**Title**

**har — Regression with HAR standard errors**

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

**Syntax**

har depvar [varlist1] (varlist2=inst list)[if] [in], kernel(string) [, noconstant level(#) ]

options Description

Model

kernel(string) set the type of kernels: Bartlett; Parzen; Quadratic Spectral; Orthonormal Series

noconstant suppress constant term

Reporting

level(#) set confidence level; default is level(95)

kernel(string ) is required.

you must tsset your data before using har; see [TS] tsset.

time-series operators are allowed.

**Menu**

statistics>Time series >Regression with HAR standard errors

**Description**

The command estimates an IV regression in the presence of heteroskedasticity and autocorrelation. Inferences are based on the fixed-smoothing asymptotics. The command is based on Sun (2013) and Sun (2014). Sun (2013) develops heteroskedasticity and autocorrelation robust F and t tests using an orthonormal series long run variance matrix estimator. The number of orthonormal bases is selected by minimizing the type II error of the associated test while controlling for its type I error. Sun (2014) introduces a new and easy-to-use asymptotic F test (and t test) based on kernel long run variance matrix estimators. The proposed bandwidth selection rule is testing-optimal and is similar to Sun (2013). The command allows for three types of kernels: the Bartlett, Parzen, and Quadratic Spectral kernels.

For the sake of simplicity, we refer to the orthonormal series long run variance estimator in Sun (2013) as the kernel long run variance estimator with kernel “Orthonormal Series.” In total, the command allows for four types of kernels: Bartlett, Parzen, Quadratic Spectral, and Orthonormal Series.

**Options**

Model

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

noconstant; see [R] estimation options.

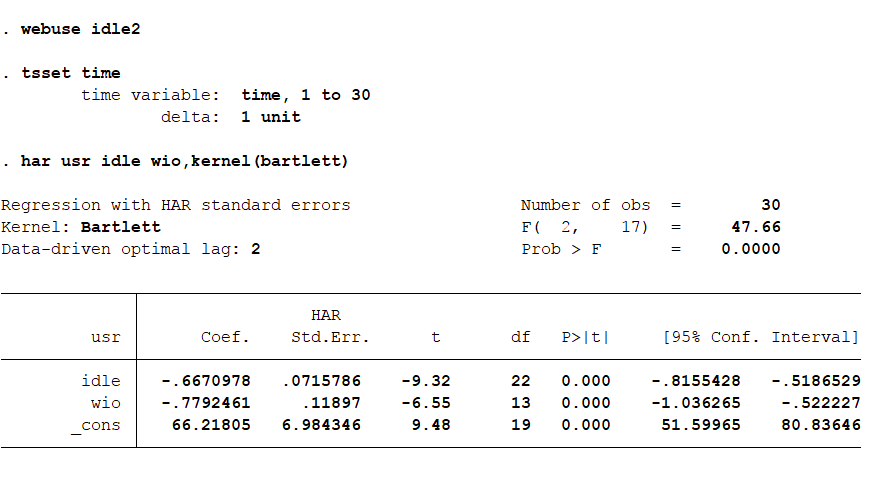
Reporting

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

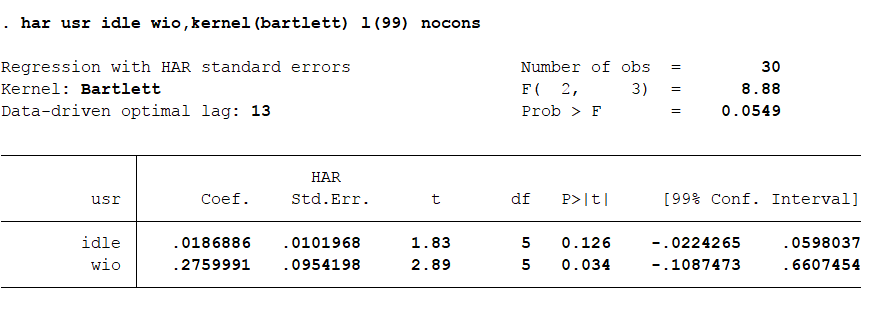
**Remarks and examples**

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 associated test subject to a control of its type I error. The tests proposed by Sun (2013) are asymptotic F and t tests in an exact sense: the asymptotic distributions 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 distributions, but they can be well approximated by the standard F and t distributions.

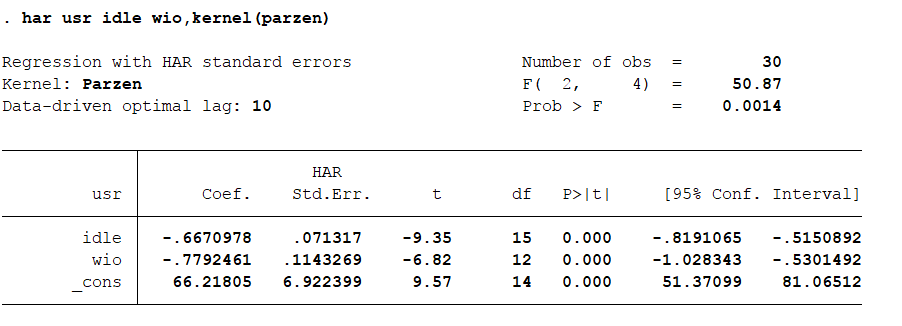
**Example 1**



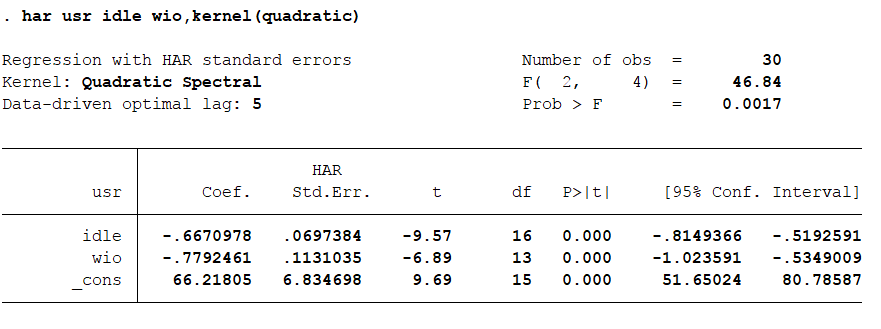
**Example 2**



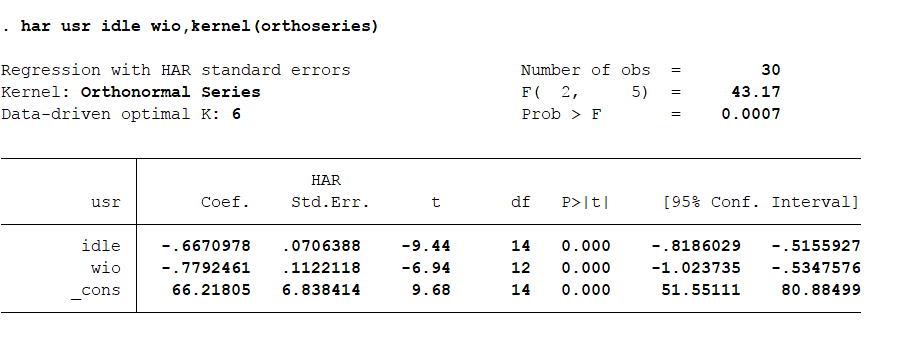
**Example 3**



**Example 4**



**Example 5**



**Stored results**

The command har uses ivregress command to get the estimated coefficient. So, in addition to the standard stored results from ivregress, har also stores the following results in e():

Scalars

e(N) the number of observations

\*# e(fdf) the first value of degrees of freedom

\*e(sF) the adjusted F statistic

\*e(ssdf) the second value of degrees of freedom

\*e(kopt) the data-driven optimal number k of orthonormal bases

#e(kF) the adjusted F statistic

#e(ksdf) the second degrees of freedom

#e(lag) the data-driven optimal truncation lag

Macros

e(cmd) har

e(cmdline) command as typed

e(varline) variable line as typed

e(carg) nocons or " " if specified

e(depvar) name of dependent variable

e(title) title in estimation output

e(vcetype) title used to label Std. Err.

e(kerneltype) kernel in the estimation

Matrices

e(b) coefficient vector

\*e(sstderr) the adjusted std error for each individual coefficient

\*e(sdf) the degrees of freedom of the approximating t distribution

\*e(st) the t statistic

\*e(sbetahat) the IV coefficient vector

#e(kbetahat) the IV coefficient vector

#e(kstderr) the adjusted std error for each individual coefficient

#e(kdf) the degrees of freedom of the approximating t distribution

#e(kt) the t statistic

Functions

e(sample) marks the estimation sample

note：for the other results in ereturn list see[R] ivregress;

\* only for Orthonormal Series;

# only for Bartlett, Parzen, Quadratic Spectral.

**Methods and formulas**

Consider the regression model:

where is a zero-mean process that may be correlated with the covariate process

There are instruments such that the moment conditions:

hold if and only if . We allow the process to have the autocorrelation of unknown forms. The model may be over-identified with the degree of over-identification .

Define:

, ,

Then the IV estimator of is

where , .

We are interested in testing the null against the alternative , where and 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

where and is the transformed moment

process. It then follows that , where is the long run variance of . The Wald statistic for testing against is

Let and. We consider the estimator of the form

where is a weighting function, and 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 and for a kernel function with being the so-called truncation lag. Define

, .

For the Bartlett kernel, , ; For the Parzen kernel, , For the Quadratic Spectral kernel,, . Let

,

where is the ceiling function, and

.

Based on the kernel estimator , Sun (2014) shows that

where is the 100 quantile of the standard distribution

Sun (2014) obtains the testing-optimal bandwidth :

where is the permitted tolerance,is the pdf of the noncentral distribution with degrees of freedom and noncentrality parameter . In the above formula, is the quantile of the distribution with degrees of freedom and is chosen to satisfy, where . In addition,

where *q* is the order of the kernel used, and is the Parzen characteristic exponent of the kernel. For the Bartlett kernel, ; For the Parzen kernel, ; For the QS kernel, .

The parameter is estimated by a standard VAR(1) plug-in procedure. This is what we opt for in the new command. Plugging the estimate of into yields The data-driven choice of is then given by 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 , and ,

where are orthonormal basis functions on satisfying for

. Sun (2013) shows that the usual Wald statistic satisfies

where is the distribution with degrees of freedom).

Sun (2013) obtains the testing-optimal as follows

As before, the parameter is estimated by a standard VAR(1) plug-in procedure. Plugging the estimate of into yields .We truncate to be between and *T*. That is, we take

Imposing the lower bound ensures that the variance of the approximating distribution is finite and that power loss is not very large. Finally, we round to the greatest even number less than . We take this greatest even number, denoted by 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.