IV and IV-GMM

Christopher F Baum

EC 823: Applied Econometrics

Boston College, Spring 2014

The IV–GMM estimator

To discuss the implementation of IV estimators and test statistics, we consider a more general framework: an instrumental variables estimator implemented using the Generalized Method of Moments (GMM). As we will see, conventional IV estimators such as two-stage least squares (2SLS) are special cases of this IV-GMM estimator.

The model:

$$y = X\beta + u, \quad u \sim (0, \Omega)$$

with X $(N \times k)$ and define a matrix Z $(N \times \ell)$ where $\ell \geq k$. This is the Generalized Method of Moments IV (IV-GMM) estimator.

The ℓ instruments give rise to a set of ℓ moments:

$$g_i(\beta) = Z_i' u_i = Z_i'(y_i - x_i \beta), \; i = 1, N$$

where each g_i is an ℓ-vector. The method of moments approach considers each of the ℓ moment equations as a sample moment, which we may estimate by averaging over N:

$$\bar{g}(\beta) = \frac{1}{N} \sum_{i=1}^{N} z_i (y_i - x_i \beta) = \frac{1}{N} Z' u$$

The GMM approach chooses an estimate that solves $\bar{g}(\hat{\beta}_{GMM}) = 0$.

If $\ell = k$, the equation to be estimated is said to be *exactly identified* by the *order condition* for identification: that is, there are as many excluded instruments as included right-hand endogenous variables. The method of moments problem is then k equations in k unknowns, and a unique solution exists, equivalent to the standard IV estimator:

$$\hat{\beta}_{IV} = (Z'X)^{-1}Z'y$$

In the case of *overidentification* ($\ell > k$) we may define a set of k instruments:

$$\hat{X} = Z(Z'Z)^{-1}Z'X = P_Z X$$

which gives rise to the *two-stage least squares* (2SLS) estimator

$$\hat{\beta}_{2SLS} = (\hat{X}'X)^{-1}\hat{X}'y = (X'P_Z X)^{-1}X'P_Z y$$

which despite its name is computed by this single matrix equation.

In the 2SLS method with overidentification, the ℓ available instruments are “boiled down" to the k needed by defining the P_Z matrix. In the IV-GMM approach, that reduction is not necessary. All ℓ instruments are used in the estimator. Furthermore, a *weighting matrix* is employed so that we may choose $\hat{\beta}_{GMM}$ so that the elements of $\bar{g}(\hat{\beta}_{GMM})$ are as close to zero as possible. With $\ell > k$, not all ℓ moment conditions can be exactly satisfied, so a criterion function that weights them appropriately is used to improve the efficiency of the estimator.

The GMM estimator minimizes the criterion

$$J(\hat{\beta}_{GMM}) = N\,\bar{g}(\hat{\beta}_{GMM})'W\,\bar{g}(\hat{\beta}_{GMM})$$

where W is a $\ell \times \ell$ symmetric weighting matrix.

Solving the set of FOCs, we derive the IV-GMM estimator of an overidentified equation:

$$\hat{\beta}_{GMM} = (X'ZWZ'X)^{-1}X'ZWZ'y$$

which will be identical for all W matrices which differ by a factor of proportionality. The *optimal* weighting matrix, as shown by Hansen (1982), chooses $W = S^{-1}$ where S is the covariance matrix of the moment conditions to produce the most *efficient* estimator:

$$S = E[Z'uu'Z] = lim_{N\to\infty}\ N^{-1}[Z'\Omega Z]$$

With a consistent estimator of S derived from 2SLS residuals, we define the feasible IV-GMM estimator as

$$\hat{\beta}_{FEGMM} = (X'Z\ \hat{S}^{-1}Z'X)^{-1}X'Z\ \hat{S}^{-1}Z'y$$

where *FEGMM* refers to the *feasible efficient* GMM estimator.

IV-GMM and the distribution of u

The derivation makes no mention of the form of Ω, the variance-covariance matrix (*vce*) of the error process u. If the errors satisfy all classical assumptions are $i.i.d.$, $S = \sigma_u^2 I_N$ and the optimal weighting matrix is proportional to the identity matrix. The IV-GMM estimator is merely the standard IV (or 2SLS) estimator.

IV-GMM robust estimates

If there is heteroskedasticity of unknown form, we usually compute *robust* standard errors in any Stata estimation command to derive a consistent estimate of the *vce*. In this context,

$$\hat{S} = \frac{1}{N} \sum_{i=1}^{N} \hat{u}_i^2 Z_i' Z_i$$

where $\hat{u}$ is the vector of residuals from any consistent estimator of β (e.g., the 2SLS residuals). For an overidentified equation, the IV-GMM estimates computed from this estimate of S will be *more efficient* than 2SLS estimates.

IV-GMM cluster-robust estimates

If errors are considered to exhibit arbitrary intra-cluster correlation in a dataset with M clusters, we may derive a *cluster-robust* IV-GMM estimator using

$$\hat{S} = \sum_{j=1}^{M} \hat{u}_j' \hat{u}_j$$

where

$$\hat{u}_j = (y_j - x_j \hat{\beta}) X' Z (Z'Z)^{-1} z_j$$

The IV-GMM estimates employing this estimate of S will be both robust to arbitrary heteroskedasticity and intra-cluster correlation, equivalent to estimates generated by Stata's `cluster(`*varname*`)` option. For an overidentified equation, IV-GMM cluster-robust estimates will be *more efficient* than 2SLS estimates.

IV-GMM HAC estimates

The IV-GMM approach may also be used to generate *HAC standard errors*: those robust to arbitrary heteroskedasticity and autocorrelation. Although the best-known *HAC* approach in econometrics is that of Newey and West, using the Bartlett kernel (per Stata's `newey`), that is only one choice of a *HAC* estimator that may be applied to an IV-GMM problem.

Baum–Schaffer–Stillman's `ivreg2` (from the SSC Archive) and Stata's `ivregress` provide several choices for kernels. For some kernels, the kernel *bandwidth* (roughly, number of lags employed) may be chosen automatically in either command.

Example of IV and IV-GMM estimation

We illustrate various forms of the IV estimator with a model of US real import growth constructed with US quarterly data from a recent edition of *International Financial Statistics*. The model seeks to explain the growth rate (change in the log) of US real imports. In the initial form of the model, we include as regressors the growth rate of real GDP, the lagged rate of change of the REER (real effective exchange rate), and the rate of change of real crude oil prices.

We first fit the relationship with the standard 2SLS estimator, assuming *i.i.d.* errors, using Baum–Schaffer–Stillman's `ivreg2` command. You could fit the same equation with `ivregress 2sls`.

We model US real import growth considering that the contemporaneous growth rate of real GDP may be endogenous to this relationship. We use the first three lags of GDP growth as instruments for the current growth rate. Some of the standard `ivreg2` output, relating to weak instruments, has been edited on the following slides.

```
. ivreg2 dlrimports (dlrgdp = L(1/3).dlrgdp) ldlreer dlroilprice
IV (2SLS) estimation
--------------------

Estimates efficient for homoskedasticity only
Statistics consistent for homoskedasticity only
                                                      Number of obs =      138
                                                      F(  3,    134) =    23.84
                                                      Prob > F      =   0.0000
Total (centered) SS     =  .1872911248                Centered R2   =   0.0706
Total (uncentered) SS   =  .2086385951                Uncentered R2 =   0.1657
Residual SS             =  .1740637032                Root MSE      =   .03552

------------------------------------------------------------------------------
  dlrimports |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      dlrgdp |   5.022829   .9923138     5.06   0.000     3.077929    6.967728
     ldlreer |  -.2971572   .0931814    -3.19   0.001    -.4797895    -.114525
 dlroilprice |   .1084789    .022928     4.73   0.000     .0635409     .153417
       _cons |  -.0245364   .0077375    -3.17   0.002    -.0397016   -.0093713
------------------------------------------------------------------------------
Sargan statistic (overidentification test of all instruments):           1.999
                                                  Chi-sq(2) P-val =    0.3680
------------------------------------------------------------------------------
Instrumented:          dlrgdp
Included instruments: ldlreer dlroilprice
Excluded instruments: L.dlrgdp L2.dlrgdp L3.dlrgdp
------------------------------------------------------------------------------
```

We may fit this equation with different assumptions about the error process. The estimates above assume *i.i.d.* errors. We may also compute robust standard errors in the 2SLS context.

We then apply IV-GMM with robust standard errors. As the equation is overidentified, the IV-GMM estimates will differ, and will be more efficient than the robust 2SLS estimates.

Last, we may estimate the equation with IV-GMM and HAC standard errors, using the default Bartlett kernel (as employed by Newey–West) and a bandwidth of 5 quarters. This corresponds to four lags in the `newey` command.

```
. estimates table IID Robust IVGMM IVGMM_HAC, b(%9.4f) t(%5.2f) ///
> title("Alternative IV estimates of real US import growth") stat(rmse)
```

Alternative IV estimates of real US import growth

Variable	IID	Robust	IVGMM	IVGMM_HAC
dlrgdp	5.0228	5.0228	5.0197	4.6662
	5.06	5.33	5.32	5.74
ldlreer	-0.2972	-0.2972	-0.3337	-0.3462
	-3.19	-2.31	-2.95	-3.40
dlroilprice	0.1085	0.1085	0.1067	0.1100
	4.73	6.51	6.44	7.78
_cons	-0.0245	-0.0245	-0.0250	-0.0232
	-3.17	-3.14	-3.22	-3.66
rmse	0.0355	0.0355	0.0355	0.0337

legend: b/t

Note that the coefficients' point estimates change when IV-GMM is employed, and that their *t*-statistics are larger than those of robust IV. The point estimates are also altered when IV-GMM with HAC VCE is computed.

Tests of overidentifying restrictions

If and only if an equation is *overidentified*, with more excluded instruments than included endogenous variables, we may test whether the excluded instruments are appropriately independent of the error process. That test should always be performed when it is possible to do so, as it allows us to evaluate the validity of the instruments.

A test of *overidentifying restrictions* regresses the residuals from an IV or 2SLS regression on all instruments in Z. Under the null hypothesis that all instruments are uncorrelated with u, the test has a large-sample $\chi^2(r)$ distribution where r is the number of overidentifying restrictions.

Under the assumption of *i.i.d.* errors, this is known as a *Sargan test*, and is routinely produced by `ivreg2` for IV and 2SLS estimates. After `ivregress`, the command `estat overid` provides the test.

If we have used IV-GMM estimation in `ivreg2`, the test of overidentifying restrictions becomes the Hansen J statistic: the GMM criterion function. Although J will be identically zero for any exactly-identified equation, it will be positive for an overidentified equation. If it is “too large”, doubt is cast on the satisfaction of the moment conditions underlying GMM.

The test in this context is known as the *Hansen test* or *J test*, and is calculated by `ivreg2` when the `gmm2s` option is employed.

The Sargan–Hansen test of overidentifying restrictions should be performed routinely in any overidentified model estimated with instrumental variables techniques. Instrumental variables techniques are powerful, but if a strong rejection of the null hypothesis of the Sargan–Hansen test is encountered, you should strongly doubt the validity of the estimates.

For instance, consider a variation of the IV-GMM model estimated above (with robust standard errors) and focus on the test of overidentifying restrictions provided by the Hansen J statistic.

In this form of the model, we also include the lagged growth rate of real oil prices as an excluded instrument. The model is overidentified by three degrees of freedom, as there is one endogenous regressor and four excluded instruments. We see that the J statistic clearly rejects its null, casting doubt on our choice of instruments.

```
. ivreg2 dlrimports (dlrgdp = L(1/3).dlrgdp L.dlroilprice) ldlreer dlroilprice,
>  robust gmm2s
2-Step GMM estimation
---------------------

Estimates efficient for arbitrary heteroskedasticity
Statistics robust to heteroskedasticity
                                                       Number of obs =       138
                                                       F(  3,    134) =     27.95
                                                       Prob > F      =    0.0000
Total (centered) SS     =  .1872911248                 Centered R2   =    0.1139
Total (uncentered) SS   =  .2086385951                 Uncentered R2 =    0.2045
Residual SS             =  .1659629526                 Root MSE      =    .03468

------------------------------------------------------------------------------
             |               Robust
  dlrimports |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      dlrgdp |   4.849372   .9087537     5.34   0.000     3.068247    6.630496
     ldlreer |  -.3379459    .108448    -3.12   0.002       -.5505   -.1253917
 dlroilprice |   .0967915   .0150484     6.43   0.000     .0672972    .1262858
       _cons |  -.0249144   .0075339    -3.31   0.001    -.0396806   -.0101482
------------------------------------------------------------------------------
Hansen J statistic (overidentification test of all instruments):        10.346
                                                   Chi-sq(3) P-val =    0.0158
------------------------------------------------------------------------------
Instrumented:         dlrgdp
Included instruments: ldlreer dlroilprice
Excluded instruments: L.dlrgdp L2.dlrgdp L3.dlrgdp L.dlroilprice
------------------------------------------------------------------------------
```

We reestimate the model, retaining real oil price growth as an exogenous variable, but including it in the estimated equation rather than applying an exclusion restriction. The resulting J statistic now fails to reject its null.

```
. ivreg2 dlrimports (dlrgdp = L(1/3).dlrgdp) ldlreer dlroilprice L.dlroilprice,
>  robust gmm2s
2-Step GMM estimation
---------------------

Estimates efficient for arbitrary heteroskedasticity
Statistics robust to heteroskedasticity
                                                      Number of obs =      138
                                                      F(  4,   133) =    33.39
                                                      Prob > F      =   0.0000
Total (centered) SS     =  .1872911248                Centered R2   =   0.2002
Total (uncentered) SS   =  .2086385951                Uncentered R2 =   0.2821
Residual SS             =  .1497892412                Root MSE      =   .03295

------------------------------------------------------------------------------
             |               Robust
  dlrimports |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      dlrgdp |     4.7493   .8256717     5.75   0.000     3.131013    6.367586
     ldlreer |  -.2660648   .1157742    -2.30   0.022    -.4929782   -.0391515
             |
 dlroilprice |
         --. |    .092877   .0130212     7.13   0.000     .0673559    .1183981
         L1. |   .0666371    .021165     3.15   0.002     .0251545    .1081197
             |
       _cons |  -.0230283   .0067559    -3.41   0.001    -.0362697   -.0097869
------------------------------------------------------------------------------
Hansen J statistic (overidentification test of all instruments):         1.816
                                                   Chi-sq(2) P-val =    0.4033
------------------------------------------------------------------------------
Included instruments: ldlreer dlroilprice L.dlroilprice
```

It is important to understand that the Sargan–Hansen test of overidentifying restrictions is a joint test of the hypotheses that the instruments, excluded and included, are independently distributed of the error process *and* that they are properly excluded from the model.

Note as well that all exogenous variables in the equation—excluded and included—appear in the set of instruments Z. In the context of single-equation IV estimation, they must. You cannot pick and choose which instruments appear in which 'first-stage' regressions.

Testing a subset of overidentifying restrictions

We may be quite confident of some instruments' independence from *u* but concerned about others. In that case a *GMM distance* or *C* test may be used. The `orthog( )` option of `ivreg2` tests whether a *subset* of the model's overidentifying restrictions appear to be satisfied.

This is carried out by calculating two Sargan–Hansen statistics: one for the full model and a second for the model in which the listed variables are (a) considered endogenous, if included regressors, or (b) dropped, if excluded regressors. In case (a), the model must still satisfy the order condition for identification. The difference of the two Sargan–Hansen statistics, often termed the *GMM distance* or Hayashi *C statistic*, will be distributed χ^2 under the null hypothesis that the specified orthogonality conditions are satisfied, with d.f. equal to the number of those conditions.

We perform the *C* test on the estimated equation by challenging the exogeneity of `ldlreer`. Is it properly considered exogenous? The `orthog()` option reestimates the equation, treating it as endogenous, and evaluates the difference in the *J* statistics from the two models. Considering `ldlreer` as exogenous is essentially imposing one more orthogonality condition on the GMM estimation problem.

```
. ivreg2 dlrimports (dlrgdp = L(1/3).dlrgdp) ldlreer dlroilprice L.dlroilprice,
>  robust gmm2s orthog(ldlreer)
2-Step GMM estimation
---------------------

Estimates efficient for arbitrary heteroskedasticity
Statistics robust to heteroskedasticity
                                                      Number of obs =       138
                                                      F(  4,    133) =     33.39
                                                      Prob > F      =    0.0000
Total (centered) SS     =  .1872911248                Centered R2   =    0.2002
Total (uncentered) SS   =  .2086385951                Uncentered R2 =    0.2821
Residual SS             =  .1497892412                Root MSE      =    .03295
...
-------------------------------------------------------------------------------
Hansen J statistic (overidentification test of all instruments):         1.816
                                                   Chi-sq(2) P-val =    0.4033
-orthog- option:
Hansen J statistic (eqn. excluding suspect orthog. conditions):          0.456
                                                   Chi-sq(1) P-val =    0.4997
C statistic (exogeneity/orthogonality of suspect instruments):           1.361
                                                   Chi-sq(1) P-val =    0.2434
Instruments tested:   ldlreer
-------------------------------------------------------------------------------
Instrumented:         dlrgdp
Included instruments: ldlreer dlroilprice L.dlroilprice
Excluded instruments: L.dlrgdp L2.dlrgdp L3.dlrgdp
-------------------------------------------------------------------------------
```

It appears that `ldlreer` may be considered exogenous in this specification.

A variant on this strategy is implemented by the `endog( )` option of `ivreg2`, in which one or more variables considered endogenous can be tested for exogeneity. The *C* test in this case will consider whether the null hypothesis of their exogeneity is supported by the data.

If all endogenous regressors are included in the `endog( )` option, the test is essentially a test of whether IV methods are required to estimate the equation. If OLS estimates of the equation are consistent, they should be preferred. In this context, the test is equivalent to a *(Durbin–Wu–)Hausman test* comparing IV and OLS estimates, as implemented by Stata's `hausman` command with the `sigmaless` option. Using `ivreg2`, you need not estimate and store both models to generate the test's verdict.

For instance, with the model above, we might question whether IV techniques are needed. We can conduct the *C* test via:

```
ivreg2 dlrimports (dlrgdp = L(1/3).dlrgdp) ldlreer ///
dlroilprice L.dlroilprice, robust gmm2s endog(dlrgdp)
```

where the `endog(dlrgdp)` option tests the null hypothesis that the variable can be treated as exogenous in this model, rather than as an endogenous variable.

```
. ivreg2 dlrimports (dlrgdp = L(1/3).dlrgdp) ldlreer dlroilprice L.dlroilprice,
>  ///
> robust gmm2s endog(dlrgdp)
2-Step GMM estimation
---------------------

Estimates efficient for arbitrary heteroskedasticity
Statistics robust to heteroskedasticity
                                                      Number of obs =      138
                                                      F(  4,    133) =    33.39
                                                      Prob > F      =   0.0000
Total (centered) SS     =  .1872911248                Centered R2   =   0.2002
Total (uncentered) SS   =  .2086385951                Uncentered R2 =   0.2821
Residual SS             =  .1497892412                Root MSE      =   .03295
...
------------------------------------------------------------------------------
Hansen J statistic (overidentification test of all instruments):         1.816
                                                   Chi-sq(2) P-val =    0.4033
-endog- option:
Endogeneity test of endogenous regressors:                              11.736
                                                   Chi-sq(1) P-val =    0.0006
Regressors tested:    dlrgdp
------------------------------------------------------------------------------
Instrumented:         dlrgdp
Included instruments: ldlreer dlroilprice L.dlroilprice
Excluded instruments: L.dlrgdp L2.dlrgdp L3.dlrgdp
------------------------------------------------------------------------------
```

In this context, it appears that we cannot consistently estimate this equation with OLS techniques, as the null hypothesis that `dlrgdp` can be treated as exogenous is strongly rejected by the data.

The weak instruments problem

Instrumental variables methods rely on two assumptions: the excluded instruments are distributed independently of the error process, and they are sufficiently correlated with the included endogenous regressors. Tests of overidentifying restrictions address the *first* assumption, although we should note that a rejection of their null may be indicative that the exclusion restrictions for these instruments may be inappropriate. That is, some of the instruments have been improperly excluded from the regression model's specification.

The specification of an instrumental variables model asserts that the excluded instruments affect the dependent variable only *indirectly*, through their correlations with the included endogenous variables. If an excluded instrument exerts both direct and indirect influences on the dependent variable, the exclusion restriction should be rejected. This can be readily tested by including the variable as a regressor.

To test the *second* assumption—that the excluded instruments are sufficiently correlated with the included endogenous regressors—we should consider the goodness-of-fit of the "first stage" regressions relating each endogenous regressor to the entire set of instruments.

It is important to understand that the theory of single-equation ("limited information") IV estimation requires that all columns of X are conceptually regressed on all columns of Z in the calculation of the estimates. We cannot meaningfully speak of "this variable is an instrument for that regressor" or somehow restrict which instruments enter which first-stage regressions. Stata's `ivregress` or `ivreg2` will not let you do that because such restrictions only make sense in the context of estimating an entire system of equations by full-information methods (for instance, with `reg3`).

The `first` and `ffirst` options of `ivreg2` (or the `first` option of `ivregress`) present several useful diagnostics that assess the first-stage regressions. If there is a single endogenous regressor, these issues are simplified, as the instruments either explain a reasonable fraction of that regressor's variability or not. With multiple endogenous regressors, diagnostics are more complicated, as each instrument is being called upon to play a role in each first-stage regression.

With sufficiently weak instruments, the asymptotic identification status of the equation is called into question. An equation identified by the order and rank conditions in a finite sample may still be *effectively unidentified* or *numerically unidentified*.

As Staiger and Stock (*Econometrica*, 1997) show, the weak instruments problem can arise even when the first-stage t- and F-tests are significant at conventional levels in a large sample. In the worst case, the bias of the IV estimator is the same as that of OLS, IV becomes inconsistent, and instrumenting only aggravates the problem.

Beyond the informal "rule-of-thumb" diagnostics such as $F > 10$, `ivreg2` computes several statistics that can be used to critically evaluate the strength of instruments. We can write the first-stage regressions as

$$X = Z\Pi + v$$

With X_1 as the endogenous regressors, Z_1 the excluded instruments and Z_2 as the included instruments, this can be partitioned as

$$X_1 = [Z_1 Z_2] \, [\Pi'_{11} \Pi'_{12}]' + v_1$$

The rank condition for identification states that the $L \times K_1$ matrix Π_{11} must be of full column rank.

We do not observe the true Π_{11}, so we must replace it with an estimate. Anderson's (1984) approach to testing the rank of this matrix (or that of the full Π matrix) considers the *canonical correlations* of the X and Z matrices. If the equation is to be identified, all K of the canonical correlations will be significantly different from zero.

The squared canonical correlations can be expressed as eigenvalues of a matrix. Anderson's *CC* test considers the null hypothesis that the minimum canonical correlation is zero. Under the null, the test statistic is distributed χ^2 with $(L - K + 1)$ d.f., so it may be calculated even for an exactly-identified equation. Failure to reject the null suggests the equation is unidentified. `ivreg2` routinely reports this Lagrange Multiplier (LM) statistic.

The C–D statistic is a closely related test of the rank of a matrix. While the Anderson *CC* test is a LR test, the C–D test is a Wald statistic, with the same asymptotic distribution. The C–D statistic plays an important role in Stock and Yogo's work (see below). Both the Anderson and C–D tests are reported by `ivreg2` with the `first` option.

Research by Kleibergen and Paap (KP) (*J. Econometrics*, 2006) has developed a robust version of a test for the rank of a matrix: e.g. testing for *underidentification*. The statistic has been implemented by Kleibergen and Schaffer as command `ranktest`, which is part of the `ivreg2` package. If non-*i.i.d.* errors are assumed, the `ivreg2` output contains the K–P `rk` statistic in place of the Anderson canonical correlation statistic as a test of underidentification.

The canonical correlations may also be used to test a set of instruments for *redundancy* by considering their statistical significance in the first stage regressions. This can be calculated, in robust form, as a K–P LM test. The `redundant( )` option of `ivreg2` allows a set of excluded instruments to be tested for relevance, with the null hypothesis that they do not contribute to the asymptotic efficiency of the equation.

Stock and Yogo (Camb. U. Press festschrift, 2005) propose testing for weak instruments by using the F-statistic form of the C–D statistic. Their null hypothesis is that the estimator is weakly identified in the sense that it is subject to bias that the investigator finds unacceptably large.

Their test comes in two flavors: maximal relative bias (relative to the bias of OLS) and maximal size. The former test has the null that instruments are weak, where weak instruments are those that can lead to an asymptotic relative bias greater than some level b. This test uses the finite sample distribution of the IV estimator, and can only be calculated where the appropriate moments exist (when the equation is suitably overidentified: the m^{th} moment of an IV estimator exists iff $m < (L - K + 1)$). The test is routinely reported in `ivreg2` and `ivregress` output when it can be calculated, with the relevant critical values calculated by Stock and Yogo.

The second test proposed by Stock and Yogo is based on the performance of the Wald test statistic for the endogenous regressors. Under weak identification, the test rejects too often. The test statistic is based on the rejection rate *r* tolerable to the researcher if the true rejection rate is 5%. Their tabulated values consider various values for *r*. To be able to reject the null that the size of the test is unacceptably large (versus 5%), the Cragg–Donald *F* statistic must exceed the tabulated critical value.

The Stock–Yogo test statistics, like others discussed above, assume *i.i.d.* errors. The Cragg–Donald *F* can be robustified in the absence of *i.i.d.* errors by using the Kleibergen–Paap `rk` statistic, which `ivreg2` reports in that circumstance.

LIML and GMM-CUE

OLS and IV estimators are special cases of *k-class estimators*: OLS with $k = 0$ and IV with $k = 1$. Limited-information maximum likelihood (LIML) is another member of this class, with k chosen optimally in the estimation process. Like any ML estimator, LIML is invariant to normalization. In an equation with two endogenous variables, it does not matter whether you specify y_1 or y_2 as the left-hand variable.

One of the other virtues of the LIML estimator is that it has been found to be more resistant to weak instruments problems than the IV estimator. On the down side, it makes the distributional assumption of normally distributed (and *i.i.d.*) errors. `ivreg2` produces LIML estimates with the `liml` option, and `liml` is a subcommand for official Stata's `ivregress`.

If the *i.i.d.* assumption of LIML is not reasonable, you may use the GMM equivalent: the *continuously updated* GMM estimator, or CUE estimator. In `ivreg2`, the `cue` option combined with `robust`, `cluster` and/or `bw( )` options specifies that non-*i.i.d.* errors are to be modeled. GMM-CUE requires numerical optimization, and may require many iterations to converge.

`ivregress` provides an iterated GMM estimator, which is not the same estimator as GMM-CUE.

Testing for *i.i.d.* errors in IV

In the context of an equation estimated with instrumental variables, the standard diagnostic tests for heteroskedasticity and autocorrelation are generally not valid.

In the case of heteroskedasticity, Pagan and Hall (*Econometric Reviews*, 1983) showed that the Breusch–Pagan or Cook–Weisberg tests (`estat hettest`) are generally not usable in an IV setting. They propose a test that will be appropriate in IV estimation where heteroskedasticity may be present in more than one structural equation. Mark Schaffer's `ivhettest`, part of the `ivreg2` suite, performs the Pagan–Hall test under a variety of assumptions on the indicator variables. It will also reproduce the Breusch–Pagan test if applied in an OLS context.

In the same token, the Breusch–Godfrey statistic used in the OLS context (`estat bgodfrey`) will generally not be appropriate in the presence of endogenous regressors, overlapping data or conditional heteroskedasticity of the error process. Cumby and Huizinga (*Econometrica*, 1992) proposed a generalization of the BG statistic which handles each of these cases.

Their test is actually more general in another way. Its null hypothesis of the test is that the regression error is a moving average of known order $q \geq 0$ against the general alternative that autocorrelations of the regression error are nonzero at lags greater than q. In that context, it can be used to test that autocorrelations beyond any q are zero. Like the BG test, it can test multiple lag orders. The C–H test is available as Baum and Schaffer's `ivactest` routine, part of the `ivreg2` suite.

For more details on IV and IV-GMM, please see

- Enhanced routines for instrumental variables/GMM estimation and testing. Baum, C.F., Schaffer, M.E., Stillman, S., *Stata Journal* 7:4, 2007.
- *An Introduction to Modern Econometrics Using Stata*, Baum, C.F., Stata Press, 2006 (particularly Chapter 8).
- Instrumental variables and GMM: Estimation and testing. Baum, C.F., Schaffer, M.E., Stillman, S., *Stata Journal* 3:1–31, 2003.

Both of the *Stata Journal* papers are freely downloadable from `http://stata-journal.com`.