Arkansas for Medical Sciences

William D. Dupont
Vanderbilt University

Charles Franklin
University of Wisconsin–Madison

Joanne M. Garrett
University of North Carolina

Allan Gregory
Queen's University

James Hardin
University of South Carolina

Ben Jann
ETH Zurich, Switzerland

Stephen Jenkins
University of Essex

Ulrich Kohler
WZB, Berlin

Stata Press Production Manager

Stata Press Copy Editor

Editor

Nicholas J. Cox
Geography Department
Durham University
South Road
Durham City DH1 3LE UK
n.j.cox@stata-journal.com

Jens Lauritsen
Odense University Hospital

Stanley Lemeshow
Ohio State University

J. Scott Long
Indiana University

Thomas Lumley
University of Washington–Seattle

Roger Newson
Imperial College, London

Marcello Pagano
Harvard School of Public Health

Sophia Rabe-Hesketh
University of California–Berkeley

J. Patrick Royston
MRC Clinical Trials Unit, London

Philip Ryan
University of Adelaide

Mark E. Schaffer
Heriot-Watt University, Edinburgh

Jeroen Weesie
Utrecht University

Nicholas J. G. Winter
University of Virginia

Jeffrey Wooldridge
Michigan State University

Lisa Gilmore
Gabe Waggoner

Copyright Statement: The Stata Journal and the contents of the supporting files (programs, datasets, and help files) are copyright © by StataCorp LP. The contents of the supporting files (programs, datasets, and help files) may be copied or reproduced by any means whatsoever, in whole or in part, as long as any copy or reproduction includes attribution to both (1) the author and (2) the Stata Journal.

The articles appearing in the Stata Journal may be copied or reproduced as printed copies, in whole or in part, as long as any copy or reproduction includes attribution to both (1) the author and (2) the Stata Journal.

Written permission must be obtained from StataCorp if you wish to make electronic copies of the insertions. This precludes placing electronic copies of the Stata Journal, in whole or in part, on publicly accessible web sites, file servers, or other locations where the copy may be accessed by anyone other than the subscriber.

Users of any of the software, ideas, data, or other materials published in the Stata Journal or the supporting files understand that such use is made without warranty of any kind, by either the Stata Journal, the author, or StataCorp. In particular, there is no warranty of fitness of purpose or merchantability, nor for special, incidental, or consequential damages such as loss of profits. The purpose of the Stata Journal is to promote free communication among Stata users.

The *Stata Journal*, electronic version (ISSN 1536-8734) is a publication of Stata Press, and Stata is a registered trademark of StataCorp LP.

Tests and confidence sets with correct size when instruments are potentially weak

Anna Mikusheva
Department of Economics
Harvard University
Boston, MA

Brian P. Poi
StataCorp
College Station, TX
bpoi@stata.com

Abstract. We consider inference in the linear regression model with one endogenous variable and potentially weak instruments. We construct confidence sets for the coefficient on the endogenous variable by inverting the Anderson–Rubin, Lagrange multiplier, and conditional likelihood-ratio tests. Our confidence sets have correct coverage probabilities even when the instruments are weak. We propose a numerically simple algorithm for finding these confidence sets, and we present a Stata command that supersedes the one presented in Moreira and Poi (*Stata Journal* 3: 57–70).

Keywords: st0033_2, condivreg, instrumental variables, weak instruments, confidence set, similar test

1 Introduction

We consider inference on the parameter of one endogenous variable in instrumental variables (IV) regression with potentially weak instruments. Most empirical applications rely on inference based on the asymptotic normal approximation of the t statistic. That is, they perform tests for significance of the coefficient by comparing the t statistic with quantiles of the normal distribution, and they use the conventional Wald-type confidence intervals. However, in many empirically relevant situations, the correlation between the instruments and the endogenous regressor is weak, and the normal approximation of the t statistic performs poorly (Nelson and Startz 1990). As a result, the conventional test of significance on the parameter of the endogenous variable has incorrect size, and the Wald-type confidence interval has low coverage probability.

Andrews and Stock (2005) and Stock, Wright, and Yogo (2002) give excellent surveys of the literature devoted to finding tests about the coefficient β on the single included endogenous regressor that are valid in the presence of potentially weak instruments. The class of tests robust to weak identification includes the Anderson–Rubin (AR) test (Anderson and Rubin 1949), the Lagrange multiplier (LM; score) test proposed by Kleibergen (2002) and Moreira (2001), and the conditional likelihood-ratio test suggested by Moreira (2003).

Confidence set construction is a well-known dual problem to hypothesis testing. If we have a procedure for testing the hypothesis $H_0: \beta = \beta_0$ with correct size even in the presence of weak instruments, then we can construct a confidence region for the parameter that is also robust to weak instruments by inverting the test. That is, a

value β_0 belongs to a confidence set if and only if the hypothesis $H_0: \beta = \beta_0$ cannot be rejected.

Moreira and Poi (2003) introduced the Stata commands `condivreg` and `condtest` implementing the AR, score, conditional likelihood-ratio, and conditional Wald tests. They also provided the command `condgraph`, which performed a series of tests $H_0: \beta = \beta_0$, where β_0 belongs to a fine grid. The user could then construct the robust confidence set by finding the area of acceptance for the given test.

However, that procedure has several drawbacks. First, performing the conditional likelihood-ratio and the conditional Wald tests for even modestly large datasets could take several hours and is not very accurate. Both tests are based on Moreira's conditional approach, and the critical value functions for these tests are simulated from the conditional distribution of the test statistic under the null hypothesis. The simulations are computationally intensive and not always accurate.

The second obstacle is that finding a confidence set by grid testing is implementable only if we can a priori restrict possible values of the coefficient to belong to a bounded set. In most applications, we cannot make such a restriction. Gleser and Hwang (1987) and Dufour (1997) showed that if the parameter set is not bounded and we can have arbitrary weak instruments, then every almost-sure finite confidence set has zero coverage probability. That is, a confidence region robust to weak instruments must be infinite with positive probability, making a grid search unfeasible in practice. Even if we do restrict the parameter space to be bounded, grid testing can be extremely time consuming.

Fortunately, several valuable results have been obtained in the past few years. Andrews, Moreira, and Stock (2006) found a way to perform the conditional likelihood-ratio test without having to perform simulations. They also showed that the conditional Wald test has extremely low power against a large range of alternatives and that its power curve can be nonmonotonic. Andrews, Moreira, and Stock (2006) recommended not using the Wald test in practice. Mikusheva (2005) proposed algorithms that allow one to construct confidence sets by quickly and accurately inverting the AR, score, and conditional likelihood-ratio tests without having to use a grid search.

We introduce a new version of `condivreg` that implements the advances mentioned above. All `condivreg` users should upgrade to this newer version.

The paper is organized as follows. Section 2 contains a brief overview of the model and definitions of the AR, the score, and the conditional likelihood-ratio tests. Section 3 provides algorithms for inverting these tests to construct weak-instrument robust confidence sets. Section 4 describes the syntax of `condivreg` and gives an example of its usage.

2 Tests robust to weak instruments

Here we introduce notations and give a brief overview of the tests that are robust to weak instruments. The model contains a structural equation and a reduced-form equation for one endogenous regressor:

$$\mathbf{y}_1 = \mathbf{y}_2\beta + \mathbf{X}\boldsymbol{\gamma}_1 + \mathbf{u} \quad (1)$$

$$\mathbf{y}_2 = \mathbf{Z}\boldsymbol{\pi} + \mathbf{X}\boldsymbol{\xi} + \mathbf{v}_2 \quad (2)$$

Vectors \mathbf{y}_1 and \mathbf{y}_2 are $n \times 1$ endogenous variables; \mathbf{X} is an $n \times p$ matrix of exogenous regressors; \mathbf{Z} is an $n \times k$ matrix of instrumental variables; and $\beta \in R$, $\boldsymbol{\gamma}_1, \boldsymbol{\xi} \in R^p$, and $\boldsymbol{\pi} \in R^k$ are unknown parameters. We assume without loss of generality that $\mathbf{Z}'\mathbf{X} = 0$. The $n \times 2$ matrix of errors $[\mathbf{u} : \mathbf{v}_2]$ is independently and identically distributed (i.i.d.) across rows, each row being normally distributed with mean zero and nonsingular covariance matrix.

We also consider the corresponding system of reduced-form equations obtained by substituting (2) into (1):

$$\mathbf{y}_1 = \mathbf{Z}\boldsymbol{\pi}\beta + \mathbf{X}\boldsymbol{\gamma} + \mathbf{v}_1$$

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

where

$$\boldsymbol{\gamma} = \boldsymbol{\gamma}_1 + \boldsymbol{\xi}\beta \quad \text{and} \quad \mathbf{v}_1 = \mathbf{u} + \beta\mathbf{v}_2$$

The reduced-form errors are assumed to be i.i.d. normal with zero mean and covariance matrix $\boldsymbol{\Omega}$. We assume $\boldsymbol{\Omega}$ to be known. Andrews, Moreira, and Stock (2006) showed that for unknown $\boldsymbol{\Omega}$, asymptotically valid tests can be obtained by replacing $\boldsymbol{\Omega}$ with a consistent estimator of $\boldsymbol{\Omega}$. Andrews, Moreira, and Stock (2004) also pointed out that the assumption of normality can be taken away at the cost of having only asymptotically valid rather than exactly valid tests. Here by ‘‘asymptotically valid’’ we mean having asymptotically correct size both in weak- and strong-instrument asymptotics. For definitions of these two types of asymptotics, see Andrews, Moreira, and Stock (2004).

We are interested in testing the hypothesis $H_0: \beta = \beta_0$. We require the testing procedure to have correct size when the instruments are weak as well as when they are strong.

Let us define the following two statistics:

$$\mathbf{S}(\beta_0) = (\mathbf{Z}'\mathbf{Z})^{-1/2}\mathbf{Z}'\mathbf{Y}\mathbf{b}_0(\mathbf{b}_0'\boldsymbol{\Omega}\mathbf{b}_0)^{-1/2}$$

and

$$\mathbf{T}(\beta_0) = (\mathbf{Z}'\mathbf{Z})^{-1/2}\mathbf{Z}'\mathbf{Y}\boldsymbol{\Omega}^{-1}\mathbf{a}_0(\mathbf{a}_0'\boldsymbol{\Omega}^{-1}\mathbf{a}_0)^{-1/2}$$

where $\mathbf{b}_0 = (1, -\beta_0)'$, $\mathbf{a}_0 = (\beta_0, 1)'$, and $\mathbf{Y} = [\mathbf{y}_1 : \mathbf{y}_2]$.

Moreira (2003) showed that $\mathbf{S}(\beta_0)$ and $\mathbf{T}(\beta_0)$ are independent, normally distributed vectors. Under the null hypothesis, the distribution of $\mathbf{T}(\beta_0)$ depends on the value of the nuisance parameter $\boldsymbol{\pi}$, but, importantly, the distribution of $\mathbf{S}(\beta_0)$ does not. Although the marginal distribution of a test statistic may depend on $\boldsymbol{\pi}$ and thus $\mathbf{T}(\beta_0)$, the conditional distribution of that test statistic—given that $\mathbf{T}(\beta_0) = \mathbf{t}$, its value based on the sample data—does not depend on $\boldsymbol{\pi}$ at all. As discussed by Gleser and Hwang (1987) and Dufour (1997), the true levels of the usual Wald tests deviate arbitrarily from their nominal levels when $\boldsymbol{\pi}$ cannot be bounded away from the origin. By conditioning on $\mathbf{T}(\beta_0) = \mathbf{t}$, we obtain test statistics whose distributions under the null hypothesis do not depend on $\boldsymbol{\pi}$ and therefore do not suffer from such size distortions.

We also define the matrix \mathbf{Q} as

$$\mathbf{Q}(\beta_0) = \{\mathbf{S}(\beta_0) : \mathbf{T}(\beta_0)\}' \{\mathbf{S}(\beta_0) : \mathbf{T}(\beta_0)\} = \begin{Bmatrix} Q_S(\beta_0) & Q_{ST}(\beta_0) \\ Q_{ST}(\beta_0) & Q_T(\beta_0) \end{Bmatrix}$$

where $Q_S(\beta_0) = \mathbf{S}(\beta_0)'\mathbf{S}(\beta_0)$, $Q_T(\beta_0) = \mathbf{T}(\beta_0)'\mathbf{T}(\beta_0)$, and $Q_{ST}(\beta_0) = \mathbf{S}(\beta_0)'\mathbf{T}(\beta_0)$. For simplicity, we will henceforth refer to \mathbf{S} and \mathbf{T} , with their dependence on β_0 implied.

The AR test rejects the null hypothesis $H_0 : \beta = \beta_0$ at significance level α if the statistic

$$\text{AR}(\beta_0) = \mathbf{S}'\mathbf{S} = Q_S(\beta_0)$$

exceeds the $(1 - \alpha)$ quantile of the χ^2 distribution with k degrees of freedom.

The LM (score) test accepts the null hypothesis if the statistic

$$\text{LM}(\beta_0) = (\mathbf{S}'\mathbf{T})(\mathbf{T}'\mathbf{T})^{-1}(\mathbf{T}'\mathbf{S}) = \frac{Q_{ST}^2(\beta_0)}{Q_T(\beta_0)}$$

is less than the $(1 - \alpha)$ quantile of the χ^2 distribution with 1 degree of freedom.

The conditional likelihood-ratio test is based on the conditional approach proposed by Moreira (2003). He suggested a whole class of tests that use, instead of one fixed critical value, critical values that are functions of the data. The conditional likelihood-ratio test uses the statistic

$$\text{LR}(\beta_0) = \frac{1}{2} \left(Q_S(\beta_0) - Q_T(\beta_0) + \left[\{Q_S(\beta_0) + Q_T(\beta_0)\}^2 - 4 \{Q_S(\beta_0)Q_T(\beta_0) - Q_{ST}^2(\beta_0)\} \right]^{1/2} \right)$$

and critical values $m_\alpha(Q_T)$, which are functions of $Q_T(\beta_0)$. For every α , the critical value $m_\alpha(q_T)$ is chosen in such a way that the conditional probability of the LR statistic exceeding $m_\alpha(q_T)$ given that $Q_T = q_T$ is equal to α :

$$P \{ \text{LR} > m_\alpha(q_T) \mid Q_T = q_T \} = \alpha$$

The conditional likelihood-ratio test accepts the null hypothesis $H_0: \beta = \beta_0$ if $\text{LR}(\beta_0) < m_\alpha\{Q_T(\beta_0)\}$.

Previously, the critical value function $m_\alpha(q_T)$ was determined by simulation. The main problem with this approach is that for an acceptable level of accuracy, one needs many simulations. Andrews, Moreira, and Stock (2006) suggested another way of implementing the conditional likelihood-ratio test by calculating the conditional p -value of the test. Let us define a p -value function, $p(m; q_T)$, by the following conditional probability:

$$p(m; q_T) = P(\text{LR} > m \mid Q_T = q_T)$$

Then the conditional likelihood-ratio test accepts the hypothesis $H_0: \beta = \beta_0$ at the α significance level if

$$p\{\text{LR}(\beta_0); Q_T(\beta_0)\} > \alpha$$

Andrews, Moreira, and Stock (2006) proved that the function $p(m; q_T)$ is equal to

$$p(m; q_T) = 1 - 2K \int_0^1 P \left(\chi_k^2 < \frac{q_T + m}{1 + q_T s_2^2 / m} \right) (1 - s_2^2)^{(k-3)/2} ds_2 \tag{3}$$

where $K = \Gamma(k/2)/[\pi^{1/2}\Gamma\{(k-1)/2\}]$ and $\Gamma(\cdot)$ is the gamma function. They also suggested a method of calculating the conditional p -value of the test by performing numerical integration. Their procedure achieves high accuracy and is fast.

The three tests described above have correct size for weak instruments. However, they possess different power properties. The AR test is robust to misspecifications of (2) and can be used as an overidentification test. The score test should probably not be used in practice, since it is dominated by the conditional likelihood-ratio test. But for historical reasons, we include it in the command accompanying this article. According to Andrews, Moreira, and Stock (2006), the conditional likelihood-ratio test is nearly optimal in a class of invariant similar tests. It possesses better power properties than the AR and score tests for many parameters.

3 Confidence sets based on tests robust to weak instruments

This section describes algorithms for construction of confidence sets for the coefficient on the single endogenous regressor β by inverting the AR, score, and conditional likelihood-ratio tests.

Given tests that are robust to weak instruments, we can construct confidence sets by inverting these tests. One way to find the acceptance region for a given test is to perform grid testing. However, such an algorithm works only if the area of search is bounded, that is, when the parameter space is bounded or we have some knowledge about the form of the set and its approximate location. In most empirical applications, we cannot a priori restrict the parameter space to be bounded. In general, we also cannot restrict the area for a grid search since a confidence set with correct coverage probability in a case with arbitrary weak instruments has infinite length with a positive probability. The inability to use a grid search leads to the necessity of finding an algorithm to invert tests.

By definition, the AR confidence set is

$$\begin{aligned} C_{\alpha}^{\text{AR}}(Y, X, Z) &= \{\beta_0: Q_S(\beta_0) < \chi_{1-\alpha, k}^2\} \\ &= [\beta_0: \mathbf{b}'_0 \{ \mathbf{Y}' \mathbf{Z} (\mathbf{Z}' \mathbf{Z})^{-1} \mathbf{Z}' \mathbf{Y} - \chi_{1-\alpha, k}^2 \boldsymbol{\Omega} \} \mathbf{b}_0 < 0] \end{aligned}$$

which can be found by solving a quadratic inequality. As a result, the AR confidence region $C_{\alpha}^{\text{AR}}(Y, X, Z)$ can have four possible forms:

- a finite interval, $C_{\alpha}^{\text{AR}}(Y, X, Z) = [x_1, x_2]$;
- a union of two infinite intervals, $C_{\alpha}^{\text{AR}}(Y, X, Z) = (-\infty, x_1] \cup [x_2, +\infty)$;
- the whole line, $C_{\alpha}^{\text{AR}}(Y, X, Z) = (-\infty, +\infty)$; or
- an empty set, $C_{\alpha}^{\text{AR}}(Y, X, Z) = \emptyset$.

The possibility of obtaining an infinite confidence set is a necessary condition for having a procedure robust to weak instruments. If instruments are weak, then the data contain little information about the coefficient of interest, resulting in infinite confidence sets. The AR test's ability to produce an empty confidence set is more confusing. It says that no value of the parameter is compatible with the data or that the model itself is rejected. An empty confidence set can happen even when the data were generated from the model (false rejection of the model).

By definition, the score confidence set is

$$C_{\alpha}^{\text{LM}}(Y, X, Z) = \{\beta_0: \text{LM}(\beta_0) < \chi_{1-\alpha, 1}^2\}$$

Finding the score region is equivalent to solving an inequality of the fourth power, which always has a solution in radicals because of Cardano's formula. However, there is a way

to rewrite the LM statistic in a way that requires solving two quadratic inequalities instead.

Let M and N denote the maximal and minimal eigenvalues of the matrix $\mathbf{Q}(\beta_0)$, respectively. Mikusheva (2005) showed that both M and N do not depend on β_0 and that the LM statistic has the following form:

$$\text{LM}(\beta_0) = -\frac{\{M - Q_T(\beta_0)\} \{N - Q_T(\beta_0)\}}{Q_T(\beta_0)}$$

Then the score confidence region is the set

$$C_\alpha^{\text{LM}}(Y, X, Z) = \left[\beta_0 : -\frac{\{M - Q_T(\beta_0)\} \{N - Q_T(\beta_0)\}}{Q_T(\beta_0)} < \chi_{1-\alpha,1}^2 \right]$$

The confidence set can be found in two steps. In the first step, we solve for the values of $Q_T(\beta_0)$ satisfying the inequality above. We have an ordinary quadratic inequality with respect to Q_T . In the second step, we find the score confidence set for β_0 by solving inequalities of the form $\{\beta_0 : Q_T(\beta_0) < q_1\} \cup \{\beta_0 : Q_T(\beta_0) > q_2\}$. As a result of this procedure, the score confidence region $C_\alpha^{\text{LM}}(Y, X, Z)$ for more than one instrument can have three possible forms:

- a union of two finite intervals, $C_\alpha^{\text{LM}}(Y, X, Z) = [x_1, x_2] \cup [x_3, x_4]$;
- a union of two infinite intervals and one finite interval, $C_\alpha^{\text{LM}}(Y, X, Z) = (-\infty, x_1] \cup [x_2, x_3] \cup [x_4, +\infty)$; or
- the whole line, $C_\alpha^{\text{LM}}(Y, X, Z) = (-\infty, +\infty)$.

The confidence set is never empty. It always contains the limited information maximum likelihood (LIML) estimator. The score confidence set always contains the points that minimize the p -value of the AR test and the conditional p -value of the conditional likelihood-ratio test. The distribution of the length of the score confidence set first-order stochastically dominates the distribution of the length of the conditional likelihood confidence set. That is, the score test tends to produce longer confidence sets than the conditional likelihood-ratio test. Because of these last two features, we do not recommend using the score confidence set in practice.

The main difficulty with finding an analytically tractable way of inverting the conditional likelihood-ratio test is that both the test statistic $\text{LR}(\beta_0)$ and the critical value function $m_\alpha\{Q_t(\beta_0)\}$ depend not only on data but also on the null value of the parameter β_0 . Mikusheva (2005) proved that the conditional likelihood-ratio confidence set is equal to the set

$$C_\alpha^{\text{CLR}}(Y, X, Z) = \{\beta_0 : Q_T(\beta_0) > C\}$$

where C is a solution to the equation $p(M - C; C) = \alpha$, where again M is the maximal eigenvalue of the matrix $\mathbf{Q}(\beta_0)$ and the function p was defined in (3). Thus the conditional likelihood-ratio confidence set can be found as a solution to a quadratic

inequality. As a result, the conditional likelihood-ratio confidence region $C_\alpha^{\text{CLR}}(Y, X, Z)$ can have three possible forms:

- a finite interval, $C_\alpha^{\text{CLR}}(Y, X, Z) = [x_1, x_2]$;
- a union of two infinite intervals, $C_\alpha^{\text{CLR}}(Y, X, Z) = (-\infty, x_1] \cup [x_2, +\infty)$; or
- the whole line, $C_\alpha^{\text{CLR}}(Y, X, Z) = (-\infty, +\infty)$.

The conditional likelihood-ratio confidence set is never empty; it always contains the LIML estimator.

4 Stata implementation

We have enhanced the `condivreg` command introduced by Moreira and Poi (2003) to reflect the advances made in the literature since it was introduced. `condivreg` users should upgrade to the new version. Among the changes are the following:

1. The results of the tests are presented by reporting (conditional) p -values rather than test statistics and their corresponding critical values. The conditional p -value for the conditional likelihood-ratio test is calculated by numerical integration as proposed by Andrews, Moreira, and Stock (2006) rather than by simulation.
2. The option to conduct tests by using the conditional Wald procedure was removed because of its extremely poor power properties.
3. The new version of `condivreg` contains an option to perform tests of the parameter on the endogenous regressor. Thus the `condtest` command of Moreira and Poi (2003) is deprecated.
4. We implemented algorithms for producing the conditional likelihood-ratio, score, and AR confidence sets within `condivreg`. Thus the `condgraph` command of Moreira and Poi (2003) is deprecated.
5. Since the conditional likelihood-ratio test possesses better power properties than the AR and the score tests for many parameters, `condivreg` always reports the conditional likelihood-ratio confidence set and p -value. The results for the AR and score tests are available by specifying the corresponding option.
6. The LIML estimate of the parameter on the endogenous variable is reported along with the conditional likelihood-ratio results, even when the main results are obtained via two-stage least squares (2SLS).

4.1 Syntax

```
condivreg depvar [indepvars] (endogvar = varlistiv) [if] [in] [,
  [2sls | liml] noconstant noinstconstant ar lm interval level(#)
  test(#)]
```

by, *rolling*, *statsby*, and *xi* are allowed; see [U] **11.1.10 Prefix commands**.

4.2 Options

2sls requests that the 2SLS estimator be used; this option is the default.

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

noconstant indicates that no constant term is to be included in the regression equation.

The default is to include a constant term.

noinstconstant indicates that no constant term is to be included in the first-stage regression of the endogenous variable on the instruments and 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 **noinstconstant** unless **noconstant** is also specified, but we give the user the option to experiment. By default, a constant term is included.

ar provides the coverage-corrected confidence set and size-corrected *p*-value based on the AR test statistic.

lm provides the coverage-corrected confidence set and size-corrected *p*-value based on the LM (score) test statistic.

interval displays the confidence interval, which is the minimal convex interval containing the coverage-corrected confidence set.

level(#) specifies the confidence level, as a percentage, for confidence intervals. The default is **level(95)** or as set by **set level**; see [U] **23.5 Specifying the width of confidence intervals**.

test(#) contains the hypothesized value of the endogenous variable's coefficient. The default is **test(0)**.

4.3 Remarks

condivreg fits a linear regression of *depvar* on *indepvars* and *endogvar* using *varlist_{iv}* (along with *indepvars*) as instruments for *endogvar* via the 2SLS or LIML estimator. The command reports the usual output of the IV regression in the same form as *ivreg*. In particular, it reports the conventional *t* statistic, *p*-values, and conventional Wald-type interval. The *p*-value and confidence set for the parameter on the endogenous regressor could be incorrect if instruments are weak. Also, *condivreg* reports the conditional

likelihood-ratio confidence region and p -value, both of which are robust to potentially weak instruments.

4.4 Example

For illustrative purposes, we use the same dataset and regression specification as that in [R] `ivreg` and in Moreira and Poi (2003).

```
. use http://www.stata-press.com/data/r9/hsng2
(1980 Census housing data)
. condivreg rent pcturban (hsngval = faminc reg2-reg4), ar lm
Instrumental variables (2SLS) regression
```

First-stage results	Number of obs =	50
F(4, 44) = 13.30	F(2, 47) =	42.66
Prob > F = 0.0000	Prob > F =	0.0000
R-squared = 0.6908	R-squared =	0.5989
Adj R-squared = 0.6557	Adj R-squared =	0.5818
	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
Confidence set and p-value for hsngval are based on normal approximation
```

Coverage-corrected confidence sets and p-values
for Ho: `_b[hsngval] = 0`
LIML estimate of `_b[hsngval]` = .0026686

Test	Confidence Set	p-value
Conditional LR	[.002018, .0037495]	0.0000
Anderson-Rubin	empty	0.0000
Score (LM)	[-.0007683, -.0004471] U [.0019973, .003808]	0.0000

The first half of the output looks similar to the output of command `ivreg`, except that `condivreg` also reports the first-stage regression's F statistic and R^2 . The inferential statistics in the coefficient table are based on the typical normal-approximation procedures. Here the instruments are strong and the approximation is accurate. However, for weak instruments these statistics can cause misleading inference.

The command also provides statistics that are valid whether the instruments are weak or strong. The LIML estimator, the conditional likelihood-ratio test for significance, and the conditional likelihood-ratio confidence set are always reported by default. The AR and the score tests and confidence sets are reported if options `ar` and `lm` are included.

The conditional likelihood-ratio confidence set is not much different from the one based on the normal approximation, though it is shifted toward the LIML estimator relative to the conventional Wald interval. The score confidence set consists of two finite intervals, which is the only possible form of the bounded score confidence set when the number of instruments is greater than 1. Both the conditional likelihood-ratio and score confidence sets contain the LIML estimator.

Here the AR confidence set is empty; that is, no value of the parameter is compatible with the model. The AR test can produce empty confidence sets (i.e., it rejects the model) even if the model is correct.

By default, the command reports p -values for the test of $H_0: \beta = 0$. However, the `test()` option can be used to conduct tests of $H_0: \beta = \beta_0$ for other values of β_0 . For example, here we test $H_0: \beta = 0.003$:

```
. condivreg rent pcturban (hsngval = faminc reg2-reg4), ar lm test(0.003)
(output omitted)
```

5 Saved results

`condivreg` saves the following in `e()`:

Scalars

<code>e(N)</code>	number of observations	<code>e(df_r_first)</code>	first-stage residual degrees of freedom
<code>e(df_m)</code>	model degrees of freedom	<code>e(r2_first)</code>	first-stage R -squared
<code>e(df_r)</code>	residual degrees of freedom	<code>e(r2_a_first)</code>	first-stage adjusted R -squared
<code>e(F)</code>	model F statistic	<code>e(b_instd_liml)</code>	LIML estimate of β
<code>e(H0_b)</code>	value of β under null	<code>e(level)</code>	confidence level
<code>e(r2)</code>	R -squared	<code>e(p_LR)</code>	conditional LR test p -value
<code>e(r2_a)</code>	adjusted R -squared	<code>e(p_AR)</code>	AR test p -value
<code>e(rmse)</code>	root mean squared error	<code>e(p_LM)</code>	LM test p -value
<code>e(mss)</code>	model sum of squares	<code>e(LR_xi)</code>	see below
<code>e(rss)</code>	residual sum of squares	<code>e(AR_xi)</code>	see below
<code>e(F_first)</code>	first-stage F statistic	<code>e(LM_xi)</code>	see below
<code>e(df_m_first)</code>	first-stage model degrees of freedom		

Macros

<code>e(cmd)</code>	<code>condivreg</code>	<code>e(model)</code>	2SLS or LIML
<code>e(LR_type)</code>	see below	<code>e(instd)</code>	instrumented variable
<code>e(AR_type)</code>	see below	<code>e(insts)</code>	instruments
<code>e(LM_type)</code>	see below	<code>e(exog)</code>	included exogenous variables
<code>e(cons)</code>	<code>noconstant</code> or not set	<code>e(depvar)</code>	dependent variable
<code>e(instcons)</code>	flag constant among instruments	<code>e(properties)</code>	<code>b V</code>

Matrices

<code>e(b)</code>	coefficient vector	<code>e(V)</code>	variance-covariance matrix
-------------------	--------------------	-------------------	----------------------------

Functions

<code>e(sample)</code>	marks estimation sample
------------------------	-------------------------

The macros `e(LR_type)`, `e(AR_type)`, and `e(LM_type)` indicate the form of the conditional likelihood-ratio, AR, and score (LM) confidence regions, respectively:

<i>type</i>	Confidence region
1	\emptyset
2	$[x_1, x_2]$
3	$(-\infty, \infty)$
4	$(-\infty, x_1] \cup [x_2, \infty)$
5	$(-\infty, x_1] \cup [x_2, x_3] \cup [x_4, \infty)$
6	$[x_1, x_2] \cup [x_3, x_4]$

For the conditional likelihood-ratio confidence region, only types 2–4 are possible. For the AR confidence region, types 1–4 are possible. For the score (LM) confidence region, all but type 1 are possible. The finite endpoints x_1 , x_2 , x_3 , and x_4 are saved in the scalars $\mathbf{e}(test_x1)$, $\mathbf{e}(test_x2)$, $\mathbf{e}(test_x3)$, and $\mathbf{e}(test_x4)$, respectively, where *test* is LR, AR, or LM.

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.
- Andrews, D. W. K., M. Moreira, and J. Stock. 2004. Optimal two-sided invariant similar tests for instrumental variables regression. Unpublished manuscript.
- . 2006. Performance of conditional Wald tests in IV regression with weak instruments. *Journal of Econometrics*. Forthcoming.
- Andrews, D. W. K., and J. Stock. 2005. Inference with weak instruments. Discussion Paper No. 1530. New Haven, CT: Cowles Foundation.
- 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.
- Kleibergen, F. 2002. Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica* 70: 1781–1803.
- Mikusheva, A. 2005. Robust confidence sets in presence of weak instruments. Unpublished manuscript.
- Moreira, M. 2001. Tests with correct size when instruments can be arbitrarily weak. Center for Labor Economics Working Paper 37, UC Berkeley.

———. 2003. A conditional likelihood ratio test for structural models. *Econometrica* 71: 1027–1048.

Moreira, M. J., and B. P. Poi. 2003. Implementing tests with correct size in the simultaneous equations model. *Stata Journal* 3: 57–70.

Nelson, C., and R. Startz. 1990. Some further results on the exact small sample properties of the instrumental variable estimator. *Econometrica* 58: 967–976.

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.

About the authors

Anna Mikusheva received her PhD in mathematics from Moscow State University and is currently a PhD student in the economics department at Harvard University.

Brian Poi is a senior economist at StataCorp.