Title

hart — Test linear hypotheses after har estimation

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

Syntax

Basic syntax

hart coeflist

hart exp=exp[=…]

options	Description
Model	
kernel(string)	set the type of kernels: Bartlett; Parzen; Quadratic Spectral; Orthonormal Series
accumulate	test the hypothesis jointly with previously tested hypotheses
level(#)	set the confidence level; default is level(95)

kernel(string) is required.

time-series operators are allowed.

Syntax 1 tests that the coefficients are 0.

Syntax 2 tests that the linear expressions are equal.

Menu

statistics>Postestimation>Tests> Test linear hypotheses after har estimation

Description

hart performs Wald tests of simple and composite linear hypotheses about the parameters in the most recently estimated har model. See Sun (2013) and Sun (2014).

Options

Model

kernel(string) specifies the type of kernels (Bartlett; Parzen; Quadratic Spectral; Orthonormal Series) to be used in the estimation of the long run covariance matrix.

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

Reporting

level(#); see [R] estimation options.

Remarks and examples

Sun (2013) and Sun (2014) introduce simple and trustworthy inference procedures that are robust to heteroskedasticity and autocorrelation. The HAR variance estimator in Sun (2013) is based on an orthonormal series long run variance matrix estimator. The optimal number K of orthonormal bases is selected by minimizing the type II error of the test subject to a control of the type I error. The tests in Sun (2013) are asymptotic F and t tests in an exact sense: the asymptotic distribution of the adjusted F statistic and t statistic are standard F and t distributions. The HAR variance estimator in Sun (2014) is based on the more conventional kernel long run variance matrix estimators. The F and t tests in Sun (2014) are approximate tests. The asymptotic distributions of the adjusted F and t statistic are not exactly F and t distribution, but they can be well approximated by the standard F and t distributions.

Δ Example 1: Test for a single coefficient against zero

We estimate the following regression:

```
. har usr idle wio,kernel(bartlett)

Regression with HAR standard errors                 Number of obs  =        30
Kernel: Bartlett                                    F(  2,     17)  =     47.66
Data-driven optimal lag: 2                          Prob > F        =    0.0000
```

usr	Coef.	HAR Std.Err.	t	df	P>|t|	[95% Conf.	Interval]
idle	-.6670978	.0715786	-9.32	22	0.000	-.8155428	-.5186529
wio	-.7792461	.11897	-6.55	13	0.000	-1.036265	-.522227
_cons	66.21805	6.984346	9.48	19	0.000	51.59965	80.83646

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

```
. hart idle=0, kernel(bartlett)
       F(  1,     22) =    86.86
            Prob > F =    0.0000
```

The F statistic is 86.86. The degrees of freedom of the approximating F distribution are (1,22). The p-value of the test is 0%. We can 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 idle is -0.6670978 by typing:

```
. hart idle=-0.6670978, kernel(bartlett)
       F(  1,     22) =     0.00
            Prob > F =    1.0000
```

We find that we cannot 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 idle is the same as the coefficient on wio:

```
hart idle=wio, kernel(bartlett)
     F(  1,    13) =    3.20
          Prob > F =    0.0968
```

We find that we cannot reject the equality hypothesis at the 5% level, but we can at the 10% level.

Δ Example 4:

hart allows us to test any linear restrictions. For example,

```
hart 3*idle-2*wio=2*idle-1*wio, kernel(bartlett)
     F(  1,    13) =    3.20
          Prob > F =    0.0968
```

However, hart does not deal with nonlinear restrictions directly. If you attempt to test a nonlinear hypothesis, you will be told that it is not possible.

```
. hart 3*idle/2*wio=2*idle-1*wio, kernel(bartlett)
not possible with test
r(131);
```

This is not a problem specific to the command hart. Stata's command "test" exhibits the same behavior. In fact, hart uses Stata's command "test" to parse the null hypothesis. To test nonlinear restrictions, we have to convert them into linear ones before using hart.

Δ Example 5: Testing joint hypotheses

We wish to test whether idle and wio, taken as a whole, are significant by testing whether the coefficients on idle and wio are simultaneously zero. The command hart allows us to specify multiple conditions to be tested, each embedded within parentheses.

```
. hart (idle=0) (wio=0), kernel(bartlett)
       F(  2,    17) =   47.66
            Prob > F =    0.0000
```

hart displays the set of conditions and reports an F statistic of 47.66. hart also reports the degrees of freedom of the approximating F distribution, which are 2 and 17. The p-value of the test is reported to be around 0. So we can strongly reject the hypothesis of no difference between the two coefficients.

□ Technical note

An alternative method to test simultaneous hypotheses is to specify a test for the first constraint and then accumulate it with the second constraint:

```
. hart idle=0, kernel(bartlett)
       F(  1,    22) =   86.86
            Prob > F =    0.0000

. hart wio=0, kernel(bartlett) acc
       F(  2,    17) =   47.66
            Prob > F =    0.0000
```

We first test the hypothesis that the coefficient on idle is zero by typing hart idle=0. We then test whether the coefficient on wio is also zero by typing hart wio=0, accumulate. The accumulate option tells hart that this is not the start of a new test but a continuation of the previous one. hart responds by showing us the two equations and reporting an F statistic of 47.66. The p-value is about 0%.

Δ Example 6: Quickly testing coefficients against zero

It is very common to test whether some coefficents are jointly zero in applied research. The command hart has a more convenient syntax to accommodate this common case:

```
hart idle wio, kernel(bartlett)
       F(  2,     17) =    47.66
            Prob > F =     0.0000
```

Δ Example 7: Replaying the previous test

We can review our last test by typing hart without arguments.

```
hart , kernel(bartlett)
       F(  2,     17) =    47.66
            Prob > F =     0.0000
```

Δ Example 8: Testing the equality of multiple coefficients

Let's test the hypothesis that idle, wio and syslcl have the same coefficient.

```
. har usr idle wio syslcl,kernel(bartlett)

Regression with HAR standard errors                 Number of obs   =        30
Kernel: Bartlett                                    F(  3,     27)  =    323.51
Data-driven optimal lag: 1                          Prob > F        =    0.0000
```

usr	Coef.	HAR Std.Err.	t	df	P>|t|	[95% Conf.	Interval]
idle	-1.019373	.0311105	-32.77	4	0.000	-1.10575	-.9329965
wio	-1.034228	.0333176	-31.04	54	0.000	-1.101026	-.9674303
syslcl	-1.000682	.0520787	-19.21	12	0.000	-1.114152	-.8872121
_cons	102.003	3.01555	33.83	5	0.000	94.25128	109.7547

```
. hart idle=wio=syslcl, kernel(bartlett)
       F(  2,     18) =     0.65
            Prob > F =     0.5315
```

The syntax idle=wio=syslcl 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 perform the same test by using either of the following

```
hart (idle=wio)(idle=syslcl), kernel(bartlett)
       F(  2,     18) =     0.65
            Prob > F =     0.5315


. hart (idle=wio)(wio=syslcl), kernel(bartlett)
       F(  2,     18) =     0.65
            Prob > F =     0.5315
```

Stored results

hart stores the following results in r():

Scalars

r(firdf) the first degrees of freedom

r(secdf) the second degrees of freedom

r(lag) the data-driven optimal truncation lag

r(kopt) the data-driven optimal K

r(F) the adjusted F statistic

Matrices

r(thetaiv) the IV 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:

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

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 $W_{0T} = S_{ZZ} \in R^{m\times m}$, $\underset{T\to\infty}{Plim}\, W_{0T} = W_0$.

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. Under some standard high-level conditions, we have

$$\sqrt{T}R(\hat{\theta}_{IV} - \theta_0) = \sqrt{T}(R\hat{\theta}_{IV} - r) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t + o_p(1)$$

where $G_0 = ES_{ZX} \in R^{m\times d}$ and $u_t = R(G_0'W_0^{-1}G_0)^{-1}G_0'W_0^{-1}Z_t'e_t$ is the transformed moment process. It then follows that $\sqrt{T}R(\hat{\theta}_{IV} - \theta_0) \xrightarrow{d} N(0,\Omega)$, where $\Omega = \sum_{j=-\infty}^{j=+\infty} E\, u_t u_{t-j}'$ is the long run variance of $\{u_t\}$. The Wald statistic for testing H_0 against H_1 is

$$F_{IV} = \sqrt{T}\left(R\hat{\theta}_{IV} - r\right)'\left(\hat{\Omega}\right)^{-1}\sqrt{T}\left(R\hat{\theta}_{IV} - r\right)/p.$$

Let $G_T = S_{ZX}$, $\hat{u}_t = R(G_T'W_{0T}^{-1}G_T)^{-1}G_T'W_{0T}^{-1}Z_t'$, and $\hat{u}^{ave} = T^{-1}\sum_{s=1}^{T}\hat{u}_s$. We consider the estimator $\widehat{\Omega}$ of the form

$$\widehat{\Omega} = \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T} Q_h\left(\frac{s}{T},\frac{t}{T}\right)(\hat{u}_t - \hat{u}^{ave})(\hat{u}_s - \hat{u}^{ave})',$$

where $Q_h(r,s)$ is a weighting function, and h is the smoothing parameter. The above estimator includes the kernel HAR variance estimators and the orthonormal series HAR variance estimator as special cases.

For the kernel HAR variance estimator, we let $Q_h(r,s) = k((r-s)/b)$ and $h = 1/b$ for a kernel function $k(\cdot)$ with $\mathrm{M}_T = bT$ being the so-called truncation lag. Define

$$c_1 = \int_{-\infty}^{+\infty} k(x)dx,\ c_2 = \int_{-\infty}^{+\infty} k^2(x)dx.$$

For the Bartlett kernel, $c_1 = 1,\ c_2 = 2/3$; For the Parzen kernel, $c_1 = 3/4,\ c_2 = 0.539285$; For the Quadratic Spectral kernel, $c_1 = 1.25,\ c_2 = 1$. Let

$$K = \max\left(\left\lceil\frac{1}{bc_2}\right\rceil, p\right) - p + 1,$$

where $\lceil\cdot\rceil$ is the ceiling function, and

$$\kappa = \frac{1}{2}(\exp(b[c_1 + (p-1)c_2]) + (1 + b[c_1 + (p-1)c_2])).$$

Based on the kernel estimator $\hat{\Omega}$, Sun (2014) shows that

$$P\left(F_{IV} > \kappa F_{p,K}^{\alpha}\right) = \alpha + o(b) + O((bT)^{-q}) + O\left(\frac{logT}{\sqrt{T}}\right),$$

where $F_{p,K}^{\alpha}$ is the $100(1-\alpha)\%$ quantile of the standard $F_{p,K}$ distribution.

Sun (2014) obtains the testing-optimal bandwidth b_{opt}:

$$b_{opt} = \begin{cases} \left[\dfrac{2qG'_{p,\delta^2}\left(\chi_p^{1-\alpha}\right)|\bar{B}|}{\delta^2 G'_{(p+2),\delta^2}\left(\chi_p^{1-\alpha}\right)c_2}\right]^{\frac{1}{q+1}} T^{\frac{-q}{q+1}}, \bar{B} > 0 \\ \left[\dfrac{G'_p\left(\chi_p^{1-\alpha}\right)\chi_p^{1-\alpha}|\bar{B}|}{(\tau-1)\alpha}\right]^{\frac{1}{q}} \dfrac{1}{T}, \bar{B} \le 0 \end{cases}$$

where $\tau > 1$ is the permitted tolerance, $\mathrm{G}'_{p,\delta^2}(z)$ is the pdf of the noncentral χ^2 distribution with degrees of freedom p and noncentrality parameter δ^2. In the above formula, $\chi_p^{1-\alpha}$ is the $1-\alpha$ quantile of the χ^2 distribution with p degrees of freedom and δ^2 is chosen to satisfy $P\{\varkappa > \chi_p^{1-\alpha}\} = 75\%$, where $\varkappa \sim \chi_p^2(\delta^2)$. In addition,

$$\bar{B} = tr(B\Omega^{-1})/p, B = -\rho_q \sum_{j=-\infty}^{\infty} |j|^q\, Eu_t u_{t-j}'$$

where q is the order of the kernel used, and ρ_q is the Parzen characteristic exponent of the kernel. For the Bartlett kernel, $q = 1, \rho_q = 1$; For the Parzen kernel, $q = 2, \rho_q = 6$; For the QS kernel, $q = 2, \rho_q = 1.421223$.

The parameter $\bar{B}$ is estimated by a standard VAR(1) plug-in procedure. This is what we opt for in the new command. Plugging the estimate of B into b_{opt} yields b_{temp}. The data-driven choice of b is then given by $\hat{b}_{opt} = \min(b_{temp}, 0.5)$. We do not use a b larger than 0.5 in order to avoid large power loss.

For the OS HAR variance estimator, we let $Q_h(r,s) = K^{-1}\sum_{j=1}^{K} \phi_j(r)\phi_j(s)$, and $h = K$, where $\{\phi_j(\cdot)\}_{j=1}^{K}$ are orthonormal basis functions on $L^2[0,1]$ satisfying $\int_0^1 \phi_j(r)\, dr = 0$ for $j = 1,2,\cdots,K$. Sun (2013) shows that the usual Wald statistic F_{IV} satisfies

$$\frac{K-p+1}{K} F_{IV} \xrightarrow{d} F_{p,K-p+1} = \frac{\chi_p^2/p}{\chi_{K-p+1}^2/(K-p+1)}$$

where $F_{p,K-p+1}$ is the F distribution with degrees of freedom$(p, K-p+1)$.

Sun (2013) obtains the testing-optimal K_{opt} as follows

$$K_{opt} = \begin{cases} \left[\dfrac{\delta^2 G'_{(p+2),\delta^2}\left(\chi_p^{1-\alpha}\right)}{4G'_{p,\delta^2}\left(\chi_p^{1-\alpha}\right)|\bar{B}|}\right]^{\frac{1}{3}} T^{\frac{2}{3}}, \bar{B} > 0 \\ \left[\dfrac{(\tau-1)\alpha}{G'_p\left(\chi_p^{1-\alpha}\right)\chi_p^{1-\alpha}|\bar{B}|}\right]^{\frac{1}{2}} T, \bar{B} \le 0 \end{cases}$$

As before, the parameter $\bar{B}$ is estimated by a standard VAR(1) plug-in procedure. Plugging the estimate of $\bar{B}$ into K_{opt} yields $\hat{K}_{temp}$.We truncate $\hat{K}_{temp}$ to be between $p+4$ and T. That is, we take

$$\tilde{K}_{temp} = \begin{cases} p+4, & if\ \hat{K}_{temp} \le p+4 \\ \hat{K}_{temp}, & if\ \hat{K}_{temp} \in (p+4, T] \\ T, & if\ \hat{K}_{temp} > T \end{cases}$$

Imposing the lower bound $p+4$ ensures that the variance of the approximating distribution $F_{p,K-p+1}$ is finite and that power loss is not very large. Finally, we round $\tilde{K}_{temp}$to the greatest even number less than $\tilde{K}_{temp}$. We take this greatest even number, denoted by $\hat{K}_{opt}$to be our data-driven and testing-optimal choice for K.

References

Sun, Y., 2013. A heteroskedasticity and autocorrelation robust F test using an orthonormal series variance estimator. Econometrics Journal, 16: 1-26.

Sun, Y., 2014. Let’s fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust inference. Journal of Econometrics, 178: 660-677.

Also see

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