User's Guide : **Advanced Single Equation Analysis** : **Cointegrating Regression** : **Estimating a Cointegrating Regression**

Estimating A Cointegrating Regression

Specifying the Equation

Equation Specification

Cointegrating Regressors Specification

Specifying an Estimation Method

Fully Modified OLS

Canonical Cointegrating Regression

Dynamic OLS

EViews offers three methods for estimating a single cointegrating vector: Fully Modified OLS (FMOLS), Canonical Cointegrating Regression (CCR), and Dynamic OLS (DOLS). Static OLS is supported as a special case of DOLS. We emphasize again that Johansen's (1991, 1995) system maximum likelihood approach is discussed in "Vector Autoregression and Error Correction Models".


The equation object is used to estimate a cointegrating equation. First, create an equation object, select **Object/New Object…/Equation** or **Quick/Estimate Equation…** then select **COINTREG - Cointegrating Regression** in the **Method** dropdown menu. The dialog will show settings appropriate

for your cointegrating regression. Alternately, you may enter the `cointreg` keyword in the command window to perform both steps.

There are three parts to specifying your equation. First, you should use the first two sections of the dialog (**Equation specification** and **Cointegrating regressors specification**) to specify your triangular system of equations. Second, you will use the **Nonstationary estimation settings** section to specify the basic cointegrating regression estimation method. Lastly, you should enter a sample specification, then click on **OK** to estimate the equation. (We ignore, for a moment, the options settings on the **Options** tab.)

Specifying The Equation

The first two sections of the dialog (**Equation specification** and **Cointegrating regressors specification**) are used to describe your cointegrating and regressors equations.
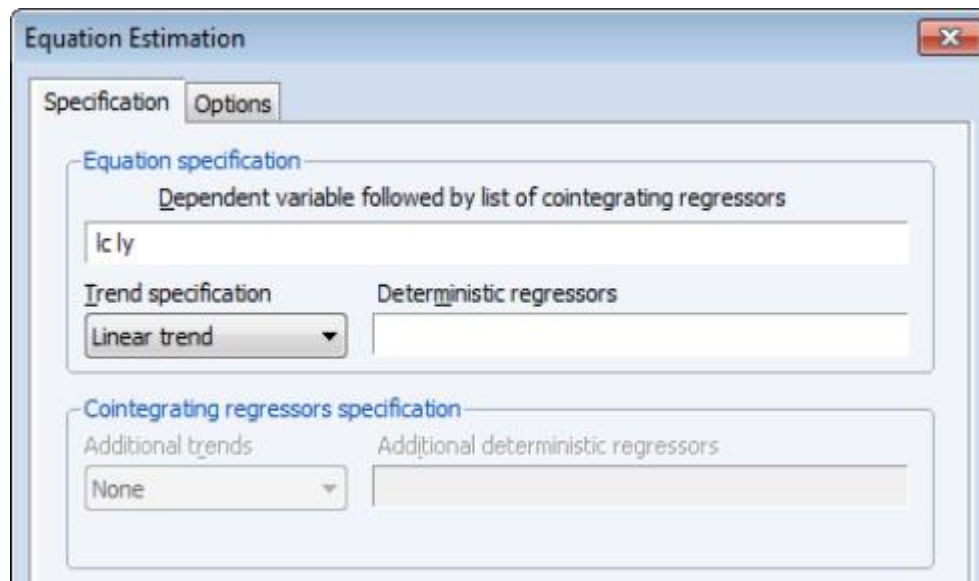
Equation Specification

The cointegrating equation is described in the **Equation specification** section. You should enter the name of the dependent variable, y, followed by a list of cointegrating regressors, X, in the edit field, then use the **Trend specification** dropdown to choose from a list of deterministic trend variable assumptions (**None**, **Constant (Level)**, **Linear Trend**, **Quadratic Trend**). The dropdown menu selections imply trends up to the specified order so that the **Quadratic Trend** selection depicted includes a constant and a linear trend term along with the quadratic.

If you wish to add deterministic regressors that are not offered in the pre-specified list to D_1, you may enter the series names in the **Deterministic regressors** edit box.

Cointegrating Regressors Specification

Cointegrating Regressors Specification section of the dialog completes the specification of the regressors equations.

First, if there are any D_2 deterministic regressors (regressors that are included in the regressors equations but not in the cointegrating equation), they should be specified here using the **Additional trends** dropdown menu or by entering regressors explicitly using the **Additional deterministic regressors** edit field.

Second, you should indicate whether you wish to estimate the regressors innovations u_{2t} indirectly by estimating the regressors equations in levels and then differencing the residuals or directly by estimating the regressors equations in differences. Check the box for **Estimate using differenced data** (which is only relevant and only appears if you are estimating your equation using FMOLS or CCR) to estimate the regressors equations in differences.

Specifying An Estimation Method

Once you specify your cointegrating and regressor equations you are ready to describe your estimation method. The EViews equation object offers three methods for estimating a single cointegrating vector: Fully Modified OLS (FMOLS), Canonical Cointegrating Regression (CCR), and Dynamic OLS (DOLS). We again emphasize that Johansen's (1991, 1995) system maximum likelihood approach is described elsewhere("Vector Error Correction (VEC) Models").

The **Nonstationary estimation settings** section is used to describe your estimation method. First, you should use the **Method** dropdown menu to choose one of the three methods. Both the main dialog page and the options page will change to display the options associated with your selection.

Fully Modified OLS

Phillips and Hansen (1990) propose an estimator which employs a semi-parametric correction to eliminate the problems caused by the long run correlation between the cointegrating equation and stochastic regressors innovations. The resulting Fully Modified OLS (FMOLS) estimator is asymptotically unbiased and has fully efficient mixture normal asymptotics allowing for standard Wald tests using asymptotic Chi-square statistical inference.

The FMOLS estimator employs preliminary estimates of the symmetric and one-sided long-run covariance matrices of the residuals. Let u_{1t} be the residuals obtained after estimating

Equation (27.1). The u_{2t} may be obtained indirectly as $u_{2t} = \Delta \hat{\epsilon}_{2t}$ from the levels regressions

$$X_t = \hat{\Gamma}_{21}' D_{1t} + \hat{\Gamma}_{22}' D_{2t} + \hat{\epsilon}_{2t} \tag{27.4}$$

or directly from the difference regressions

$$\Delta X_t = \hat{\Gamma}_{21}' \Delta D_{1t} + \hat{\Gamma}_{22}' \Delta D_{2t} + u_{2t} \tag{27.5}$$

Let $\hat{\Omega}$ and $\hat{\Lambda}$ be the long-run covariance matrices computed using the residuals $u_t = (u_{1t}, u_{2t}')'$. Then we may define the modified data

$$y_t^+ = y_t - \hat{\omega}_{12} \hat{\Omega}_{22}^{-1} u_2 \tag{27.6}$$

and an estimated bias correction term

$$\hat{\lambda}_{12}^+ = \hat{\lambda}_{12} - \hat{\omega}_{12} \hat{\Omega}_{22}^{-1} \hat{\Lambda}_{22} \tag{27.7}$$

The FMOLS estimator is given by

$$\hat{\theta} = \begin{bmatrix} \beta \\ \hat{\gamma}_1 \end{bmatrix} = \left(\sum_{t=2}^{T} Z_t Z_t' \right)^{-1} \left(\sum_{t=2}^{T} Z_t y_t^+ - T \begin{bmatrix} \hat{\lambda}_{12}^+{}' \\ 0 \end{bmatrix} \right) \tag{27.8}$$

where $Z_t = (X_t', D_t')'$. The key to FMOLS estimation is the construction of long-run covariance matrix estimators $\hat{\Omega}$ and $\hat{\Lambda}$.

Before describing the options available for computing $\hat{\Omega}$ and $\hat{\Lambda}$, it will be useful to define the scalar estimator

$$\hat{\omega}_{1.2} = \hat{\omega}_{11} - \hat{\omega}_{12} \hat{\Omega}_{22}^{-1} \hat{\omega}_{21} \tag{27.9}$$

which may be interpreted as the estimated long-run variance of u_{1t} conditional on u_{2t}. We may, if desired, apply a degree-of-freedom correction to $\hat{\omega}_{1.2}$.
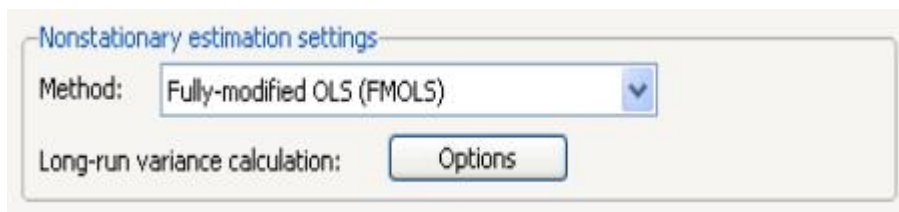
Hansen (1992) shows that the Wald statistic for the null hypothesis $R\theta = r$

$$W = (R\theta - r)'(R V(\theta)R')^{-1}(R\theta - r) \tag{27.10}$$

with

$$V(\theta) = \hat{\omega}_{1.2} \left(\sum_{t=2}^{T} Z_t Z_t' \right)^{-1} \tag{27.11}$$

has an asymptotic χ_g^2-distribution, where g is the number of restrictions imposed by R. (You should bear in mind that restrictions on the constant term and any other non-trending variables are not testable using the theory underlying Equation (27.10).)


To estimate your equation using FMOLS, select **Fully-modified OLS (FMOLS)** in the **Nonstationary estimation settings** dropdown menu. The main dialog and options pages will change to show the available settings.

To illustrate the FMOLS estimator, we employ data for (100 times) log real quarterly aggregate personal disposable income (LY) and personal consumption expenditures (LC) for the U.S. from 1947q1 to 1989q3 as described in Hamilton (2000, p. 600, 610) and contained in the workfile "Hamilton_coint.WF1".

We wish to estimate a model that includes an intercept in the cointegrating equation, has no additional deterministics in the regressors equations, and estimates the regressors equations in non-differenced form.

By default, EViews will estimate Ω and Λ using a (non-prewhitened) kernel approach with a Bartlett kernel and Newey-West fixed bandwidth. To change the whitening or kernel settings, click on the **Long-run variance calculation: Options** button and enter your changes in the sub-dialog.
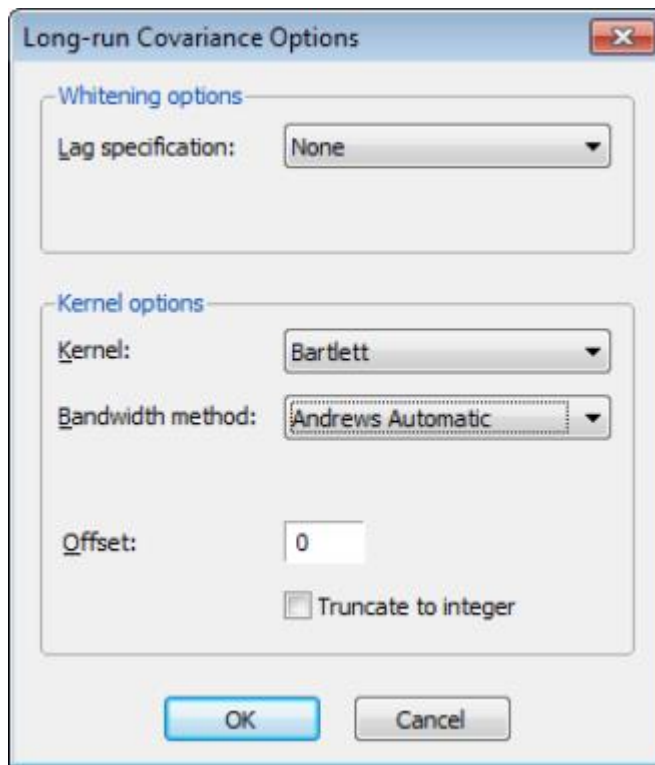
Here we have specified that the long-run variances be computed using a nonparametric method with the Bartlett kernel and a real-valued bandwidth chosen by Andrews' automatic bandwidth selection method.

In addition, you may use the **Options** tab of the **Equation Estimation** dialog to modify the computation of the coefficient covariance. By default, EViews computes the coefficient covariance by rescaling the usual OLS covariances using the $\hat{\omega}_{1.2}$ obtained from the estimated $\hat{\Omega}$ after applying a degrees-of-freedom correction. In our example, we will use the checkbox on the **Options** tab (not depicted) to remove the d.f. correction.

The estimates for this specification are given by:

Dependent Variable: LC
Method: Fully Modified Least Squares (FMOLS)
Date: 08/11/09   Time: 13:19
Sample (adjusted): 1947Q2 1989Q3
Included observations: 170 after adjustments
Cointegrating equation deterministics: C
Long-run covariance estimate (Bartlett kernel, Andrews bandwidth = 14.9878)
No d.f. adjustment for standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LY	0.987548	0.009188	107.4880	0.0000
C	-0.035023	6.715362	-0.005215	0.9958

R-squared	0.998171	Mean dependent var	720.5078
Adjusted R-squared	0.998160	S.D. dependent var	41.74069
S.E. of regression	1.790506	Sum squared resid	538.5929
Durbin-Watson stat	0.406259	Long-run variance	25.46653

The top portion of the results describe the settings used in estimation, in particular, the specification of the deterministic regressors in the cointegrating equation, the kernel nonparametric method used to compute the long-run variance estimators $\hat{\Omega}$ and $\hat{\Lambda}$, and the no-d.f. correction option used in the

calculation of the coefficient covariance. Also displayed is the bandwidth of 14.9878 selected by the Andrews automatic bandwidth procedure.

The estimated coefficients are presented in the middle of the output. Of central importance is the coefficient on LY which implies that the estimated cointegrating vector for LC and LY (1, -0.9875). Note that we present the standard error, *t*-statistic, and *p*-value for the constant even though they are not, strictly speaking, valid.

The summary statistic portion of the output is relatively familiar but does require a bit of comment. First, all of the descriptive and fit statistics are computed using the original data, not the FMOLS transformed data. Thus, while the measures of fit and the Durbin-Watson stat may be of casual interest, you should exercise extreme caution in using these measures. Second, EViews displays a "Long-run variance" value which is an estimate of the long-run variance of u_{1t} conditional on u_{2t}. This statistic, which takes the value of 25.47 in this example, is the $\hat{\omega}_{1.2}$ employed in forming the coefficient covariances, and is obtained from the $\hat{\Omega}$ and $\hat{\Lambda}$ used in estimation. Since we are not d.f. correcting the coefficient covariance matrix the $\hat{\omega}_{1.2}$ reported here is not d.f. corrected.

Once you have estimated your equation using FMOLS you may use the various cointegrating regression equation views and procedures. We will discuss these tools in greater depth in ("Working with an Equation"), but for now we focus on a simple Wald test for the coefficients. To test for whether the cointegrating vector is (1, -1), select **View/Coefficient Diagnostics/Wald Test - Coefficient Restrictions** and enter "C(1)=1" in the dialog. EViews displays the output for the test:

Wald Test:
Equation: FMOLS
Null Hypothesis: C(1)=1

Test Statistic	Value	df	Probability
t-statistic	-1.355362	168	0.1771
F-statistic	1.837006	(1, 168)	0.1771
Chi-square	1.837006	1	0.1753

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
-1 + C(1)	-0.012452	0.009188

Restrictions are linear in coefficients.

The *t*-statistic and Chi-square *p*-values are both around 0.17, indicating that the we cannot reject the null hypothesis that the cointegrating regressor coefficient value is equal to 1.

Note that this Wald test is for a simple linear restriction. Hansen points out that his theoretical results do not directly extend to testing nonlinear hypotheses in models with trend regressors, but EViews does allow tests with nonlinear restrictions since others, such as Phillips and Loretan (1991) and Park (1992) provide results in the absence of the trend regressors. We do urge caution in interpreting nonlinear restriction test results for equations involving such regressors.

Canonical Cointegrating Regression

Park's (1992) Canonical Cointegrating Regression (CCR) is closely related to FMOLS, but instead employs stationary transformations of the (y_{1t}, X_t') data to obtain least squares estimates to remove the long run dependence between the cointegrating equation and stochastic regressors innovations. Like FMOLS, CCR estimates follow a mixture normal distribution which is free of non-scalar nuisance parameters and permits asymptotic Chi-square testing.

As in FMOLS, the first step in CCR is to obtain estimates of the innovations $u_t = (u_{1t}, u_{2t}')'$ and corresponding consistent estimates of the long-run covariance matrices $\hat{\Omega}$ and $\hat{\Lambda}$. Unlike FMOLS, CCR also requires a consistent estimator of the contemporaneous covariance matrix $\hat{\Sigma}$.

Following Park, we extract the columns of $\hat{\Lambda}$ corresponding to the one-sided long-run covariance matrix of u_t and (the levels and lags of) u_{2t}

$$\hat{\Lambda}_2 = \begin{bmatrix} \hat{\lambda}_{12} \\ \hat{\Lambda}_{22} \end{bmatrix} \tag{27.12}$$

and transform the (y_{1t}, X_t') using

$$
\begin{aligned}
X_t^* &= X_t - (\hat{\Sigma}^{-1}\hat{\Lambda}_2)'\, u_t \\
y_t^* &= y_t - \left(\hat{\Sigma}^{-1}\hat{\Lambda}_2 \tilde{\beta} + \begin{bmatrix} 0 \\ \hat{\Omega}_{22}^{-1} \hat{\omega}_{21} \end{bmatrix} \right)' u_t
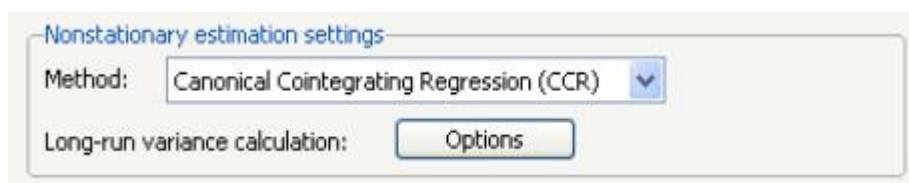\end{aligned}
\tag{27.13}
$$

where the $\tilde{\beta}$ are estimates of the cointegrating equation coefficients, typically the SOLS estimates used to obtain the residuals u_{1t}.

The CCR estimator is defined as ordinary least squares applied to the transformed data

$$\begin{bmatrix} \beta \\ \hat{\gamma}_1 \end{bmatrix} = \left(\sum_{t=1}^{T} Z_t^* Z_t^{*'} \right)^{-1} \sum_{t=1}^{T} Z_t^* y_t^* \tag{27.14}$$

where $Z_t^* = (Z_t^{*'}, D_{1t}')'$.

Park shows that the CCR transformations asymptotically eliminate the endogeneity caused by the long run correlation of the cointegrating equation errors and the stochastic regressors innovations, and simultaneously correct for asymptotic bias resulting from the contemporaneous correlation between the regression and stochastic regressor errors. Estimates based on the CCR are therefore fully efficient and have the same unbiased, mixture normal asymptotics as FMOLS. Wald testing may be carried out as in Equation (27.10) with Z_t^* used in place of Z_t in Equation (27.11).

Nonstationary estimation settings
Method: Canonical Cointegrating Regression (CCR) ▾
Long-run variance calculation: Options

To estimate your equation using CCR, select **Canonical Cointegrating Regression (CCR)** in the **Nonstationary estimation settings** dropdown menu. The main dialog and options pages for CCR are identical to those for FMOLS.

To continue with our consumption and disposable income example, suppose we wish to estimate the same specification as before by CCR, using prewhitened Quadratic-spectral kernel estimators of the long-run covariance matrices. Fill out the equation specification portion of the dialog as before, then click on the **Long-run variance calculation: Options** button to change the calculation method. Here, we have specified a (fixed lag) VAR(1) for the prewhitening method and have changed our kernel shape to quadratic spectral. Click on **OK** to accept the covariance options

Once again go to the **Options** tab to turn off d.f. correction for the coefficient covariances so that they match those from FMOLS. Click on **OK** again to accept the estimation options.

The results are presented below:

Dependent Variable: LC
Method: Canonical Cointegrating Regression (CCR)
Date: 08/11/09   Time: 13:25
Sample (adjusted): 1947Q2 1989Q3
Included observations: 170 after adjustments
Cointegrating equation deterministics: C
Long-run covariance estimate (Prewhitening with lags = 1, Quadratic
    -Spectral kernel, Andrews bandwidth = 1.5911)
No d.f. adjustment for standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LY	0.988975	0.007256	136.3069	0.0000
C	-1.958828	5.298819	-0.369673	0.7121

R-squared	0.997780	Mean dependent var	720.5078
Adjusted R-squared	0.997767	S.D. dependent var	41.74069
S.E. of regression	1.972481	Sum squared resid	653.6343
Durbin-Watson stat	0.335455	Long-run variance	15.91571

The first thing we note is that the VAR prewhitening has a strong effect on the kernel part of the calculation of the long-run covariances, shortening the Andrews optimal bandwidth from almost 15 down to 1.6. Furthermore, as a result of prewhitening, the estimate of the conditional long-run variance changes quite a bit, decreasing from 25.47 to 15.92. This decrease contributes to estimated coefficient standard errors for CCR that are smaller than their FMOLS counterparts. Differences aside, however, the

estimates of the cointegrating vector are qualitatively similar. In particular, a Wald test of the null hypothesis that the cointegrating vector is equal to (1, -1) yields a *p*-value of 0.1305.

Dynamic OLS

A simple approach to constructing an asymptotically efficient estimator that eliminates the feedback in the cointegrating system has been advocated by Saikkonen (1992) and Stock and Watson (1993). Termed Dynamic OLS (DOLS), the method involves augmenting the cointegrating regression with lags *and leads* of ΔX_t so that the resulting cointegrating equation error term is orthogonal to the entire history of the stochastic regressor innovations:

$$y_t = X_t'\beta + D_{1t}'\gamma_1 + \sum_{j=-q}^{r} \Delta X_{t+j}'\delta + v_{1t} \tag{27.15}$$

Under the assumption that adding q lags and r leads of the differenced regressors soaks up all of the long-run correlation between u_{1t} and u_{2t}, least-squares estimates of $\theta = (\beta', \gamma')'$ using Equation (27.15) have the same asymptotic distribution as those obtained from FMOLS and CCR.

An estimator of the asymptotic variance matrix of $\hat{\theta}$ may be computed by computing the usual OLS coefficient covariance, but replacing the usual estimator for the residual variance of v_{1t} with an estimator of the long-run variance of the residuals. Alternately, you could compute a robust HAC estimator of the coefficient covariance matrix.

To estimate your equation using DOLS, first fill out the equation specification, then select **Dynamic OLS (DOLS)** in the **Nonstationary estimation settings** dropdown menu. The dialog will change to display settings for DOLS.

By default, the **Lag & lead method** is **Fixed** with **Lags** and **Leads** each set to 1. You may specify a different number of lags or leads or you can use the dropdown to elect automatic information criterion selection of the lag and lead orders by selecting **Akaike**, **Schwarz**, or **Hannan-Quinn**. If you select **None**, EViews will estimate SOLS.

If you select one of the info criterion selection methods, you will be prompted for a maximum lag and lead length. You may enter a value, or you may retain the default entry "*" which instructs EViews to use an arbitrary observation-based rule-of-thumb:

$$\text{int}(\min((T-k)/3, 12) \cdot (T/100)^{1/4}) \tag{27.16}$$

to set the maximum, where k is the number of coefficients in the cointegrating equation. This rule-of-thumb is a slightly modified version of the rule suggested by Schwert (1989) in the context of unit root testing. (We urge careful thought in the use of automatic selection methods since the purpose of including leads and lags is to remove long-run dependence by orthogonalizing the equation residual with respect to the history of stochastic regressor innovations; the automatic methods were not designed to produce this effect.)

For DOLS estimation we may also specify the method used to compute the coefficient covariance matrix. Click on the **Options** tab of the dialog to see the relevant options.
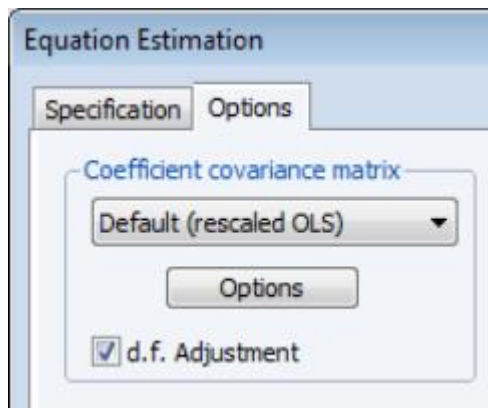
The dropdown menu allows you to choose between the **Default (rescaled OLS)**, **Ordinary Least Squares**, **White**, or **HAC - Newey West.** The default computation method re-scales the ordinary least squares coefficient covariance using an estimator of the long-run variance of DOLS residuals (multiplying by the ratio of the long-run variance to the ordinary squared standard error). Alternately, you may employ a sandwich-style **HAC (Newey-West)** covariance matrix estimator. In both cases, the **HAC Options** button may be used to override the default method for computing the long-run variance (non-prewhitened Bartlett kernel and a Newey-West fixed bandwidth). In addition, EViews offers options for estimating the coefficient covariance using the **White** covariance or **Ordinary Least Squares** methods. These methods are offered primarily for comparison purposes.

Lastly, the **Options** tab may be used to remove the degree-of-freedom correction that is applied to the estimate of the conditional long-run variance or robust coefficient covariance.

We illustrate the technique by estimating an example from Hamilton (19.3.31, p. 611) using the consumption and income data discussed earlier. The model employs an intercept-trend specification for the cointegrating equation, with no additional deterministics in the regressors equations, and four lags and leads of the differenced cointegrating regressor to eliminate long run correlation between the innovations.

Here, we have entered the cointegrating equation specification in the top portion of the dialog, and chosen **Dynamic OLS (DOLS)** as our estimation method, and specified a **Fixed** lag and lead length of 4.

In computing the covariance matrix, Hamilton computes the long-run variance of the residuals using an AR(2) whitening regression with no d.f. correction. To match Hamilton's computations, we click on the **Options** tab to display the covariance. First, turn off the adjustment for degrees of freedom by unchecking the **d.f. Adjustment** box. Next, with the dropdown set to **Default (rescaled OLS)**, click on the **HAC Options** button to display the **Long-run Variance Options** dialog. Select a **Fixed** lag specification of 2, and choose the **None** kernel. Click on **OK** to accept the HAC settings, then on **OK** again to estimate the equation.

The estimation results are given below:

```
Dependent Variable: LC
Method: Dynamic Least Squares (DOLS)
Date: 08/11/09   Time: 13:37
Sample (adjusted): 1948Q2 1988Q3
Included observations: 162 after adjustments
Cointegrating equation deterministics: C @TREND
Fixed leads and lags specification (lead=4, lag=4)
Long-run variance estimate (Prewhitening with lags =2, None kernel)
No d.f. adjustment for standard errors & covariance
```

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LY	0.681179	0.071981	9.463267	0.0000
C	199.1406	47.20878	4.218297	0.0000
@TREND	0.268957	0.062004	4.337740	0.0000

R-squared	0.999395	Mean dependent var	720.5532
Adjusted R-squared	0.999351	S.D. dependent var	39.92349
S.E. of regression	1.017016	Sum squared resid	155.1484
Durbin-Watson stat	0.422921	Long-run variance	10.19830

The top portion describes the settings used in estimation, showing the trend assumptions, the lag and lead specification, and method for computing the long-run variance used in forming the coefficient covariances. The actual estimate of the latter, in this case 10.198, is again displayed in the bottom portion of the output (if you had selected OLS as your coefficient covariance methods, this value would be simply be the ordinary S.E. of the regression; if you had selected White or HAC, the statistic would not have been computed).

The estimated coefficients are displayed in the middle of the output. First, note that EViews does not display the results for the lags and leads of the differenced cointegrating regressors since we cannot perform inference on these short-term dynamics nuisance parameters. Second, the coefficient on the linear trend is statistically different from zero at conventional levels, indicating that there is a deterministic time trend common to both LC and LY. Lastly, the estimated cointegrating vector for LC and LY is (1, -0.6812), which differs qualitatively from the earlier results. A Wald test of the restriction that the cointegrating vector is (1, -1) yields a *t*-statistic of -4.429, strongly rejecting that null hypothesis.

While EViews does not display the coefficients for the short-run dynamics, the short-run coefficients are used in constructing the fit statistics in the bottom portion of the results view (we again urge caution in using these measures). The short-run dynamics are also used in computing the residuals used by various equation views and procs such as the residual plot or the gradient view.

The short-run coefficients are not included in the representations view of the equation, which focuses only on the estimates for [Equation (27.1)](#). Furthermore, forecasting and model solution using an equation estimated by DOLS are also based on the long-run relationship. If you wish to construct forecasts that incorporate the short-run dynamics, you may use least squares to estimate an equation that explicitly includes the lags and leads of the cointegrating regressors.

Last updated: Thu, 04 Apr 2019 22:48:40 PST

Back to top