Lecture 22. Cointegration. Modelling with Nonstationary Time Series.

Cointegration

Let $X_1, \ldots, X_n$ be integrated series. In general, if you take a linear combination of these series, the order of integration of the linear combination will be equal to the maximum order of integration among the series. To illustrate, if series X_1 and X_2 are integrated of the order 1 and 2 respectively, then their linear combination will be integrated of the order 2. This seems to imply that a linear combination of series with the same order of integration k will be integrated of the order k. **However,** this is not always the case, and if the series have some long-run relationship, the order of integration of their linear combination can be lower.

Two series (with the same order of integration $k \geq 1$) are called **cointegrated** if there exists their stationary linear combination.

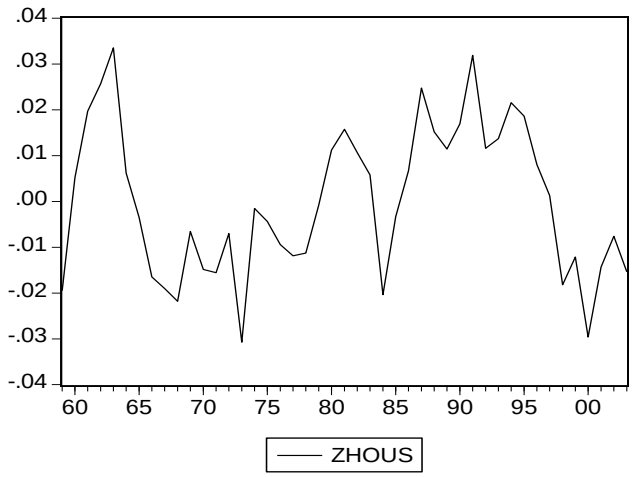
We will consider a particular case of cointegration of I(1) series which are rather typical for economic data. Two or more I(1) series are cointegrated if there exists their I(0), i.e. **stationary**, linear combination. Even though each of them has a Random Walk type behaviour, they, nevertheless, stay rather close to each other in the long run, which means an existence of actual (not spurious) long run relationship.

Example

A logarithmic regression of expenditure on housing on *DPI* and the relative price of housing was estimated for the USA, 1959-2003:

$$\widehat{LGHOUS} = \underset{}{0.006} + \underset{(0.17)}{1.03}\, LGDPI - \underset{(0.007)}{0.48}\, LGPRHOUS \qquad \underset{(0.04)}{R^2 = 0.999.}$$

The residuals for this regression are presented on the graph below.



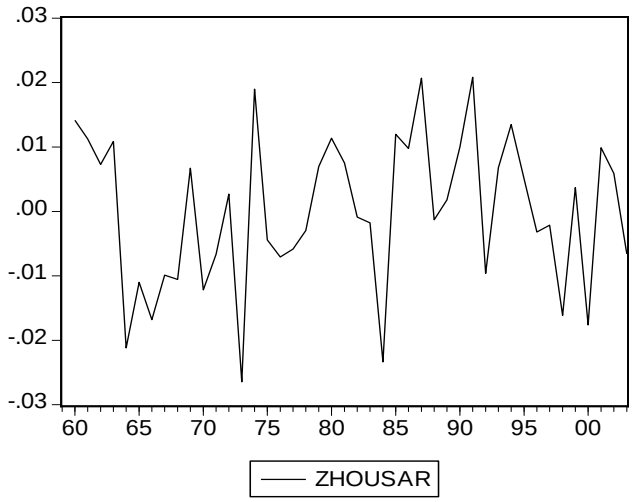
The residuals' behaviour looks more or less stationary. The ADF test statistic (intercept, no trend) is –2.91, while the asymptotic critical value for two explanatory variables is –3.34 (there are critical values for cointegrating relationships which are even higher in absolute value than those for ordinary ADF tests). So, it is not significant even at the 5 percent level, and we cannot reject a hypothesis of nonstationarity for the residuals. This would mean that the variables *LGHOUS*, *LGDPI* and *LGPRHOUS* are not cointegrated, but it is likely due to the low power of the test. The estimated coefficient of the lagged residuals is –0.33 which corresponds to the AR(1) process with ρ equal to about 0.67. It is quite possible that the residuals are stationary but autocorrelated with high ρ coefficient, and the variables are in fact cointegrated.

Durbin d-statistic equal to 0.63 indicates that it is possible that the disturbance term is subject to positive autocorrelation, with the cointegrating relationship being correct.

If the AR(1) term is added to the initial model, then we get

$$\widehat{LGHOUS} = \underset{(0.35)}{0.155} + \underset{(0.02)}{1.01}\, LGDPI - \underset{(0.09)}{0.48}\, LGPRHOUS \qquad R^2 = 0.999.$$

With the estimate of ρ equal to 0.72 (s.e. 0.16); d=1.82. All the coefficients estimated are close to those received before. New residuals look stationary:



and the ADF test statistic for them is −6.3. Though we cannot do the formal test without knowing the critical level for this particular case, which is different from the standard one, it seems that cointegrating relationship has been found.

Once more, cointegrating relationship represents a long-run link between the variables.

Fitting Models with Nonstationary Time Series

Since estimating models with nonstationary time series may rather often lead to spurious regressions, the idea is to transform the model in such a way that the series in it become stationary. We will consider three approaches: Detrending, Differencing and Error Correction Models.

Detrending

In the models with variables which include time trends, removal of the trends, or detrending, allows to avoid getting spurious regressions. On the detrending procedure, see the Lecture 19. As indicated, detrending of each variable in the model is equivalent to including the time trend as one of the explanatory variables. The coefficients would be the same for these two cases while the standard errors slightly differ (they are correct when just time variable is included).

Economic indicators rather often behave not as series including time trends, but as random walks. If you detrend a series which is in fact a random walk with a drift, then its variance still increases in time proportionally to the variable *t*, the series does not become stationary, and hence the problem of spurious regressions is not resolved.

Differencing

If having random walk time series, differencing is a procedure which can be applied:

subtracting $Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + u_{t-1}$ from $Y_t = \beta_1 + \beta_2 X_t + u_t$, we get

$$\Delta Y_t = \beta_2 \Delta X_t + \Delta u_t = \beta_2 \Delta X_t + (\rho - 1) u_{t-1} + \varepsilon_t.$$

The series ΔY_t and ΔX_t are stationary, and the coefficient β_2 can be estimated from this model. At the same time, the new disturbance term Δu_t is subject to autocorrelation, and appropriate remedial measures should be applied. Only in the case of severe autocorrelation of u_t in the initial model (ρ is close to 1) differencing also helps to reduce the autocorrelation influence. If Y_t and X_t are unrelated I(1) processes, absence of their relationship will be revealed in the differenced model, so the problem of spurious regressions will be resolved.

At the same time, there are a few more shortcomings in the differenced model: the constant disappears (though it is usually of low interest) and only short-run relationships can be investigated with it. In the long-run equilibrium $\Delta Y = \Delta X = 0$, and hence no conclusions about long-run relationship can be made. The approach for including the long-run relationships is the Error correction model.

Error correction model

Error correction model involves transforming the original model with nonstationary time series in such a way that all the series in the transformed model are stationary, and at the same time it includes the description of both short-run and long-run relationship between the variables. But every model development has its price: here this is an assumption about the particular form of relationship between the variables, the ADL(1,1) one.

So, let the model be ADL(1,1), and X_t and Y_t are both I(1) series:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + \varepsilon_t$$

Assume that in the long-run, all variables reach their steady states. This means that

$$X_t = X_{t-1} = \bar{X}$$

$$Y_t = Y_{t-1} = \bar{Y}$$

So, in the long-run, the equation looks like this:

$$\bar{Y} = \beta_0 + \beta_1 \bar{Y} + \beta_2 \bar{X} + \beta_3 \bar{X}$$

$$\bar{Y}(1 - \beta_1) = \beta_0 + (\beta_2 + \beta_3)\bar{X}$$

$\bar{Y} = \frac{\beta_0}{1-\beta_1} + \frac{\beta_2+\beta_3}{1-\beta_1} \bar{X}$ - long-run relationship between *X* and *Y*

Now it is assumed that the long-run relationship is the cointegrating relationship. This implies that in the expression

$$Y_t = \frac{\beta_0}{1-\beta_1} + \frac{\beta_2+\beta_3}{1-\beta_1} X_t + \upsilon_t$$

the disturbance term υ_t is **stationary**.

The model can be transformed as follows:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + \varepsilon_t$$

Subtract Y_{t-1} from both sides:

$$Y_t - Y_{t-1} = \beta_0 + \beta_1 Y_{t-1} - Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + \varepsilon_t$$

$$\Delta Y_t = \beta_0 + (\beta_1 - 1) Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + \varepsilon_t$$

Add and subtract $\beta_2 X_{t-1}$:

$$\Delta Y_t = \beta_0 + (\beta_1 - 1) Y_{t-1} + \beta_2 X_t - \beta_2 X_{t-1} + \beta_3 X_{t-1} + \beta_2 X_{t-1} + \varepsilon_t$$

$$\Delta Y_t = \beta_0 + (\beta_1 - 1) Y_{t-1} + \beta_2 \Delta X_t + (\beta_3 + \beta_2) X_{t-1} + \varepsilon_t$$

Take $(\beta_1 - 1)$ out of the brackets

$$\Delta Y_t = (\beta_1 - 1)\left(Y_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1} X_{t-1}\right) + \beta_2 \Delta X_t + \varepsilon_t$$

Thus, ΔY_t and ΔX_t are stationary, as the original series are I(1) and the expression in the brackets is also stationary, since it is the cointegrating relationship for $(t-1)$ time unit.

So, all the series in the transformed model are stationary and spurious regression is no longer a threat. The only problem is that the cointegrating relationship parameters are unknown. To deal with this, **Engle-Granger two-step procedure** is used in practice:

Step1. Estimate the cointegrating relationship with OLS.

Step 2. Using the estimated relationship, fit the error correction model.

Engle and Granger showed that the results of the two-step procedure are asymptotically the same as in the case when the true cointegrating relationship is used.

Let us consider the model be ADL(1,1) with two explanatory variables, X_t and Z_t, and Y_t as the dependent variable, all are I(1) series:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + \beta_4 Z_t + \beta_5 Z_{t-1} + \varepsilon_t$$

After the same transformations as before, we have

$$\Delta Y_t = (\beta_1 - 1)\left(Y_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1} X_{t-1} - \frac{\beta_4+\beta_5}{1-\beta_1} Z_{t-1}\right) + \beta_2 \Delta X_t + \beta_4 \Delta Z_t + \varepsilon_t$$

We have already estimated the cointegrating relationship for the logarithmic function of demand for housing (USA, 1959-2003) above, and now we estimate the error correction model (EViews printout below):

Dependent Variable: DLGHOUS
Method: Least Squares
Sample (adjusted): 1960 2003

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ZHOUS(-1)	-0.311355	0.111123	-2.801888	0.0077
DLGDPI	0.938132	0.048823	19.21503	0.0000
DLGPRHOUS	-0.498342	0.122453	-4.069678	0.0002
R-squared	0.249635	Durbin-Watson stat		1.626460

Here ZHOUS(-1) is the lagged residual of the cointegrating relationship. The estimation shows that about 0.31 of the short-run deviation from the equilibrium is covered each year, and the estimates of the short-run income and price elasticities of demand for housing are 0.94 and −0.50 which are rather close to the estimates of the long-run elasticities (1.01 and −0.48).