The Stata Journal (2001) **1**, Number 1, pp. 1–11

Tests and confidence sets with correct size in the simultaneous equations model with potentially weak instruments

Anna Mikusheva
Department of Economics
Harvard University

Brian P. Poi
StataCorp LP

Abstract.

We consider inference in the linear regression model with a single endogenous variable and potentially weak instruments. We construct confidence sets for the coefficient on 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 algorithms for finding these confidence sets, and we present a Stata command that supersedes the one presented in Moreira and Poi (2003).

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

1 Introduction

We consider inference on the parameter of a single 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 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 instrumented variable has incorrect size, and the Wald-type confidence interval has low coverage probability.

A wide literature is devoted to finding tests about the coefficient β on the single included endogenous regressor that are valid in the presence of potentially weak instruments. Andrews and Stock (2005) and Stock et al. (2002) give excellent surveys of the literature. The class of tests robust to weak identification includes the Anderson and Rubin (1949) test, the Lagrange multiplier (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 also robust to weak instruments by inverting the test. That is, a value β_0

 st0001

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 Anderson-Rubin, score, conditional likelihood ratio, and conditional Wald tests. They also provided the command `condgraph` that 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 data sets could take several hours and is not very accurate. Both of those 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. The simulations are computationally intensive and not always accurate.

The second obstacle is the fact 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 towards weak instruments must be infinite with positive probability, making a grid search non-feasible 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 et al. (2005) 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 its power curve can be non-monotonic. Andrews et al. (2005) recommended not using Wald test in practice. Mikusheva (2005) proposed algorithms that allow one to construct confidence sets by inverting the Anderson-Rubin, score, and conditional likelihood ratio tests in a fast and accurate way without having to use a grid search.

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

The paper is organized as follows. Section 2 contains a brief overview of the model and definitions of the Anderson-Rubin, the score and the conditional likelihood ratio tests. Section 3 provides algorithms for inverting these tests in order to construct weak instrument robust confidence sets. In section 4 we describe the syntax of the Stata command `condivreg` and provide an example of its usage.

2 Tests robust to weak instruments

In this section we introduce notations and give a brief overview of the tests robust towards weak instruments. The model contains a structural equation and a reduced form equation for a single endogenous regressor:

$$\mathbf{y}_1 = \mathbf{y}_2\beta + \mathbf{X}\gamma_1 + \mathbf{u} \tag{1}$$
$$\mathbf{y}_2 = \mathbf{Z}\pi + \mathbf{X}\xi + \mathbf{v}_2 \tag{2}$$

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; $\beta \in R$, $\gamma_1, \xi \in R^p$ and $\pi \in R^k$ are unknown parameters. We assume without loss of generality that $\mathbf{Z}'\mathbf{X} = \mathbf{0}$. The $n \times 2$ matrix of errors $[\mathbf{u} : \mathbf{v}_2]$ is i.i.d across rows, each row being normally distributed with mean zero and non-singular covariance matrix.

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

$$\begin{aligned}\mathbf{y}_1 &= \mathbf{Z}\pi\beta + \mathbf{X}\gamma + \mathbf{v}_1\\ \mathbf{y}_2 &= \mathbf{Z}\pi + \mathbf{X}\xi + \mathbf{v}_2\end{aligned}$$

where

$$\gamma = \gamma_1 + \xi\beta \quad \text{and} \quad \mathbf{v}_1 = \mathbf{u} + \beta\mathbf{v}_2$$

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 et al. (2005) showed that in the case of unknown $\boldsymbol{\Omega}$ asymptotically valid tests can be received by replacing $\boldsymbol{\Omega}$ with a consistent estimator of $\boldsymbol{\Omega}$. Andrews et al. (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 we refer readers to Andrews et al. (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 introduce the following statistics, properties of which are discussed in Moreira (2003):

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

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 $\mathbf{Y} = [\mathbf{y}_1 : \mathbf{y}_2]$.

We also consider the matrix $\mathbf{Q}$ defined as

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

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 notational simplicity, in the remainder of the article we shall simply refer to $\mathbf{S}$ and $\mathbf{T}$, with their dependence on β_0 implied.

The Anderson-Rubin test rejects the null hypthesis $H_0 : \beta = \beta_0$ at significance level α if the statistic

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

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

The Lagrange multiplier (score) test accepts the null if the statistic

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

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 a single fixed critical value, critical values that are functions of the data. The conditional likelihood ratio test uses the statistic

$$\begin{aligned} LR(\beta_0) \quad = \quad & \frac{1}{2}\Big[Q_S(\beta_0) - Q_T(\beta_0) + \\ & \qquad \left[\{Q_S(\beta_0) + Q_T(\beta_0)\}^2 - 4\left\{Q_S(\beta_0)Q_T(\beta_0) - Q_{ST}^2(\beta_0)\right\}\right]^{1/2}\Big] \end{aligned}$$

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:

$$P\left\{LR > m_\alpha(q_T) \,|\, Q_T = q_T\right\} = \alpha.$$

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

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 a large number of simulations. Andrews et al. (2005) 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:

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

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

$$p(LR(\beta_0); Q_T(\beta_0)) > \alpha.$$

Andrews et al. (2005) 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 takes almost no time.

The three tests described above have a correct size in a case of weak instruments. However, they possess different power properties. The Anderson-Rubin test is robust to misspecifications of equation (2) and can be used as an over-identification 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, it is included in the package. According to Andrews et al. (2005), the conditional likelihood ratio test is nearly optimal in a class of invariant similar tests. It possesses better power properties than the Anderson-Rubin and the score tests for a wide range of parameters.

3 Confidence sets based on tests robust towards weak instruments.

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

Given the tests robust towards 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 a 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 application 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. It leads to the necessity to find an algorithm of inverting tests without employing a grid testing.

By definition, the Anderson-Rubin confidence set is the set

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

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

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

We want to emphasize that 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 contains very little information about the coefficient of interest, and that results in infinite confidence sets. The ability of the Anderson-Rubin test 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. It can happen even when the data was in fact generated from the model (false rejection of the model).

By definition, the score confidence set is the set

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

Finding the score region is equivalent to solving an inequality of the fourth power, which always has a solution in radicals due to 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 statistics has the following form:

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

Then the score confidence region is the set

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

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 $C_\alpha^{LM}(Y, X, Z)$ in the case of more than one instrument can have one of three possible forms:

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

Several features of the score confidence set should be emphasized. First, the confidence set is never empty. It always contains the limited information maximum likelihood (LIML) estimator. Second, the score confidence set always contains the points that minimize the p-value of the Anderson-Rubin test and the conditional p-value of the conditional likelihood ratio test. Finally, 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 $LR(\beta_0)$ and the critical value function $m_\alpha(Q_t(\beta_0))$ depend not only on data, but on the null value of the parameter β_0. Mikusheva (2005) proved that the conditional likelihood ratio confidence set is equal to the set

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

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 equation (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 $C_\alpha^{CLR}(Y, X, Z)$ can have one of three possible forms:

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

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. We strongly encourage existing users of `condivreg` to 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 et al. (2005) rather than by simulation.
2. The option to conduct tests 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 Anderson-Rubin, score, and conditional likelihood ratio 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 Anderson-Rubin and the score tests for a wide range of parameters, `condivreg` always reports the conditional likelihood ratio confidence set and p-value. The results for the Anderson-Rubin 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 2SLS.

4.1 Syntax

`condivreg` *depvar* [*varlist*] `(`*endovar* `=` *varlist*$_{iv}$`)` [*if*] [*in*] [`, {2sls | liml}`
`nocons noinstcons ar lm interval level(`#`) test(`#`)`]

by:, rolling:, statsby:, and xi: may be used with condivreg.

4.2 Options

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

liml requests that the LIML estimator 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 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.

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

lm provides the coverage-corrected confidence set and size-corrected p-value based on the Lagrange multiplier (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, in percent, 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 *varlist* and *endogvar* using $varlist_{iv}$ (along with *varlist*) 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-statistics, p-values, and conventional Wald-type interval. We emphasize that the p-value and confidence set for the parameter on the endogenous regressor could be incorrect if instruments are weak. Additionally, 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 data set and regression specification as in [R] **ivreg** and in Moreira and Poi (2003).

```
. use http://www.stata-press.com/data/r7/hsng2.dta,clear
(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(  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
Confidence set and p-value for instrumented variable are based on normal apprx
------------------------------------------------------------------------------

Coverage-corrected confidence set for the coefficient on hsngval
and conditional p-value for Ho:  b[hsngval] = 0.0000   based on CLR test:
------------------------------------------------------------------------------
 LIML                      Confidence Region                  p-value
------------------------------------------------------------------------------
 .00266862               [  0.0020 ,   0.0037]                 0.0000
------------------------------------------------------------------------------

------------------------------------------------------------------------------
Additional test(s)         Confidence Region                  p-value
------------------------------------------------------------------------------
AR                               empty                         0.0000
Score(LM)          [-0.0008,-0.0004] U [ 0.0020,0.0038080]       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. In this example the instruments are strong, and the approximation is quite accurate. However, in the case of weak instruments these statistics can lead to 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 Anderson-Rubin and the score tests and the confidence sets are reported if options `ar` and `lm` are included.

We want to point out that in this example 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.

In this example the Anderson-Rubin confidence set is empty; that is, no value of the parameter is compatible with the model. We already pointed out that the Anderson-Rubin test can produce empty confidence sets (i.e. it rejects the model) even if the model is correct.

The command also allows the user to perform testing of the hypothesis $H_0 : \beta = \beta_0$ using the conditional likelihood ratio, Anderson-Rubin, and score tests.

```
. condivreg rent pcturban (hsngval = faminc reg2-reg4), ar lm test(0.00266862)
> interval

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
Confidence set and p-value for instrumented variable are based on normal apprx
------------------------------------------------------------------------------

Coverage-corrected confidence interval for the coefficient on hsngval
and conditional p-value for Ho:  b[hsngval] = 0.0027   based on CLR test:
------------------------------------------------------------------------------
 LIML                      Confidence Region                   p-value
------------------------------------------------------------------------------
 .00266862             [  0.0020 ,    0.0037]                   0.9917
------------------------------------------------------------------------------

------------------------------------------------------------------------------
Additional test(s)         Confidence Region                   p-value
------------------------------------------------------------------------------
AR                               empty                          0.0260
Score(LM)                 [-0.0008,0.0038080]                   1.0000
------------------------------------------------------------------------------
```

In this example we tested the hypothesis that the parameter of interest is equal to the LIML estimator. The LIML estimator maximizes the p-values for all three tests. The p-value for the score test and conditional p-value for the conditional likelihood ratio at the LIML estimator are approximately equal to one. The p-value of the Anderson-

Rubin test at the LIML estimator is below 5%; it is equivalent to having an empty confidence set.

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.

Andrews, D. W. K., M. Moreira, and J. Stock. 2004. Optimal Two-Sided Invariant Similar Tests for Instrumental Variables Regression. *unpublished manuscript* .

—. 2005. 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. *Journal of Econometrics* Cowles Foundation Discussion Paper No.1530.

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 Error-in-variables and Related Models. *Annals of Statistics* 15: 1351–62.

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 Arbitrary Weak*. UC Berkeley: Unpublished manuscrip.

—. 2003. A Conditional Likelihood Ratio Test for Structural Models. *Econometrica* 71(4): 1027–1048.

Moreira, M. and B. Poi. 2003. Implementing Tests with Correct Size in the Simultaneous Equations Model. *Stata Journal* 3(1): 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., J. 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 Author

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 Senior Economist at StataCorp LP.