Title

gmmhart— HAR Test for linear hypotheses in a two-step efficient gmm framework after performing gmmhar estimation

Syntax Menu Description Options Remarks and example Stored results Methods and formula References Also see

Syntax

Basic syntax

gmmhart coeflist

gmmhart exp=exp[=…]

options	Description
Model	
accumulate	test the hypothesis jointly with previously tested hypotheses

time-series operators are allowed.

Syntax 1 tests that coefficients are 0.

Syntax 2 tests that linear expressions are equal.

Menu

statistics>Postestimation>Tests> GMM HAR test

Description

The Stata command gmmhart performs the Wald tests of simple and composite linear hypotheses about the parameters in the most recently estimated gmmhar model. See Hwang and Sun (2017) for details.

Options

Model

accumulate allows a hypothesis to be tested jointly with the previously tested hypotheses.

Remarks and examples

Hwang and Sun (2017) considers the more accurate fixed-smoothing asymptotics in the two-step efficient GMM framework where the weighting matrix and the asymptotic variance matrix are based on the orthonormal series long run variance estimator. Hwang and Sun (2017) propose some modifications to the usual Wald statistic and show that the modified test statistic is asymptotically standard F distributed under the fixed-smoothing asymptotics. The modified statistic is a rescaled version of the original test statistic with the rescaling factor depending on the J statistic (Sun and Kim, 2012) for testing over-identifying restrictions. The data-driven K, the number of orthonormal bases used in the long run variance estimator, is selected to minimize the AMSE of the estimator of the long run variance of the moment process.

Δ Example 1: Test for a single coefficient against zero

We estimate the following regression:

```
. use http://fmwww.bc.edu/ec-p/data/stockwatson/macrodat

.
. generate inf =100 * log( CPI / L4.CPI )
(4 missing values generated)

.
. generate ggdp=100 * log( GDP / L4.GDP )
(10 missing values generated)


. gmmhar  D.inf L.ER (UR=L2.ggdp L.TBILL L.TBON)

Two-step Efficient GMM Estimation                  Number of obs  =       158
Data-driven optimal K: 48                          F(  2,     45)  =      2.42
                                                   Prob > F        =    0.1007

------------------------------------------------------------------------------------------
                          HAR
       D.inf       Coef.    std.Err.       t       df   P>|t|     [95% Conf. Interval]
------------------------------------------------------------------------------------------
          UR   -.1993698   .0940415     -2.12      46   0.039    -.3886655   -.0100741
        L.ER   -.0017818   .0015101     -1.18      46   0.244    -.0048215    .001258
       _cons    1.536211   .8235799      1.87      46   0.069    -.1215691   3.193992
------------------------------------------------------------------------------------------

HAR J statistic  =  .72631984
Reference Dist for the J test: F(  2,     47)
P-value of the J test =    0.4890
Instrumented: UR
Instruments: L2.ggdp L.TBILL L.TBON
```

We can test the hypothesis that the coefficient on UR is zero by typing:

```
. gmmhart UR=0
       F(  1,     46) =     4.49
            Prob > F =     0.0394
```

The F statistic is 4.49. The approximating F distribution has degrees of freedom 1 and 46. The p-value of the test is 3.94%. We fail to reject the null hypothesis at the 1% level.

Δ Example 2: Testing that a coefficient is equal to a given value

We can test the hypothesis that the coefficient on UR is -0.1993698 by typing:

```
gmmhart UR=-0.1993698
      F(  1,     46) =     0.00
           Prob > F =     1.0000
```

We find that we fail to reject that hypothesis.

Δ Example 3: Testing the equality of two coefficients

Now let’s test something a bit more difficult: whether the coefficient on UR is the same as the coefficient on L.ER:

```
. gmmhart UR=L.ER
       F(  1,    46) =    4.52
            Prob > F =    0.0388
```

We fail to reject the equality hypothesis at the 1% level, but we find evidence to reject it at the 5% level.

Δ Example 4:

When we test the equality of the UR and L.ER coefficients, the command gmmhart rearranges and simplifies the equality and test whether the coefficient on UR minus the coefficient on L.ER is zero or not. The arrangement is innocuous. In fact, gmmhart can test the null of more complicated forms. For example,

```
gmmhart 3*UR-2*L.ER=2*UR-1*L.ER
       F(  1,    46) =    4.52
            Prob > F =    0.0388
```

Although we request what appears to be a lengthy hypothesis, once gmmhart simplifies the expression, it realizes that all we want to do is to test whether the coefficient on UR is the same as the coefficient on L.ER.

The gmmhart's ability to simplify and test complex hypotheses is limited to linear hypotheses. If one attempts to test a nonlinear hypothesis, an error message will be displayed. This is not a problem specific to the command gmmhart. Stata's command "test" exhibits the same behavior. In fact, gmmhart uses Stata's command "test" to parse the null hypothesis.

```
. gmmhart 3*UR/2*L.ER=2*UR-1*L.ER
not possible with test
r(131);
```

Δ Example 5: Testing joint hypotheses

We wish to test whether UR and L.ER, taken as a whole, are jointly significant by testing whether the coefficients on UR and L.ER are simultaneously zero. The command gmmhart allows us to specify multiple restrictions to be tested, each embedded within parentheses.

```
gmmhart (UR=0) (L.ER=0)
       F(  2,    45) =    2.42
            Prob > F =    0.1007
```

gmmhart displays the set of restrictions and reports an F statistic of 2.42. gmmhart also reports the key information on the approximating F distribution: its degrees of freedom are (2, 45). The p-value of the test is close to 10.7%.

☐ Technical note

An alternative method to test simultaneous hypotheses is to specify a test for each constraint and accumulate it with the previous constraints:

```
gmmhart UR=0
     F(  1,    46) =     4.49
          Prob > F =     0.0394

gmmhart L.ER=0,acc
     F(  2,    45) =     2.42
          Prob > F =     0.1007
```

We test the hypothesis that the coefficient on UR is zero by typing gmmhart UR=0. We then test whether the coefficient on L.ER is also zero by typing gmmhart L.ER=0, accumulate. The accumulate option tells gmmhart that this is not the start of a new test but a continuation of the previous one. gmmhart responds by showing us the two equations and reporting an F statistic of 2.42. The p-value of the test is 10.7%.

Δ Example 6: Testing whether coefficients are zero

Testing whether the coefficients are zero are very common in applied statistics. The gmmhart command has a more convenient syntax to accommodate this common case:

```
gmmhart UR L.ER
     F(  2,    45) =     2.42
          Prob > F =     0.1007
```

Δ Example 7: Replaying the previous test

We just used gmmhart to test the hypothesis that the coefficients on UR and L.ER are jointly zero. We can review our last test by typing gmmhart without any argument.

```
gmmhart
     F(  2,    45) =     2.42
          Prob > F =     0.1007
```

Δ Example 8: Testing the equality of multiple coefficients

Let's test the hypothesis that UR, L.ER, and L2.ggdp have the same coefficient

```
. gmmhar D.inf L2.ggdp L.ER (UR=L.TBILL L.TBON)

Two-step Efficient GMM Estimation                   Number of obs   =       158
Data-driven optimal K: 48                           F(  3,     45)  =      1.60
                                                    Prob > F        =    0.2037
```

D.inf	Coef.	HAR std.Err.	t	df	P>|t|	[95% Conf.	Interval]
UR	-.1956027	.0986306	-1.98	47	0.053	-.394022	.0028165
L2.ggdp	.0025137	.0191045	0.13	47	0.896	-.0359196	.040947
L.ER	-.0018674	.0016471	-1.13	47	0.263	-.0051809	.0014462
_cons	1.518585	.8321348	1.82	47	0.074	-.1554543	3.192624

```
HAR J statistic  =  1.4651778
Reference Dist for the J test: F(  1,    48)
P-value of the J test =    0.2320
Instrumented: UR
Instruments: L.TBILL L.TBON


. gmmhart UR=L2.ggdp=L.ER
       F(  2,    46) =     2.21
            Prob > F =     0.1216
```

The syntax UR=L.ER=L2.ggdp with multiple = operators is just a convenient shorthand for typing that the first expression equals the second expression and that the first expression equals the third expression.

We can obtain the same test by typing

```
gmmhart (UR=L.ER) (UR=L2.ggdp)
     F(  2,    46) =    2.21
          Prob > F =    0.1216
```

Equivalently, we can type

```
gmmhart (UR=L.ER) (L.ER=L2.ggdp)
     F(  2,    46) =    2.21
          Prob > F =    0.1216
```

Stored results

gmmhart stores the following results in r():

Scalars

r(firdf) the first degrees of freedom

r(secdf) the second degrees of freedom

r(kopt) the data-driven optimal K

r(F) the adjusted F statistic

Matrices

r(thetagmm) the two-step gmm coefficient vector

Methods and formulas

Consider the regression model:

$$Y_t = X_t\theta_0 + e_t, t = 1,2,\cdots,T$$

where $\{e_t\}$ is a zero-mean process that may be correlated with the covariate process $\{X_t \in R^{1\times d}\}$. There are instruments $\{Z_t \in R^{1\times m}\}$ such that the moment conditions:

$$EZ_t'(Y_t - X_t\theta_0) = 0$$

hold if and only if $\theta = \theta_0$. We allow the process $\{Z_t'e_t\}$ to have the autocorrelation of unknown forms. The model may be over-identified with the degree of over identification $q = m - d \geq 0$.

Define: $S_{ZX} = \frac{1}{T}\sum_{t=1}^{T} Z_t' X_t,\ S_{ZZ} = \frac{1}{T}\sum_{t=1}^{T} Z_t' Z_t,\ S_{ZY} = \frac{1}{T}\sum_{t=1}^{T} Z_t' Y_t.$

Then the IV estimator of θ_0 is

$$\hat{\theta}_{IV} = [S_{ZX}' W_{0T}^{-1} S_{ZX}]^{-1}[S_{ZX}' W_{0T}^{-1} S_{ZY}],$$

where $W_{0T} = S_{ZZ} \in R^{m\times m}$.

Hwang and Sun (2017) consider two-step efficient GMM estimation and inference where the weighting matrix and asymptotic variance matrix are based on the orthonormal series long run variance estimator. In its general form, the two-step GMM estimator is given by

$$\hat{\theta}_{GMM} = arg \min_{\theta\in\Theta} g_T(\theta)' w_T^{-1}(\hat{\theta}_{IV}) g_T(\theta) = \{S_{ZX}'[w_T(\hat{\theta}_{IV})]^{-1} S_{ZX}\}^{-1}\{S_{ZX}'[w_T(\hat{\theta}_{IV})]^{-1} S_{ZY}\}$$

where

$$w_T(\theta) = \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T} Q_K\left(\frac{t}{T},\frac{s}{T}\right)\left(v_t(\theta) - \bar{v}(\theta)\right)\left(v_t(\theta) - \bar{v}(\theta)\right)', \quad v_t(\theta) = Z_t{}'(Y_t - X_t\theta),$$

and $\bar{v}(\theta) = \sum_{t=1}^{T} v_t(\theta)/T$. Note that $w_T\left(\hat{\theta}_{IV}\right)$ is a quadratic estimator of the long run variance of moment process $\{v_t(\theta_0)\}$. $Q_K(r,s)$ is a symmetric weighting function that depends on the smoothing parameter K. Hwang and Sun (2017) focus on the orthonormal series LRV estimation with $Q_K(r,s) = \frac{1}{K}\sum_{j=1}^{K} \phi_j(r)\,\phi_j(s)$, where $\left\{\phi_j(\cdot)\right\}_{j=1}^{K}$ are orthonormal basis functions on $L^2[0,1]$ satisfying $\int_0^1 \phi_j(r)\,dr = 0$ for $j = 1,2,\cdots,K$.

We are interested in testing the null $H_0{:}\,R\theta_0 = r$ against the alternative $H_1{:}\,R\theta_0 \neq r$, where $r \in R^{p\times 1}$ and $R \in R^{p\times d}$ is a matrix of full row rank. Nonlinear restrictions can be converted into linear ones via the delta method. The Wald statistic is given by

$$\mathbb{W}_T\left(\hat{\theta}_{GMM}\right) = \sqrt{T}\left(R\hat{\theta}_{GMM} - r\right)'\{R\left(G_T{}'w_T^{-1}(\hat{\theta}_{GMM})G_T\right)^{-1}R'\}\sqrt{T}\left(R\hat{\theta}_{GMM} - r\right)/p$$

where $G_T = S_{ZX}$. Under the assumptions in Hwang and Sun (2017), the following result holds:

$$\mathbb{W}_T^C\left(\hat{\theta}_{GMM}\right) \overset{d}{\to} F_{p,K-p-q+1},$$

where

$$\mathbb{W}_T^C\left(\hat{\theta}_{GMM}\right) = \frac{K-p-q+1}{K}\frac{\mathbb{W}_T\left(\hat{\theta}_{GMM}\right)}{1+\frac{1}{K}J_T\left(\hat{\theta}_{GMM}\right)}$$

is the modified Wald statistic and

$$J_T\left(\hat{\theta}_{GMM}\right) = \frac{K-q+1}{K}\cdot\frac{Tg_T\left(\hat{\theta}_{GMM}\right)'w_T^{-1}\left(\hat{\theta}_{GMM}\right)g_T\left(\hat{\theta}_{GMM}\right)}{q}$$

is the J statistics for testing overidentifying restrictions. Sun and Kim (2012) show that $J_T\left(\hat{\theta}_{GMM}\right)$ converges in distribution to $F(q, K-q+1)$. Following Sun and Kim (2012) and Hwang and Sun (2017), we select K based on the AMSE criterion implemented using the VAR(1) plug-in procedure.

References

Huang, J. and Y. Sun, 2017.Asymptotic F and t Tests in an Efficient GMM Setting. Journal of Econometrics, 198(2): 277-295.

Sun, Y. and M. S. Kim, 2012. Simple and Powerful GMM Over-identification Tests with Accurate Size. Journal of Econometrics, 166(2): 267-281.

Also see

[TS]tsset—Declare data to be time series data.