The Stata Journal (2015)
15, Number 1, pp. 216–225

A robust test for weak instruments in Stata

Carolin E. Pflueger
University of British Columbia
Vancouver, Canada
carolin.pflueger@sauder.ubc.ca

Su Wang
London School of Economics
London, UK
s.wang50@lse.ac.uk

Abstract. We introduce a routine, `weakivtest`, that implements the test for weak instruments by Montiel Olea and Pflueger (2013, *Journal of Business and Economic Statistics* 31: 358–369). `weakivtest` allows for errors that are not conditionally homoskedastic and serially uncorrelated. It extends the Stock and Yogo (2005, Testing for weak instruments in linear IV regression. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. D. W. K. Andrews and J. J. Stock, 80–108. [Cambridge University Press]) weak-instrument tests available in `ivreg2` and in the `ivregress` postestimation command `estat firststage`. `weakivtest` tests the null hypothesis that instruments are weak or that the estimator's Nagar (1959, *Econometrica* 27: 575–595) bias is large relative to a benchmark for both two-stage least-squares estimation and limited-information maximum likelihood with one endogenous regressor. The routine can accommodate Eicker–Huber–White heteroskedasticity robust estimates, Newey and West (1987, *Econometrica* 55: 703–708) heteroskedasticity- and autocorrelation-consistent estimates, and clustered variance estimates.

Keywords: st0377, weakivtest, F statistic, heteroskedasticity, autocorrelation, clustered, weak instruments, testing

1 Introduction

In this article, we describe the weak-instrument test by Montiel Olea and Pflueger (2013) and introduce a new routine, `weakivtest`, for implementing this test. `weakivtest` is a postestimation routine for `ivreg2` and `ivregress`.[1]

Weak instruments can bias point estimates and cause substantial test size distortions (Nelson and Startz 1990; Stock and Yogo 2005). Departures from the homoskedastic serially uncorrelated framework are extremely common in practice and can also further bias estimates and distort test sizes when instruments are weak (Montiel Olea and Pflueger 2013). We provide a user-friendly routine for heteroskedasticity, autocorrelation, and cluster–robust weak-instrument tests. These tests apply to two-stage least-squares (TSLS) estimation and limited-information maximum likelihood (LIML) with one endogenous regressor.

Under strong instruments, TSLS and LIML are asymptotically unbiased. However, under weak instruments, this is not the case. For more on inference with potentially weak instruments, see Stock, Wright, and Yogo (2002) and Andrews and Stock (2006).

1. While this article provides a pretest for weak instruments, methods for weak-instrument robust inference are also available and are implemented in the command `weakiv` (Finlay, Magnusson, and Schaffer 2014).

Staiger and Stock (1997) and Stock and Yogo (2005) proposed widely used pretests for weak instruments under the assumption of conditionally homoskedastic, serially uncorrelated model errors. These tests reject the null hypothesis of weak instruments when the Cragg and Donald (1993) statistic exceeds a given threshold. This test statistic reduces to the first-stage F statistic with one endogenous regressor. The null hypothesis of weak instruments can be defined in terms of either estimator bias or test size distortions.

The `ivreg2` suite, described in Baum, Schaffer, and Stillman (2007, 2010), implements the Stock and Yogo (2005) weak-instrument test for conditionally homoskedastic, serially uncorrelated model errors.

Practitioners frequently report the robust or nonrobust first-stage F statistic as an ad hoc way of adjusting the Stock and Yogo (2005) tests for heteroskedasticity, autocorrelation, and clustering. However, Montiel Olea and Pflueger (2013) show that both the robust and the nonrobust F statistics may be high even when instruments are weak. Baum, Schaffer, and Stillman (2007) also emphasize that the Kleibergen and Paap (2006) rank using the Wald statistic does not provide a formal test for weak instruments with heteroskedastic, serially correlated, or clustered model errors.

`weakivtest` tests the null hypothesis that the estimator's approximate asymptotic bias (or Nagar [1959] bias) exceeds a fraction τ of a "worst-case" benchmark (BM). This BM agrees with the ordinary least-squares (OLS) bias when errors are conditionally homoskedastic and serially uncorrelated. The test rejects the null hypothesis when the test statistic, the effective F statistic, exceeds a critical value. The critical value depends on the significance level, α, and the desired threshold, τ.

When data are conditionally homoskedastic and serially uncorrelated, the effective F statistic is identical to the Cragg and Donald (1993) statistic recommended by Stock and Yogo (2005). We can compare `weakivtest` critical values for the null hypothesis that the TSLS approximate asymptotic bias (henceforth, the Nagar bias) exceeds 10% of the BM with Stock and Yogo (2005) critical values for the null hypothesis that the TSLS bias exceeds 10% of the OLS bias. With conditionally homoskedastic and serially uncorrelated errors, `weakivtest` critical values with significance level 5% increase from 8.53 for 3 instruments to 12.27 for 30 instruments. By comparison, the corresponding Stock and Yogo (2005) critical values increase from 9.08 for 3 instruments to 11.32 for 30 instruments.

2 Linear IV with potentially weak instruments

We consider the following standard linear instrumental-variables (IV) setup with one endogenous regressor and K instruments. We write the linear IV model in reduced form, as follows:

$$\mathbf{y} = \mathbf{Z\Pi}\beta + \mathbf{X}\boldsymbol{\gamma}_1 + \mathbf{v}_1 \qquad (1)$$

$$\mathbf{Y} = \mathbf{Z\Pi} + \mathbf{X}\boldsymbol{\gamma}_2 + \mathbf{v}_2 \qquad (2)$$

Equation (1) denotes the reduced-form second-stage relationship. Equation (2) denotes the reduced-form first-stage relationship between the instruments and the endogenous regressor. We want to estimate the structural parameter β. $\mathbf{\Pi} \in \mathbb{R}^K$ denotes the vector of unknown first-stage parameters, and $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ denote the vector of coefficients on the included exogenous regressors.

We want to observe the outcome variable $\mathbf{y}_s$, the endogenous regressor $\mathbf{Y}_s$, the vector of K instruments $\mathbf{Z}_s$, and the vector of L included exogenous regressors $\mathbf{X}_s$ for $s = 1, \dots, S$. The unobserved reduced-form errors have realizations $\mathbf{v}_{js}, j \in 1, 2$. We stack the realized variables in matrices $\mathbf{y} \in \mathbb{R}^S$, $\mathbf{Z} \in \mathbb{R}^{s\times K}$, and $\mathbf{v}_j \in \mathbb{R}^S$, $j \in (1, 2)$.

TSLS and LIML estimators depend on realized variables only through their projection residuals with respect to $\mathbf{X}$. Saving notation, we let $\mathbf{y}$, $\mathbf{Y}$, and $\mathbf{v}_j$, $j = 1, 2$ denote their projection errors onto $\mathbf{X}$. For instance, we replace the endogenous regressor $\mathbf{Y}$ by $\mathbf{M_X Y}$, where $\mathbf{M_X} = \mathbb{I}_S - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$. We also normalize the vector of instruments $\mathbf{Z}$ such that $\mathbf{Z}'\mathbf{Z}/S = \mathbb{I}_S$, which again leaves TSLS and LIML estimators unchanged.

We denote the projection matrix onto $\mathbf{Z}$ by $\mathbf{P_Z} = \mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'$ and the complementary matrix by $\mathbf{M_Z} = \mathbb{I}_S - \mathbf{P_Z}$. The TSLS estimator of β is

$$\widehat{\beta}_{\text{TSLS}} \equiv (\mathbf{Y}'\mathbf{P_Z}\mathbf{Y})^{-1}(\mathbf{Y}'\mathbf{P_Z}\mathbf{y})$$

The LIML estimator of β is

$$\widehat{\beta}_{\text{LIML}} = \{\mathbf{Y}'(\mathbb{I}_S - k_{\text{LIML}}\mathbf{M_Z})\mathbf{Y}\}^{-1}\{\mathbf{Y}'(\mathbb{I}_S - k_{\text{LIML}}\mathbf{M_Z})\mathbf{y}\}$$

where k_{LIML} is the smallest root of the determinantal equation

$$\Big| (\mathbf{y}, \mathbf{Y})'(\mathbf{y}, \mathbf{Y}) - k(\mathbf{y}, \mathbf{Y})'\mathbf{M_Z}(\mathbf{y}, \mathbf{Y}) \Big| = 0$$

The robust weak-instrument pretest relies on two assumptions. First, we model weak instruments by assuming that the IV first-stage relation is local to zero, following the modeling strategy in Staiger and Stock (1997). Intuitively, the vector of first-stage coefficients is small in magnitude relative to the sample size.

Assumption $\mathbf{L_\Pi}$. (**Local to Zero**) $\mathbf{\Pi} = \mathbf{\Pi}_S = \mathbf{C}/\sqrt{S}$, where $\mathbf{C}$ is a fixed vector $\mathbf{C} \in \mathbb{R}^K$.

Second, we make high-level assumptions about the variances and covariances of the reduced-form residuals and the residuals interacted with the vector of instruments.

Assumption high level. The following limits hold as $S \to \infty$:

1. $\begin{pmatrix} \mathbf{Z}'\mathbf{v}_1/\sqrt{S} \\ \mathbf{Z}'\mathbf{v}_2/\sqrt{S} \end{pmatrix} \xrightarrow{d} \mathcal{N}_{2K}(\mathbf{0}, \mathbf{W})$ for positive-definite $\mathbf{W} = \begin{pmatrix} \mathbf{W}_1 & \mathbf{W}_{12} \\ \mathbf{W}'_{12} & \mathbf{W}_2 \end{pmatrix}$

2. $(\mathbf{v}_1, \mathbf{v}_2)'(\mathbf{v}_1, \mathbf{v}_2)/S \xrightarrow{p} \mathbf{\Omega}$ for positive-definite $\mathbf{\Omega} \equiv \begin{pmatrix} \omega_1^2 & \omega_{12} \\ \omega_{12} & \omega_2^2 \end{pmatrix}$

3. There exists a sequence of positive-definite estimates $\{\widehat{\mathbf{W}}(S)\}$, measurable with respect to $(y_s, Y_s, \mathbf{Z}_s)_{s=1}^S$, such that $\widehat{\mathbf{W}}(S) \xrightarrow{p} \mathbf{W}$ as $S \to \infty$.

2.1 Testing procedure

`weakivtest` tests the null hypothesis that instruments are weak. When the null hypothesis is rejected, we can conclude that instruments are strong and proceed using standard inference.

Montiel Olea and Pflueger (2013) use the standard Nagar (1959) methodology to obtain a tractable proxy for the asymptotic estimator bias. They define the Nagar bias as the expectation of the first three terms in the Taylor series expansion of the asymptotic estimator distribution under weak-instrument asymptotics. The Nagar Bias is always defined and bounded for both TSLS and LIML. Montiel Olea and Pflueger (2013) define the null hypothesis of weak instruments such that the Nagar bias may be large. Under the alternative hypothesis, the Nagar bias is bounded relative to the BM.

Montiel Olea and Pflueger (2013) use the BM to test the TSLS Nagar bias, N_{TSLS}, and the LIML Nagar bias, N_{LIML}, against the following function: $\text{BM}(\beta, \mathbf{W}) \equiv \sqrt{tr(\mathbf{W}_1 - 2\beta\mathbf{W}_{12} + \beta^2\mathbf{W}_2)/tr(\mathbf{W}_2)}$. Intuitively, the BM captures the "worst-case" situation when instruments are completely uninformative and when first-stage and second-stage errors are perfectly correlated. It is also a natural extension of using the BM against the OLS bias when reduced-form errors are conditionally homoskedastic and serially uncorrelated, as in the tests proposed by Stock and Yogo (2005).

Under the weak-instrument null hypothesis, the Nagar bias exceeds a fraction τ of the BM for some value of the structural parameter β and some direction of the first-stage coefficients $\mathbf{\Pi}$. Under the alternative, the Nagar bias is at most a fraction τ of the BM for any values for the structural parameter β and for any direction of the first-stage coefficients $\mathbf{\Pi}$.

The robust weak-instrument test rejects the null hypothesis of weak instruments when the test statistic, the effective F statistic $\widehat{F}_{\text{eff}}$,

$$\widehat{F}_{\text{eff}} \equiv \frac{\mathbf{Y}'\mathbf{P_Z}\mathbf{Y}}{\text{tr}(\widehat{\mathbf{W}}_2)}$$

exceeds a critical value. When there is one instrument, as shown above, the effective F statistic equals the robust F statistic, but it generally differs from both the nonrobust F,

$$\widehat{F} \equiv \frac{\mathbf{Y}'\mathbf{P_Z}\mathbf{Y}}{K\widehat{\omega}_2^2}$$

and the robust F statistic,

$$\widehat{F}_r \equiv \frac{\mathbf{Y}'\mathbf{Z}\widehat{\mathbf{W}}_2^{-1}\mathbf{Z}'\mathbf{Y}}{K \times S}$$

The critical value (c) depends on the significance level (α), the desired threshold (τ), the estimated variance–covariance matrix ($\widehat{W}$), and on the estimator (TSLS or LIML). Both a generalized and a simplified conservative critical value are available for TSLS.

3 Implementation

1. `weakivtest` uses Stata's built-in `regress` routine to estimate (1) and (2) using equation-by-equation OLS. `weakivtest` estimates the matrix $\widehat{\mathbf{W}}$ using the same level of robustness as the preceding `ivreg2` or `ivregress` command with Baum and Schaffer's (2013) program `avar`. The estimate $\widehat{\mathbf{W}}$ equals the robust estimated variance–covariance matrix times a degrees-of-freedom adjustment, $S/(S-K-L-1)$. `weakivtest` supports estimating $\widehat{\mathbf{W}}$ with Eicker–Huber–White heteroskedasticity robust, Newey and West (1987) heteroskedasticity-consistent and autocorrelation-consistent, or clustered variance–covariance matrix estimates. $\widehat{\boldsymbol{\Omega}}$ is obtained as the cross-product of $\widehat{\mathbf{v}}_1$ and $\widehat{\mathbf{v}}_2$.

2. `weakivtest` obtains the effective F statistic as a scaled version of the nonrobust first-stage F statistic, $\widehat{F}$, with $\widehat{F}_{\text{eff}} = \widehat{F} \times K\widehat{\omega}_2^2/tr(\widehat{\mathbf{W}_2})$, where $\widehat{\omega}_2$ is the consistent estimate of ω_2. Note the following:

   a. The asymptotic distribution of $\widehat{F}_{\text{eff}}$—denoted F^*_{eff}—is a weighted sum of noncentral χ^2 random variables (see Montiel Olea and Pflueger [2013, lemma 1, part 5, 362]). One of the challenges in the implementation of our testing procedure is approximating the quantiles of such distribution.

   b. Large values of the expectation of F^*_{eff} correspond to small values of the approximate asymptotic bias (or Nagar Bias) for both TSLS and LIML (see Montiel Olea and Pflueger [2013, theorem 1, 362]). This observation explains the selection of the test statistic.

3. `weakivtest` computes two quantities that are used to approximate the upper α point of F^*_{eff}: a noncentrality parameter x that bounds the mean of F^*_{eff} under the null hypothesis and the effective degrees of freedom $\widehat{K}_{\text{eff}}$. The rationale for these parameters is as follows. Patnaik (1949) and Imhof (1961) approximate the critical values of a weighted sum of independent noncentral χ^2 distributions by a central χ^2 with the same first and second moments. Building on this result, Montiel Olea and Pflueger (2013) approximate F^*_{eff} by the following noncentral χ^2:

$$\frac{1}{\widehat{K}_{\text{eff}}}\chi^2_{\widehat{K}_{\text{eff}}}\left(\widehat{K}_{\text{eff}}x\right)$$

   Here

$$\begin{aligned}
\widehat{K}_{\text{eff}} &\equiv \frac{\left\{\text{tr}\left(\widehat{\mathbf{W}}_2\right)\right\}^2(1+2x)}{\text{tr}\left(\widehat{\mathbf{W}}_2'\widehat{\mathbf{W}}_2'\right)+2x\text{tr}\left(\widehat{\mathbf{W}}_2\right)\text{max eval}\left(\widehat{\mathbf{W}}_2\right)} \\
x &= \frac{B_e\left(\widehat{\mathbf{W}},\widehat{\boldsymbol{\Omega}}\right)}{\tau} \quad \text{for } e \in (\text{TSLS}, \text{LIML})
\end{aligned}$$

and $B_e(\widehat{\mathbf{W}}, \widehat{\mathbf{\Omega}})$ is closely related to the supremum of the Nagar bias relative to the BM.

Computing x requires maximizing the ratio of the Nagar (1959) bias divided by the BM over all values of the structural parameter, β, and all directions for the first-stage coefficients, $\mathbf{\Pi}$. As shown in Montiel Olea and Pflueger (2013), this step reduces to a numerical maximization over the real line, and `weakivtest` implements it using the built-in function `optimize()`.

4. Given the noncentrality parameter, x, and the effective degrees of freedom, $\widehat{K}_{\text{eff}}$, the critical values can be calculated as the upper α point of $\chi^2_{\widehat{K}_{\text{eff}}}(x\widehat{K}_{\text{eff}})/\widehat{K}_{\text{eff}}$, following the curve-fitting methodology of Patnaik (1949).

5. `weakivtest` calls routine `invnchi2` if Stata version is 13.1 or higher and calls routine `invnFtail` for lower Stata versions to compute the inverse noncentral chi-squared cumulative distribution function. At the time of writing, the built-in routine `invnchi2` does not support noninteger degrees of freedom for Stata versions lower than 13.1. Setting the denominator degrees of freedom in `invnFtail` to a sufficiently large positive number approximates the noncentral chi-squared cumulative distribution function.

3.1 Syntax

`weakivtest` $\big[$ `, level(`#`) eps(`#`) n2(`#`)` $\big]$

3.2 Description

`weakivtest` is a postestimation command for `ivreg2` and `ivregress`.

`weakivtest` reports the effective F statistic. It reports generalized TSLS and LIML critical values for threshold values $\tau \in (5\%, 10\%, 20\%, 30\%)$.

Montiel Olea and Pflueger (2013) provide both the generalized value and the simplified conservative critical value for TSLS. The simplified critical value exploits an analytical conservative bound for the Nagar Bias of TSLS (see Montiel Olea and Pflueger [2013], theorem 1, part 3). The simplified procedure follows the same steps as the generalized procedure, but it sets the noncentrality parameter to $x = 1/\tau$. Hence, simplified critical values are computationally less demanding. For completeness, `weakivtest` saves both types of TSLS critical values. However, the TSLS generalized critical value provides a slightly more powerful test and should be used when available. Therefore, `weakivtest`, displays only the TSLS generalized critical value.

3.3 Options

`level(#)` specifies the confidence level. The default is `level(0.05)`.

`eps(#)` specifies the input parameter for the Nelder–Mead optimization technique. The default is `eps(10^(-3))`.

`n2(#)` specifies the denominator degrees of freedom for the inverse noncentral F distribution. Set `n2(#)` to a large positive number to approximate an inverse noncentral chi-squared distribution. The default is `n2(10^(7))`.

The weak-instrument test can adjust for a variety of violations of conditionally homoskedastic, independent, identically distributed model errors.

3.4 Options for ivreg2 and ivregress

`weakivtest` estimates the variance–covariance matrix of errors as specified in the preceding `ivreg2` or `ivregress` command. The following options are supported:

`robust` estimates an Eicker–Huber–White heteroskedasticity robust variance–covariance matrix.

`cluster(`*varname*`)` estimates a variance–covariance matrix clustered by the specified variable.

`robust bw(#)` (for `ivreg2`) estimates a heteroskedasticity- and autocorrelation-consistent variance–covariance matrix computed with a Bartlett (Newey–West) kernel with bandwidth #.

`bw(#)`, without the `robust` option (for `ivreg2`), requests estimates that are autocorrelation consistent but not heteroskedasticity consistent.

`vce(hac nw #)` (for `ivregress`) estimates a heteroskedasticity- and autocorrelation-consistent variance–covariance matrix computed with a Bartlett (Newey–West) kernel with number of lags #. The bandwidth of a kernel is equal to the number of lags plus one.

4 Example

Here we implement `weakivtest` as a postestimation command for `ivreg2` using the dataset of Campbell (2003) and Yogo (2004). The IV setup is identical to that in table 2A of Montiel Olea and Pflueger (2013). This baseline example uses a Bartlett (Newey–West) kernel with a bandwidth of 7 and a significance level of 5%, and it focuses on a weak-instrument threshold of $\tau = 10\%$.

By comparison, Montiel Olea and Pflueger (2013) report an effective F statistic of 7.94, a TSLS critical value of 15.49, and a LIML critical value of 9.68 for $\tau = 10\%$. `weakivtest` cannot reject the null of weak instruments for TSLS or for LIML for a weak-instrument threshold of $\tau = 10\%$, consistent with the findings in Montiel Olea and Pflueger (2013).

```
. use data_usaq
(Campbell (2003) and Yogo (2004) US quarterly data 1947.3-1998.4)
. quietly ivreg2 dc (rrf=z1 z2 z3 z4), robust bw(7)
. weakivtest
(obs=206)
Montiel-Pflueger robust weak instrument test
--------------------------------------------
Effective F statistic:        7.942
Confidence level alpha:          5%
--------------------------------------------

--------------------------------------------
Critical Values              TSLS      LIML
--------------------------------------------
% of Worst Case Bias
tau=5%                     25.848    15.245
tau=10%                    15.486     9.684
tau=20%                     9.817     6.569
tau=30%                     7.744     5.408
--------------------------------------------
```

5 Stored results

`weakivtest` stores the following in `r()`:

Scalars

`r(N)`	number of observations
`r(K)`	number of instruments
`r(n2)`	denominator degrees of freedom noncentral F
`r(level)`	test significance level
`r(eps)`	optimization parameter
`r(F_eff)`	effective F statistic
`r(c_TSLS_5)`	TSLS critical value for $\tau = 5\%$
`r(c_TSLS_10)`	TSLS critical value for $\tau = 10\%$
`r(c_TSLS_20)`	TSLS critical value for $\tau = 20\%$
`r(c_TSLS_30)`	TSLS critical value for $\tau = 30\%$
`r(c_LIML_5)`	LIML critical value for $\tau = 5\%$
`r(c_LIML_10)`	LIML critical value for $\tau = 10\%$
`r(c_LIML_20)`	LIML critical value for $\tau = 20\%$
`r(c_LIML_30)`	LIML critical value for $\tau = 30\%$
`r(c_simp_5)`	TSLS simplified conservative critical value for $\tau = 5\%$
`r(c_simp_10)`	TSLS simplified conservative critical value for $\tau = 10\%$
`r(c_simp_20)`	TSLS simplified conservative critical value for $\tau = 20\%$
`r(c_simp_30)`	TSLS simplified conservative critical value for $\tau = 30\%$
`r(x_TSLS_5)`	TSLS noncentrality parameter for $\tau = 5\%$
`r(x_TSLS_10)`	TSLS noncentrality parameter for $\tau = 10\%$
`r(x_TSLS_20)`	TSLS noncentrality parameter for $\tau = 20\%$
`r(x_TSLS_30)`	TSLS noncentrality parameter for $\tau = 30\%$

`r(K_eff_TSLS_5)`	TSLS effective degrees of freedom for $\tau = 5\%$
`r(K_eff_TSLS_10)`	TSLS effective degrees of freedom for $\tau = 10\%$
`r(K_eff_TSLS_20)`	TSLS effective degrees of freedom for $\tau = 20\%$
`r(K_eff_TSLS_30)`	TSLS effective degrees of freedom for $\tau = 30\%$
`r(x_LIML_5)`	LIML noncentrality parameter for $\tau = 5\%$
`r(x_LIML_10)`	LIML noncentrality parameter for $\tau = 10\%$
`r(x_LIML_20)`	LIML noncentrality parameter for $\tau = 20\%$
`r(x_LIML_30)`	LIML noncentrality parameter for $\tau = 30\%$
`r(K_eff_LIML_5)`	LIML effective degrees of freedom for $\tau = 5\%$
`r(K_eff_LIML_10)`	LIML effective degrees of freedom for $\tau = 10\%$
`r(K_eff_LIML_20)`	LIML effective degrees of freedom for $\tau = 20\%$
`r(K_eff_LIML_30)`	LIML effective degrees of freedom for $\tau = 30\%$
`r(x_simp_5)`	TSLS simplified noncentrality parameter for $\tau = 5\%$
`r(x_simp_10)`	TSLS simplified noncentrality parameter for $\tau = 10\%$
`r(x_simp_20)`	TSLS simplified noncentrality parameter for $\tau = 20\%$
`r(x_simp_30)`	TSLS simplified noncentrality parameter for $\tau = 30\%$
`r(K_eff_simp_5)`	TSLS simplified effective degrees of freedom for $\tau = 5\%$
`r(K_eff_simp_10)`	TSLS simplified effective degrees of freedom for $\tau = 10\%$
`r(K_eff_simp_20)`	TSLS simplified effective degrees of freedom for $\tau = 20\%$
`r(K_eff_simp_30)`	TSLS simplified effective degrees of freedom for $\tau = 30\%$

6 References

Andrews, D. W. K., and J. H. Stock. 2006. Inference with weak instruments. In *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress, Vol. 3*, ed. R. Blundell, W. K. Newey, and T. Persson, 122–173. Cambridge: Cambridge University Press.

Baum, C. F., and M. E. Schaffer. 2013. avar: Stata module to perform asymptotic covariance estimation for iid and non-iid data robust to heteroskedasticity, autocorrelation, 1- and 2-way clustering, and common cross-panel autocorrelated disturbances. Statistical Software Components S457689, Department of Economics, Boston College. http://ideas.repec.org/c/boc/bocode/s457689.html.

Baum, C. F., M. E. Schaffer, and S. Stillman. 2007. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *Stata Journal* 7: 465–506.

———. 2010. ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression. Statistical Software Components S425401, Department of Economics, Boston College.
http://ideas.repec.org/c/boc/bocode/s425401.html.

Campbell, J. Y. 2003. Consumption-based asset pricing. In *Handbook of the Economics of Finance: Financial Markets and Asset Pricing, Vol. 1B*, ed. G. Constantinides, M. Harris, and R. M. Stulz, 803–887. Amsterdam: Elsevier.

Cragg, J. G., and S. G. Donald. 1993. Testing identifiability and specification in instrumental variables models. *Econometric Theory* 9: 222–240.

Finlay, K., L. Magnusson, and M. E. Schaffer. 2014. weakiv: Stata module to perform weak-instrument-robust tests and confidencce intervals for instrumental-variable(IV) estimation of linear, probit and tobit models. Statistical Software Components S457684, Department of Economics, Boston College. http://ideas.repec.org/c/boc/bocode/s457684.html.

Imhof, J. P. 1961. Computing the distribution of quadratic forms in normal variables. *Biometrika* 48: 419–426.

Kleibergen, F., and R. Paap. 2006. Generalized reduced rank tests using the singular-value decomposition. *Journal of Econometrics* 133: 97–126.

Montiel Olea, J. L., and C. E. Pflueger. 2013. A robust test for weak instruments. *Journal of Business and Economic Statistics* 31: 358–369.

Nagar, A. L. 1959. The bias and moment matrix of the general k-class estimators of the parameters in simultaneous equations. *Econometrica* 27: 575–595.

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

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703–708.

Patnaik, P. 1949. The non-central χ^2 and F-distributions and their applications. *Biometrika* 36: 202–232.

Staiger, D., and J. H. Stock. 1997. Instrumental variables regression with weak instruments. *Econometrica* 65: 557–86.

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.

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

Yogo, M. 2004. Estimating the elasticity of intertemporal substitution when instruments are weak. *Review of Economics and Statistics* 86: 797–810.

About the authors

Carolin Pflueger is an assistant professor of finance in the Sauder School of Business at the University of British Columbia.

Su Wang is a PhD candidate in the finance department at the London School of Economics and Political Science, London, UK.