

Chapter 6

Multivariate Time Series Analysis

6.0.1 Introduction

Multivariate analysis investigates dependence and interactions among a set of variables in multi-values processes. One of the most powerful method of analyzing multivariate time series is the vector autoregression model. It is a natural extension of the univariate autoregressive model to the multivariate case.

In this chapter we cover concepts of VAR modelling, non-stationary multivariate time series and cointegration.

More detailed discussion can be found in Hamilton (1994), Harris (1995), Enders (2004), Tsay (2002), Zivot and Wang (2006).

6.1 Vector Autoregression Model

Let $\mathbf{Y}_t = (Y_{1,t}, Y_{2,t}, \dots, Y_{n,t})'$ denote an $k \times 1$ vector of time series variables. The basic vector autoregressive model of order p , $VAR(p)$, is

$$\mathbf{Y}_t = \mathbf{c} + \Pi_1 \mathbf{Y}_{t-1} + \Pi_2 \mathbf{Y}_{t-2} + \dots + \Pi_p \mathbf{Y}_{t-p} + \mathbf{u}_t, \quad t = 1, \dots, T, \quad (6.1.1)$$

where Π_i are $k \times k$ matrices of coefficients, \mathbf{c} is a $k \times 1$ vector of constants and \mathbf{u}_t is an $k \times 1$ unobservable zero mean white noise vector process with covariance matrix Σ .

If we consider a special case of two dimensional vector \mathbf{Y} , the VAR consists of two equation (also called a bivariate VAR)

$$\begin{aligned} \mathbf{Y}_t &= \begin{pmatrix} Y_{1,t} \\ Y_{2,t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \pi_{11}^1 & \pi_{12}^1 \\ \pi_{21}^1 & \pi_{22}^1 \end{pmatrix} \begin{pmatrix} Y_{1,t-1} \\ Y_{2,t-1} \end{pmatrix} \\ &+ \begin{pmatrix} \pi_{11}^2 & \pi_{12}^2 \\ \pi_{21}^2 & \pi_{22}^2 \end{pmatrix} \begin{pmatrix} Y_{1,t-2} \\ Y_{2,t-2} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} \end{aligned} \quad (6.1.2)$$

with $\text{cov}(u_{1,t}, u_{2,s}) = \sigma_{12}$ for $t \neq s$.

As in the univariate case with AR processes, we can use the lag operator to represent $VAR(p)$

$$\Pi(L)\mathbf{Y}_t = \mathbf{c} + \mathbf{u}_t,$$

where $\Pi(L) = I_n - \Pi_1 L - \dots - \Pi_p L^p$.

If we impose stationarity on \mathbf{Y}_t in (6.1.2), the unconditional expected value is given by

$$\boldsymbol{\mu} = (I_n - \Pi_1 - \dots - \Pi_p)^{-1} \mathbf{c}.$$

Very often other deterministic terms or stochastic exogenous variables may be included into the VAR specification to represent. More general form of the $VAR(p)$ model is

$$\mathbf{Y}_t = \Pi_1 \mathbf{Y}_{t-1} + \Pi_2 \mathbf{Y}_{t-2} + \dots + \Pi_p \mathbf{Y}_{t-p} + \Gamma \mathbf{X}_t + \mathbf{u}_t,$$

where \mathbf{X}_t represents an $m \times 1$ matrix of exogenous or deterministic variables, and Γ is a matrix of parameters.



Brain power

By 2020, wind could provide one-tenth of our planet's electricity needs. Already today, SKF's innovative know-how is crucial to running a large proportion of the world's wind turbines.

Up to 25 % of the generating costs relate to maintenance. These can be reduced dramatically thanks to our systems for on-line condition monitoring and automatic lubrication. We help make it more economical to create cleaner, cheaper energy out of thin air.

By sharing our experience, expertise, and creativity, industries can boost performance beyond expectations. Therefore we need the best employees who can meet this challenge!

The Power of Knowledge Engineering

Plug into The Power of Knowledge Engineering.
Visit us at www.skf.com/knowledge

SKF



6.1.1 Estimation of VARs and Inference on coefficients

Since the $VAR(p)$ may be written as a system of equations with the same sets of explanatory variables, its coefficients can be efficiently and consistently estimated by estimating each of the components using the OLS method (see Hamilton (1994)). Under standard assumptions regarding the behavior of stationary and ergodic VAR models (see Hamilton (1994)) the estimators of the coefficients are asymptotically normally distributed.

An element of $\hat{\Pi}_i$ is asymptotically normally distributed, so asymptotically valid t-tests on individual coefficients may be constructed in the usual way (see Chapter 2). More general linear hypotheses can also be tested using the Wald statistic.

Lag Length Selection A reasonable strategy how to determine the lag length of the VAR model is to fit $VAR(p)$ models with different orders $p = 0, \dots, p_{max}$ and choose the value of p which minimizes some model selection criteria. Model selection criteria for $VAR(p)$ could be based on Akaike (AIC), Schwarz-Bayesian (BIC) and Hannan-Quinn (HQ) information criteria:

$$\begin{aligned} AIC(p) &= \ln |\bar{\Sigma}(p)| + \frac{2}{T}pn^2 \\ BIC(p) &= \ln |\bar{\Sigma}(p)| + \frac{\ln T}{T}pn^2 \\ HQ(p) &= \ln |\bar{\Sigma}(p)| + \frac{2 \ln \ln T}{T}pn^2 \end{aligned}$$

Forecasting We can use VAR model to forecast times series in a similar way to forecasting from a univariate AR model.

The one-period-ahead forecast based on information available at time T is

$$\mathbf{Y}_{T+1|T} = \mathbf{c} + \Pi_1 \mathbf{Y}_T + \dots + \Pi_p \mathbf{Y}_{T-p+1}$$

while h -step forecast is

$$\mathbf{Y}_{T+h|T} = \mathbf{c} + \Pi_1 \mathbf{Y}_T + \dots + \Pi_p \mathbf{Y}_{T-p+1},$$

where $\mathbf{Y}_{T+j|T} = \mathbf{Y}_{T+j}$ for $j < 0$. The h -step forecast errors may be expressed as

$$\mathbf{Y}_{T+h} - \mathbf{Y}_{T+h|T} = \sum_{s=0}^{h-1} \Psi_s \varepsilon_{T+h-s},$$

where the matrices Ψ_s are determined by recursive substitution

$$\Psi_s = \sum_{j=1}^{p-1} \Psi_{s-j} \Pi_j \quad (6.1.3)$$

with $\Psi_0 = I_n$ and $\Pi_j = 0$ for $j > p$. The forecasts are unbiased since all of the forecast errors have expectation zero and the MSE matrix for $\mathbf{Y}_{t+h|T}$ is

$$\Sigma(h) = \text{MSE}(\mathbf{Y}_{T+h} - \mathbf{Y}_{T+h|T}) = \sum_{s=0}^{h-1} \Psi_s \Sigma \Psi_s'$$

The h -step forecast in the case of estimated parameters is

$$\hat{\mathbf{Y}}_{T+h|T} = \hat{\Pi}_1 \hat{\mathbf{Y}}_{T+h-1|T} + \dots + \hat{\Pi}_p \hat{\mathbf{Y}}_{T+h-p|T},$$

where $\hat{\Pi}_j$ are the estimated matrices of parameters. The h -step forecast error is now

$$\mathbf{Y}_{T+h} - \hat{\mathbf{Y}}_{T+h|T} = \sum_{s=0}^{h-1} \Psi_s \varepsilon_{T+h-s} + \left(\mathbf{Y}_{t+h} - \hat{\mathbf{Y}}_{T+h|T} \right)$$

The estimate of the MSE matrix of the h -step forecast is then

$$\hat{\Sigma}(h) = \sum_{s=0}^{h-1} \hat{\Psi}_s \hat{\Sigma} \hat{\Psi}_s'$$

with $\Psi_s = \sum_{j=1}^s \hat{\Psi}_{s-j} \hat{\Pi}_j$.

6.1.2 Granger Causality

One of the main uses of VAR models is forecasting. The structure of the VAR model provides information about a variable's or a group of variables' forecasting ability for other variables. The following intuitive notion of a variable's forecasting ability is due to Granger (1969). If a variable, or group of variables, Y_1 is found to be helpful for predicting another variable, or group of variables, Y_2 then Y_1 is said to Granger-cause Y_2 ; otherwise it is said to fail to Granger-cause Y_2 . Formally, Y_1 fails to Granger-cause Y_2 if for all $s > 0$ the MSE of a forecast of $Y_{2,t+s}$ based on $(Y_{2,t}, Y_{2,t-1}, \dots)$ is the same as the MSE of a forecast of $Y_{2,t+s}$ based on $(Y_{2,t}, Y_{2,t-1}, \dots)$ and $(Y_{1,t}, Y_{1,t-1}, \dots)$. Note that the notion of Granger causality only implies forecasting ability.

In a bivariate $VAR(p)$ model for $\mathbf{Y}_t = (Y_{1t}, Y_{2t})'$, Y_2 fails to Granger-cause Y_1 if all of the p VAR coefficient matrices Π_1, \dots, Π_p are lower triangular. That is, all of the coefficients on lagged values of Y_2 are zero in the equation for Y_1 . The p linear coefficient restrictions implied by Granger non-causality may be tested using the Wald statistic. Notice that if Y_2 fails to Granger-cause Y_1 and Y_1 fails to Granger-cause Y_2 , then the VAR coefficient matrices Π_1, \dots, Π_p are diagonal.

6.1.3 Impulse Response and Variance Decompositions

As in the univariate case, a $VAR(p)$ process can be represented in the form of a vector moving average (VMA) process.

$$\mathbf{Y}_t = \boldsymbol{\mu} + \mathbf{u}_t + \Psi_1 \mathbf{u}_{t-1} + \Psi_2 \mathbf{u}_{t-2} + \dots,$$

where the $k \times k$ moving average matrices Ψ_s are determined recursively using (6.1.3).

The elements of coefficient matrices Ψ_s mean effects of \mathbf{u}_{t-s} shocks on \mathbf{Y}_t . That is, the (i, j) -th element, ψ_{ij}^s , of the matrix Ψ_s is interpreted as the impulse response

$$\frac{\partial Y_{i,t+s}}{\partial u_{j,t}} = \frac{\partial Y_{i,t}}{\partial u_{j,t-s}} = \psi_{ij}^s, \quad i, j = 1, \dots, T.$$

Sets of coefficients $\psi_{ij}(s) = \psi_{ij}^s$, $i, j = 1, \dots, T$ are called the **impulse response functions**.

It is possible to decompose the h -step-ahead forecast error variance into the proportions due to each shock u_{jt} .

The forecast variance decomposition determines the proportion of the variation Y_{jt} due to the shock u_{jt} versus shocks of other variables u_{it} for $i \neq j$.

“I studied English for 16 years but...
...I finally learned to speak it in just six lessons”
Jane, Chinese architect

ENGLISH OUT THERE

Click to hear me talking before and after my unique course download

6.1.4 VAR in EViews

As an example of VAR estimation in EViews, consider two time series of returns of monthly IBM stocks and the market portfolio returns from Fama-French database (data is contained in IBM1.wf1).

There are several ways to estimate VAR model in EViews. The first one is through the main menu. Clicking on **View/Estimate VAR...** will open a dialog window for VAR model estimation.

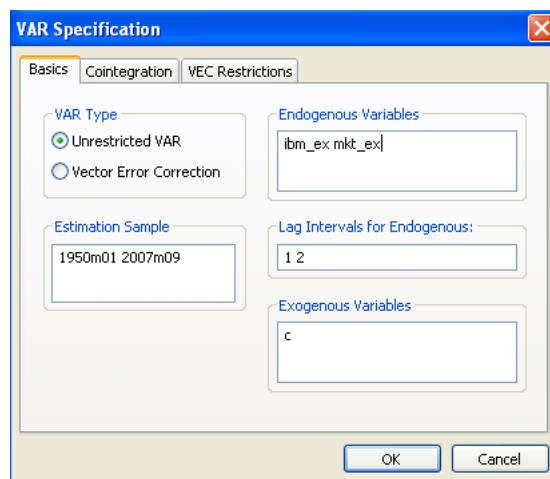


Figure 6.1: VAR model estimation dialog window

We choose **Unrestricted VAR** and in the **Endogenous Variables** box we have to specify the list of endogenous time series variables to be included in the VAR model. We consider two excess return series of the IBM stock `IBM_ex` and the market portfolio `Mkt_ex`.

In the **Lag Intervals for Endogenous** we have to specify the order of the model, that is interval of lags to be included in the model. If we want to build a model with only two lags, we write `1 2`. This means, we include all lags beginning from the first one and ending with the lag of order 2. We do not specify any exogenous variables apart from the intercept term `c`.

Another way of calling the VAR estimation dialog window is to select both endogenous variables in the workfile and in the context menu (right button click) choose **Open/as VAR...** The **Endogenous Variables** box will be filled in automatically.

Finally, we can estimate VAR model from the command line. There is a separate object, called `var`, to declare the VAR model. The estimation of the above mentioned example will look like

```
var ibm2.ls 1 2 ibm_ex mkt_ex
```

Here `ibm2` is a name of the var-object which will be saved in the workfile, `ls` indicates the estimation method; in this case it is OLS estimation method of the unrestricted VAR model. Then, specifications of the lags pairs and the list of endogenous variables follow. If one wishes to include exogenous variables besides the intercept, it can be done by typing a symbol `@` followed by a list of exogenous variables. For example,

```
var ibm2.ls 1 2 ibm_ex mkt_ex @ exvar1 exvar2
```

Click **OK** and EViews produces an estimation output for the specified VAR model.

	IBM_EX	MKT_EX
IBM_EX(-1)	0.009969 (0.04688) [0.21267]	0.004618 (0.02849) [0.16208]
IBM_EX(-2)	0.061395 (0.04689) [1.30944]	0.011865 (0.02850) [0.41630]
MKT_EX(-1)	0.069736 (0.07747) [0.90017]	0.074885 (0.04709) [1.59023]
MKT_EX(-2)	-0.198868 (0.07745) [-2.56754]	-0.044936 (0.04708) [-0.95443]
C	0.008388 (0.00268) [3.12617]	0.001999 (0.00163) [1.22556]
R-squared	0.011416	0.007309
Adj. R-squared	0.005651	0.001520
Sum sq. resids	3.306425	1.221712
S.E. equation	0.069425	0.042201
F-statistic	1.980399	1.262645
Log likelihood	865.2686	1209.253
Akaike AIC	-2.489923	-3.485538
Schwarz SC	-2.457086	-3.452700
Mean dependent	0.008725	0.002207
S.D. dependent	0.069622	0.042233
Determinant resid covariance (dof adj.)	5.64E-06	
Determinant resid covariance	5.56E-06	
Log likelihood	2219.343	
Akaike information criterion	-6.394625	
Schwarz criterion	-6.328950	

Figure 6.2: Output for the VAR model estimation

Two columns correspond to two equation in the VAR model. The only significant coefficient besides the intercept one is at the second lag of the market portfolio

returns in the IBM equation. As expected, there is a unidirectional dynamic relationship from the market portfolio returns to the IBM returns, Thus, the IBM return is affected by the past movements of the market while past movements of IBM stock returns do not affect the market portfolio returns. The second equation (for market portfolio) is not significant as suggested by the F-statistics. This means that the the estimated model cannot explain variation in the market portfolio returns. This can happen because we possibly omitted some important exogenous variables or the order of the model is inappropriately selected. EViews provides a tool to choose the most suitable lag order. In the workfile menu choose **View/Lag Structure/Lag Length Criteria...** to determine the optimal model structure. In the appeared **Lag Specification** window we choose $p_{max} = 8$ (maximal lag order).

All criteria indicate that the optimal lag order of the model is 0. This means that the VAR model is inappropriate model to explain IBM and market portfolio returns. Indeed, we know from the CAPM that market portfolio returns affect the stock returns contemporaneously and are not in lag relationship. Thus, either additional exogenous factors should be found to include in the model or another structure of the model should be employed in this case.



What do you want to do?

No matter what you want out of your future career, an employer with a broad range of operations in a load of countries will always be the ticket. Working within the Volvo Group means more than 100,000 friends and colleagues in more than 185 countries all over the world. We offer graduates great career opportunities – check out the Career section at our web site www.volvogroup.com. We look forward to getting to know you!

VOLVO
 AB Volvo (publ)
www.volvogroup.com

VOLVO TRUCKS | RENAULT TRUCKS | MACK TRUCKS | VOLVO BUSES | VOLVO CONSTRUCTION EQUIPMENT | VOLVO PENTA | VOLVO AERO | VOLVO IT
 VOLVO FINANCIAL SERVICES | VOLVO 3P | VOLVO POWERTRAIN | VOLVO PARTS | VOLVO TECHNOLOGY | VOLVO LOGISTICS | BUSINESS AREA ASIA

Download free eBooks at bookboon.com



Var Lag Order Selection Criteria
 Endogenous variables: IBM_EX MKT_EX
 Exogenous variables: C
 Date: 01/03/09 Time: 12:27
 Sample: 1950M01 2007M09
 Included observations: 685

Lag	LogL	LR	FPE	AIC	SC	HQ
0	2194.622	NA*	5.69e-06*	-6.401817*	-6.388593*	-6.396700*
1	2196.775	4.287075	5.72e-06	-6.396424	-6.356751	-6.381073
2	2199.864	6.131986	5.73e-06	-6.393763	-6.327640	-6.368177
3	2201.623	3.481326	5.77e-06	-6.387219	-6.294647	-6.351399
4	2203.219	3.150236	5.81e-06	-6.380200	-6.261179	-6.334146
5	2205.939	5.354029	5.83e-06	-6.376465	-6.230995	-6.320176
6	2206.141	0.395622	5.90e-06	-6.365375	-6.193455	-6.298852
7	2206.432	0.568871	5.96e-06	-6.354545	-6.156177	-6.277787
8	2207.274	1.642798	6.02e-06	-6.345326	-6.120508	-6.258334

* indicates lag order selected by the criterion
 LR: sequential modified LR test statistic (each test at 5% level)
 FPE: Final prediction error
 AIC: Akaike information criterion
 SC: Schwarz information criterion
 HQ: Hannan-Quinn information criterion

Figure 6.3: Output for the lag length selection procedure

Lag selection can be programmed manually in the same way as it is done for ARMA model (see Chapter 3). There are some command references given below which can be used to assess various statistic values in the VAR analysis in EViews.

6.2 Cointegration

The assumption of stationary of regressors and regressands is crucial for the properties and the OLS estimators discussed in Chapter 2. In this case, the usual statistical results for the linear regression model and consistency of estimators hold. However, when variables are non-stationary then the usual statistical results may not hold.

6.2.1 Spurious Regression

If there are trends in the data (deterministic or stochastic) this can lead to a spurious results when running OLS regression. This is because time trend will dominate other stationary variables and the OLS estimators will pick up covariances generated by time trends only. While the effects of deterministic trends can be removed from the regression by either including time trend regressor or simply de-trending variables, non-stationary variables with stochastic trends may lead to invalid inferences.

Var Data Members

Data Member	Description
@eqlogl(k)	log likelihood for equation k
@eqncoef(k)	number of estimated coefficients in equation k
@eqregobs(k)	number of observations in equation k
@meandep(k)	mean of the dependent variable in equation k
@r2(k)	R-squared statistic for equation k
@rbar2(k)	adjusted R-squared statistic for equation k
@sddep(k)	standard deviation of dependent variable in equation k
@se(k)	standard error of the regression in equation k
@ssr(k)	sum of squared residuals in equation k
@aic	Akaike information criterion for the system
@detresid	determinant of the residual covariance matrix
@hq	Hannan-Quinn information criterion for the system
@logl	log likelihood for system
@ncoefs	total number of estimated coefficients in the var
@neqn	number of equations
@regobs	number of observations in the var
@sc	Schwarz information criterion for the system
@svarcvgtype	Returns an integer indicating the convergence type of the structural decomposition estimation: 0 (convergence achieved), 2 (failure to improve), 3 (maximum iterations reached), 4 (no convergence-structural decomposition not estimated)
@svaroverid	over-identification LR statistic from structural factorization
@totalobs	sum of "@eqregobs" from each equation ("@regobs*@neqn")
@coefmat	coefficient matrix (as displayed in output table)
@coefse	matrix of coefficient standard errors (corresponding to the output table)
@impfact	factorization matrix used in last impulse response view
@lrrsp	accumulated long-run responses from last impulse response view
@lrrspse	standard errors of accumulated long-run responses
@residcov	covariance matrix of the residuals
@svaramat	estimated A matrix for structural factorization
@svarbmat	estimated B matrix for structural factorization
@svarcovab	covariance matrix of stacked A and B matrix for structural factorization
@svarrcov	restricted residual covariance matrix from structural factorization

Consider, for example,

$$\begin{aligned} Y_{1,t} &= Y_{1,t-1} + u_{1,t}, & u_{1,t} &\sim IN(0, 1) \\ Y_{2,t} &= Y_{2,t-1} + u_{2,t}, & u_{2,t} &\sim IN(0, 1) \end{aligned}$$

Both of the variables are non-stationary and independent from each other. In the regression $Y_{1,t} = \beta_0 + \beta_1 Y_{2,t} + \varepsilon_t$, the value of true slope parameter $\beta_1 = 0$. Thus, the value of the OLS estimate $\hat{\beta}_1$ should be insignificant. The actual estimations produce high R^2 coefficients and highly significant β_1 .

The problem with the spurious regression is that t- and F-statistics do not follow standard distributions. As shown in Phillips (1986), $\hat{\beta}_1$ does not converge in

probability to zero, R^2 converges to unity as $T \rightarrow \infty$ so that the model will appear to fit well even though it is misspecified.

Regression with $I(1)$ data only makes sense when the data are *cointegrated*.

6.2.2 Cointegration

Let $\mathbf{Y}_t = (Y_{1t}, \dots, Y_{kt})'$ denote an $k \times 1$ vector of $I(1)$ time series. \mathbf{Y}_t is cointegrated if there exists an $k \times 1$ vector $\beta = (\beta_1, \dots, \beta_k)'$ such that

$$\mathbf{Z}_t = \beta' \mathbf{Y}_t = \beta_1 Y_{1t} + \dots + \beta_k Y_{kt} \sim I(0). \quad (6.2.1)$$

The non-stationary time series in \mathbf{Y}_t are cointegrated if there is a linear combination of them that is stationary. If some elements of β are equal to zero then only the subset of the time series in \mathbf{Y}_t with non-zero coefficients is cointegrated.

There may be different vectors β such that $\mathbf{Z}_t = \beta' \mathbf{Y}_t$ is stationary. In general, there can be $0 < r < k$ linearly independent cointegrating vectors. All cointegrating vectors form a *cointegrating matrix* \mathbf{B} . This matrix is again not unique. Some normalization assumption is required to eliminate ambiguity from the definition.

gaiteye
Challenge the way we run

**EXPERIENCE THE POWER OF
FULL ENGAGEMENT...**

.....

**RUN FASTER.
RUN LONGER..
RUN EASIER...**

**READ MORE & PRE-ORDER TODAY
WWW.GAITEYE.COM**

A typical normalization is

$$\beta = (1, -\beta_2, \dots, -\beta_k)'$$

so that the cointegration relationship may be expressed as

$$\mathbf{Z}_t = \beta' \mathbf{Y}_t = Y_{1t} - \beta_2 Y_{2t} - \dots - \beta_k Y_{kt} \sim I(0).$$

6.2.3 Error Correction Models

Engle and Granger (1987) state that if a bivariate $I(1)$ vector $\mathbf{Y}_t = (Y_{1t}, Y_{2t})'$ is cointegrated with cointegrating vector $\beta = (1, -\beta_2)'$ then there exists an error correction model (ECM) of the form

$$\Delta Y_{1t} = \delta_1 + \phi_1(Y_{1,t-1} - \beta_1 Y_{2,t-1}) + \sum_{j=1}^p \alpha_{11}^j \Delta Y_{1,t-j} + \sum_{s=1}^p \alpha_{12}^j \Delta Y_{2,t-j} + \varepsilon_{1t} \quad (6.2.2)$$

$$\Delta Y_{2t} = \delta_2 + \phi_2(Y_{1,t-1} - \beta_2 Y_{2,t-1}) + \sum_{j=1}^p \alpha_{21}^j \Delta Y_{1,t-j} + \sum_{s=1}^p \alpha_{22}^j \Delta Y_{2,t-j} + \varepsilon_{2t} \quad (6.2.3)$$

that describes the long-term relations of Y_{1t} and Y_{2t} . If both time series are $I(1)$ but are cointegrated (have a long-term stationary relationship), there is a force that brings the error term back towards zero. If the cointegrating parameter β_1 or β_2 is known, the model can be estimated by the OLS method.

6.2.4 Tests for Cointegration: The Engle-Granger Approach

Engle and Granger (1987) show that if there is a cointegrating vector, a simple two-step residual-based testing procedure can be employed to test for cointegration. In this case, a long-run equilibrium relationship between components of \mathbf{Y}_t can be estimated by running

$$Y_{1,t} = \beta \mathbf{Y}_{2,t} + u_t, \quad (6.2.4)$$

where $\mathbf{Y}_{2,t} = (Y_{2,t}, \dots, Y_{k,t})'$ is an $(k-1) \times 1$ vector. To test the null hypothesis that \mathbf{Y}_t is not cointegrated, we should test whether the residuals $\hat{u}_t \sim I(1)$ against the alternative $\hat{u}_t \sim I(0)$. This can be done by any of the tests for unit roots. The most commonly used is the augmented Dickey-Fuller test with the constant term and without the trend term. Critical values for this test is tabulated in Phillips and Ouliaris (1990) or MacKinnon (1996).

Potential problems with Engle-Granger approach is that the cointegrating vector will not involve $Y_{1,t}$ component. In this case the cointegrating vector will not be consistently estimated from the OLS regression leading to spurious results. Also, if

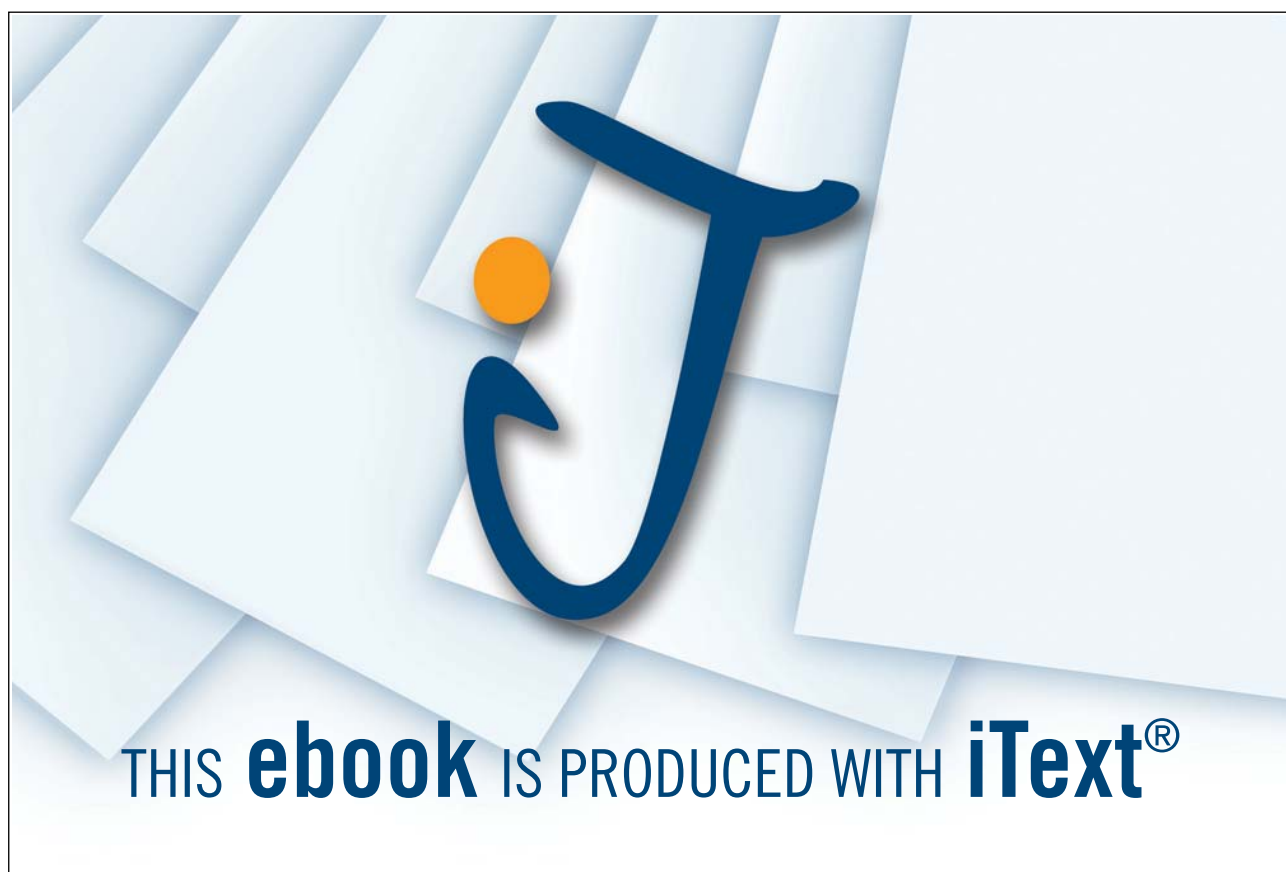
there are more than one cointegrating relation, the Engle-Granger approach cannot detect all of them.

Estimation of the static model (6.2.4) is equivalent to omitting the short-term components from the error-correction model (6.2.3). If this results for autocorrelation in residuals, although the results will still hold asymptotically, it might create a severe bias in finite samples. Because of this, it makes sense to estimate the full dynamic model. Since all variables in the ECM are $I(0)$, the model can be consistently estimated using the OLS method. This approach leads to a better performance as it does not push the short-term dynamics into residuals.

6.2.5 Example in EViews: Engle-Granger Approach

Consider as an example the Forward Premium Puzzle. Due to rational expectation hypothesis, forward rate should be unbiased predictor of future spot exchange rate. This means that in the regression of levels of spot S_{t+1} on forward rate F_t the intercept coefficient should be equal to zero and the slope coefficient should be equal to unity.

Consider monthly data of the USG/GBP spot and forward exchange rate for the period from January 1986 to November 2008 (the data is in FPP.wf1 file).



Unit roots are often found in in levels of spot and forward exchange rates. Augmented Dickey-Fuller test statistic values are -2.567 and -2.688 which are high enough to fail rejecting the null hypothesis at 5% significance level. Phillips-Perron test produces test statistic which value os on the border of the rejection region. Thus, if two series are not cointegrated, there is a danger to obtain spurious results from the OLS regression. However, if we look at plots of the two series we can see that they co-move together very closely, so we can expect existence of cointegrating relation between them.

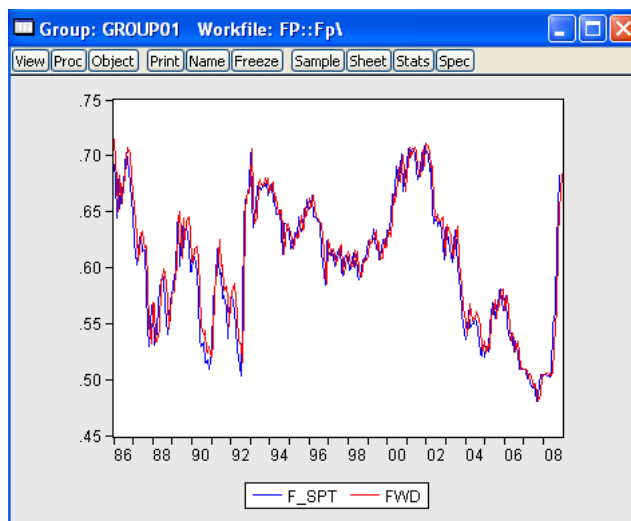


Figure 6.4: Plots of forward and future spot USD/GBP exchange rates

To perform Engle-Granger test for cointegration let us run OLS regression $S_{t+1} = \beta F_t + u_t$ in EViews and generate residuals from the model.

```
ls f_spt fwd
series resid1=resid
```

The second step is to test the residuals for stationarity. Augmented Dickey-Fuller test strongly rejects the presence of a unit root in the residual series in the favour of stationarity hypothesis.

Similar results are generated by other testing procedures. Thus, we conclude that future spot and forward exchange rates are cointegrated. Hence, the OLS results are valid for the regression in levels as well. In this case the slope coefficient is equal to 0.957 which is positive and close to unity. However, we reject the null hypothesis $H_0: \beta_1 = 1$ with the Wald test.

Thus, the forward premium puzzle also exists even for the model in levels for the exchange rates.

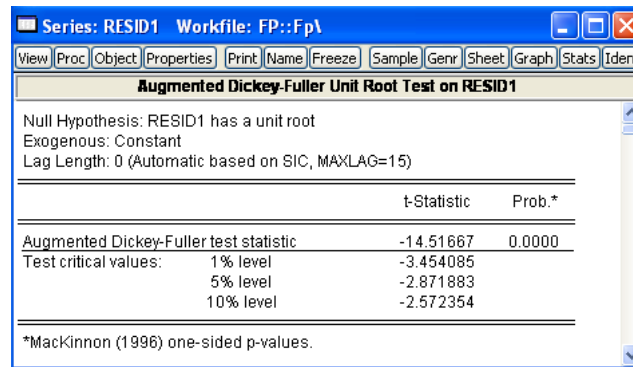


Figure 6.5: Results of Augmented Dickey-Fuller test for residuals from the long-run equilibrium relationship

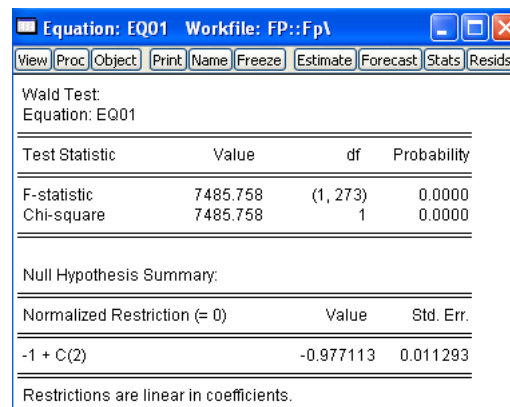


Figure 6.6: Wald test results for testing $H_0: \beta_1 = 1$

Another way of estimating cointegrating equation is to estimate a vector error correction model. To do this, open both forward and spot series as VAR system (select both series and in the context menu choose **Open/as VAR...**). In the VAR type box select **Vector Error Correction** and in the **Cointegration** tab click on **Intercept (no trend) in CE - no intercept in VAR**. EViews' output is given in Figure ??.

As expected, the output shows that the stationary series is approximately $S_{t+1} - F_t$ with the mean around zero. Deviations from the long-run equilibrium equation have significant effect on changes of the spot exchange rate. Another highly significant coefficient α_{22}^1 indicates a significant impact of ΔS_t on ΔF_t which is not surprising. This underlies the relationships between the spot and forward rate through the Covered Interest rate Parity condition (CIP).

The following subsection introduces an approach of testing for cointegration

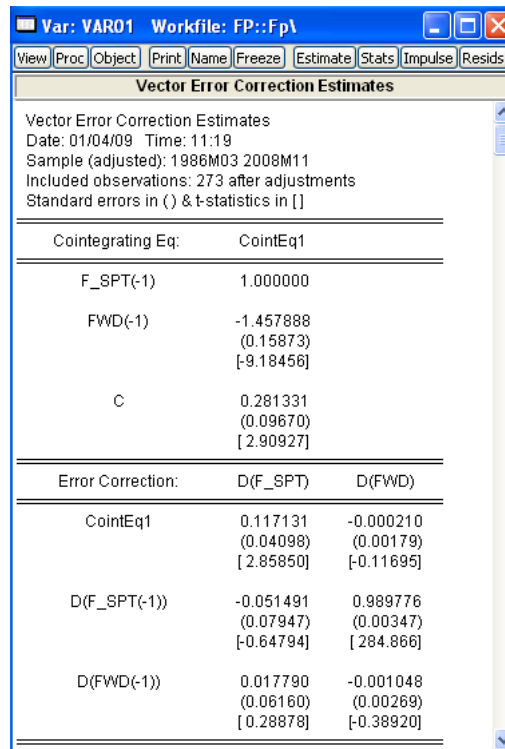


Figure 6.7: Output of the vector error correction model

when there exists more than one cointegrating relationship.

6.2.6 Tests for Cointegration: The Johansen's Approach

An alternative approach to test for cointegration was introduced by Johansen (1988). His approach allows to avoid some drawbacks existing in the Engle-Granger's approach and test the number of cointegrating relations directly. The method is based on the VAR model estimation.

Consider the $VAR(p)$ model for the $k \times 1$ vector \mathbf{Y}_t

$$\mathbf{Y}_t = \Pi_1 \mathbf{Y}_{t-1} + \dots + \Pi_p \mathbf{Y}_{t-p} + \mathbf{u}_t, \quad t = 1, \dots, T, \quad (6.2.5)$$

where $\mathbf{u}_t \sim IN(0, \Sigma)$.

Since levels of time series \mathbf{Y}_t might be non-stationary, it is better to transform Equation (6.2.5) into a dynamic form, calling vector error correction model (VECM)

$$\Delta \mathbf{Y}_t = \Pi \mathbf{Y}_{t-1} + \Gamma_1 \Delta \mathbf{Y}_{t-1} + \dots + \Gamma_{p-1} \Delta \mathbf{Y}_{t-p+1} + \mathbf{u}_t,$$

where $\Pi = \Pi_1 + \dots + \Pi_p - I_n$ and $\Gamma_k = -\sum_{j=k+1}^p \Pi_j$, $k = 1, \dots, p-1$.

Let us assume that \mathbf{Y}_t contains non-stationary $I(1)$ time series components. Then in order to get a stationary error term \mathbf{u}_t , $\Pi\mathbf{Y}_{t-1}$ should also be stationary. Therefore, $\Pi\mathbf{Y}_{t-1}$ must contain $r < k$ cointegrating relations. If the $VAR(p)$ process has unit roots then Π has reduced rank $\text{rank}(\Pi) = r < k$. Effectively, testing for cointegration is equivalent to checking out the rank of the matrix Π .

If Π has a full rank then all time series in \mathbf{Y} are stationary, if the rank of Π is zero then there are no cointegrating relationships.

If $0 < \text{rank}(\Pi) = r < k$. This implies that \mathbf{Y}_t is $I(1)$ with r linearly independent cointegrating vectors and $k - r$ non-stationary vectors. Since Π has rank r it can be written as the product

$$\underset{(k \times k)}{\Pi} = \underset{(k \times r)}{\alpha} \underset{(r \times k)}{\beta'},$$

where α and β are $k \times r$ matrices with $\text{rank}(\alpha) = \text{rank}(\beta) = r$. The matrix β is a matrix of long-run coefficients and α represents the speed of adjustment to disequilibrium. The VECM model becomes

$$\Delta\mathbf{Y}_t = \alpha\beta'\mathbf{Y}_{t-1} + \Gamma_1\mathbf{Y}_{t-1} + \dots + \Gamma_{p-1}\Delta\mathbf{Y}_{t-p+1} + \mathbf{u}_t, \quad (6.2.6)$$

with $\beta'\mathbf{Y}_{t-1} \sim I(0)$.



www.sylvania.com

**We do not reinvent
the wheel we reinvent
light.**

Fascinating lighting offers an infinite spectrum of possibilities: Innovative technologies and new markets provide both opportunities and challenges. An environment in which your expertise is in high demand. Enjoy the supportive working atmosphere within our global group and benefit from international career paths. Implement sustainable ideas in close cooperation with other specialists and contribute to influencing our future. Come and join us in reinventing light every day.

Light is OSRAM

**OSRAM
SYLVANIA** 



Johansen's methodology of obtaining estimates of α and β is given below.

Johansen's Methodology

Specify and estimate a $VAR(p)$ model (6.2.5) for \mathbf{Y}_t .

Determine the rank of Π ; the maximum likelihood estimate for β equals the matrix of eigenvectors corresponding to the r largest eigenvalues of a $k \times k$ residual matrix (see Hamilton (1994), Lutkepohl (1991), Harris (1995) for more detailed description).

Construct likelihood ratio statistics for the number of cointegrating relationships. Let estimated eigenvalues are $\hat{\lambda}_1 > \hat{\lambda}_2 > \dots > \hat{\lambda}_k$ of the matrix Π .

Johansen's likelihood ratio statistic tests the nested hypotheses

$$H_0: r \leq r_0 \text{ vs. } H_1: r > r_0$$

The likelihood ratio statistic, called the *trace statistic*, is given by

$$LR_{trace}(r_0) = -T \sum_{i=r_0+1}^k \log(1 - \hat{\lambda}_i).$$

It checks whether the smallest $k - r_0$ eigenvalues are statistically different from zero. If $\text{rank}(\Pi) = r_0$ then $\hat{\lambda}_{r_0+1}, \dots, \hat{\lambda}_k$ should all be close to zero and $LR_{trace}(r_0)$ should be small. In contrast, if $\text{rank}(\Pi) > r_0$ then some of $\hat{\lambda}_{r_0+1}, \dots, \hat{\lambda}_k$ will be nonzero (but less than 1) and $LR_{trace}(r_0)$ should be large.

We can also test $H_0: r = r_0$ against $H_1: r_0 = r_0 + 1$ using so called the *maximum eigenvalue statistic*

$$LR_{max}(r_0) = -T \log(1 - \hat{\lambda}_{r_0+1}).$$

Critical values for the asymptotic distribution of $LR_{trace}(r_0)$ and $LR_{max}(r_0)$ statistics are tabulated in Osterwald-Lenum (1992) for $k - r_0 = 1, \dots, 10$.

In order to determine the number of cointegrating vectors, first test $H_0: r_0 = 0$ against the alternative $H_1: r_0 > 0$. If this null is not rejected then it is concluded that there are no cointegrating vectors among the k variables in \mathbf{Y}_t . If $H_0: r_0 = 0$ is rejected then there is at least one cointegrating vector. In this case we should test $H_0: r_0 \leq 1$ against $H_1: r_0 > 1$. If this null is not rejected then we say that there is only one cointegrating vector. If the null is rejected then there are at least two cointegrating vectors. We test $H_0: r_0 \leq 2$ and so on until the null hypothesis is not rejected.

In a small samples tests are biased if asymptotic critical values are used without a correction. Reinsel and Ahn (1992) and Reimars (1992) suggested small samples bias correction by multiplying the test statistics with $T - kp$ instead of T in the construction of the likelihood ratio tests.

6.2.7 Example in EViews: Johansen's Approach

A very good example of a model with several cointegrating equations has been given by Johansen and Juselius (1990) (1992) (see also Harris (1995)). They considered a single equation approach to combine both Purchasing Power Parity and Uncovered Interest rate Parity condition in one model.

In this model we expect two cointegrating equations between the UK consumer price index P , the US consumer price index P^* , USD/GBP exchange rate S and two interest rates I and I^* in the domestic and foreign countries respectively. If we denote their log counterparts by the corresponding small letter, the theory suggest that the following two relationships should hold in efficient markets with rational investors: $p_t - p_t^* = s_t$ and $\Delta s_{t+1} = i_t - i_t^*$. The data is considered within the range from January 1989 to November 2008 is given in PPPFP1.wf1 file.

We create the log counterparts of the variables in the standard ways, like

```
series lcp_i_uk=log(cpi_uk)
```

and so on. In order to check for cointegration we can either estimate VECM (open 5 series as VAR model) or create a Group with the variables. Johansen and Juselius (1990) included into the model seasonal dummy variables as well as crude oil prices. We restrict ourself with only seasonal dummy for simplicity. We can create dummy variables by using a command `@expand`, which allows to create a group of dummy variables by expanding out one or more series into individual categories. For this purposes we need first to create a variable indicating the quarter of the observation. We do it in the following way

```
series quarter=@quarter(cpi_uk)
```

The command `@quarter` returns the quarter of the year in which the current observation begins. The second step is to create the dummy variables:

```
group dum=@expand(quarter)
```

EViews will create a new group object `dum` containing four dummy variables for each of the quarter of the observation.

In both cases, either with VAR or with group objects, one can perform Johansen's test procedure by clicking on **View/Cointegration Test...**

The dialog window will ask offer to specify the form of the VECM and the cointegrating equation (with or without intercept or trend components). We choose the first option with no trend and intercept to avoid perfect collinearity since we include four dummy variables as exogenous in the model. In the box **Exogenous Variables** enter the name of the dummy variables group `dum`.

In the box **Lag Intervals for D(Endogenous)** we set 1 4 – we include 4 lags

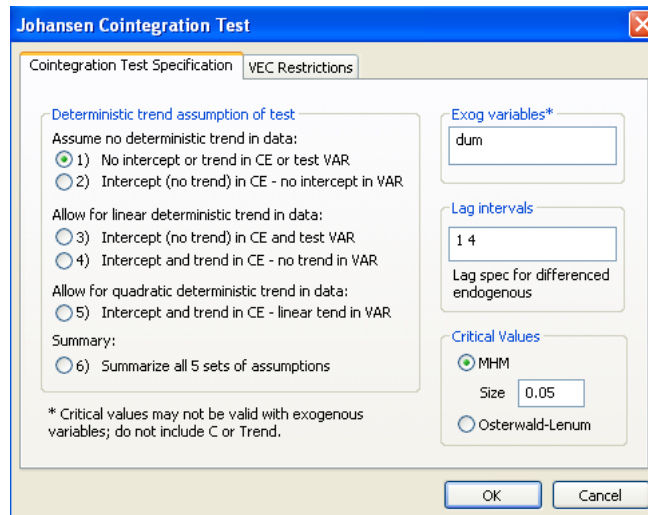


Figure 6.8: Johansen’s Cointegration test dialog window

in the model. This is determined by EViews as optimal according to 3 criteria (first estimate VAR with any of the lag specifications, check the optimality of the lag order in **View/Lag Structure/Lag Specification/Lag Length Criteria** and then re-estimate the VECM with the optimal lag order).

360° thinking.

Deloitte.

Discover the truth at www.deloitte.ca/careers

© Deloitte & Touche LLP and affiliated entities.



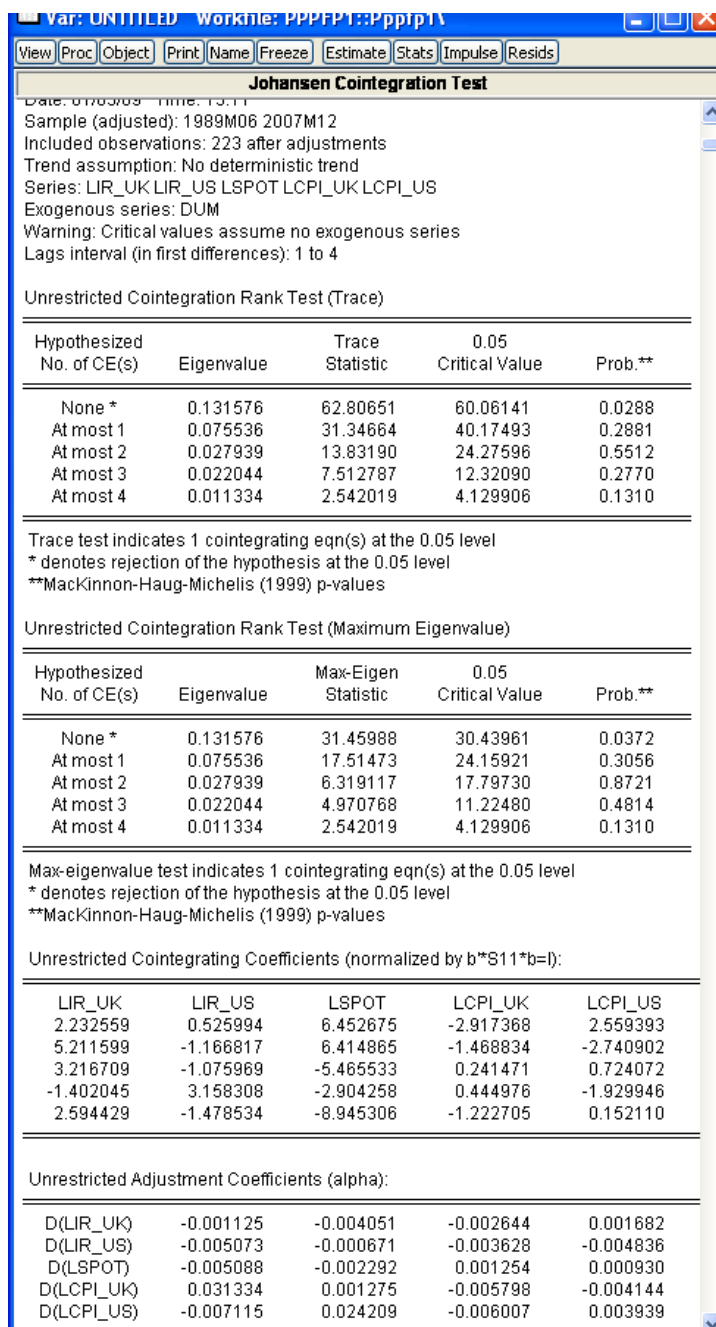


Figure 6.9: Output for Johansen’s Cointegration test

EViews produces results for various hypothesis tested, from no cointegration ($r = 0$) to increasing number of cointegrating vectors (see Figure ??). The eigenvalues of matrix $\hat{\Pi}$ is given in the second column. In the third column λ_{trace} statistic is higher than the corresponding critical value at 5% significance for the first hypothesis. This means that we reject the null hypothesis of no cointegration.

However, we cannot reject the hypothesis that there is at most one cointegrating equation. On the basis of λ_{max} statistics (the second panel) it is also possible to accept that there is only one cointegrating relationship. The following two panels provide estimates of matrices β and α respectively.

Note the warning on the top of the output window that saying that critical values assume no exogenous series. This means that we have to take into account that the critical values we are using might not be fully correct as we included exogenous dummy variables in the model. This may give as an explanation why we detected only one cointegrating equation instead of two which were expected. Another reason may be that the second relation based on the UIP condition involves changes of exchange rate rather than levels considered in the VAR model.