APPENDIX C

COINTEGRATION TEST

The economic interpretation of cointegration is that if two (or more) series are linked to form an long-run equilibrium relationship, then they will move closely together overtime and a linear combination between them will be stationary even when series may certain stochastic non-stationary trends. Hence, cointegration would represent a long-run steady-state relationship. The presence of a cointegration cannot imply accurately the causality relationship. In other words, it can either that Granger causality must exist either undirectionally or bidirectionally (Granger, 1986, 1988). Additionally, evidence of cointegration among variables rules out spurious regression.

Enger and Granger (1987) gave the definition of any variables to be cointegrated as follows:

The components of the vector x_t are said to be cointegrated of order (d, b) denoted $x_t \sim CI(d, b)$, if

(i) All components of x_i are $I(d)$;

(ii) There is vector $\alpha (\neq 0)$ so that $\alpha' x_t \sim I(0)$.

The vector α is called the cointegrating vector.

Since there are three variables (F, G, Z) to test for cointegration in our study, we use the Johansen (1988) maximum likelihood procedure in stead of the Engle and Granger (1997) procedure. One problem of the Engle and Granger procedure is that it relies on a 'two-step' estimator. The first step is to generate the error series $\{\hat{\varepsilon}_t\}$ and the second step will use these generated errors to estimate the regression of the form $\Delta \hat{\varepsilon}_t = a_1 \hat{\varepsilon}_{t-1} + ...$ hence, the coefficient a_1 is obtained by estimating the regression using the residuals from another regression. As a result, any error in step 1 is carried into step 2. Another defect is that the estimator of the long-run equilibrium regression requires that we have to place one variable on the left-hand side and use the others as regressors. In practice, it is possible to find that one regression shows that variables are cointegrated while the reversed order exhibits no cointegration. To avoid these problems, Johansen maximum likelihood procedure is utilized.

To begin, consider an unrestricted n equation vector autoregression (VAR) involving up to k lags of $\mathbf{x}_t$:

$$\mathbf{x_t} = \mathbf{A_t x_{t-1}} + \ldots + \mathbf{A_k x_{t-k}} + \boldsymbol{\varepsilon}_\mathbf{t} \quad ;\boldsymbol{\varepsilon}_\mathbf{t} \sim \mathbf{IN(0,\Sigma)} \qquad \text{(c.1)}$$

Where, for example, in our study,

$$\mathbf{x_t} = \begin{bmatrix} \mathbf{F_t} \\ \mathbf{G_t} \\ \mathbf{Z_t} \end{bmatrix}$$

where $\mathbf{F_t}$ is the composite indicator of financial development

$\mathbf{G_t}$ is the real GDP per capita in term of natural log

$\mathbf{z_t}$ is the control variable which includes saving, investment, trade openness and interest rate.

$\mathbf{A_i}$ is an (3×3) matrix of parameters ; i = 1,…,k, and

$$\Sigma = \begin{bmatrix} \sigma_F^2 & 0 & 0 \\ 0 & \sigma_G^2 & 0 \\ 0 & 0 & \sigma_Z^2 \end{bmatrix}$$

This demonstrates how to estimate the dynamic relationship among jointly endogenous variables without imposing strong a priori restrictions. The system is in reduced form with variable in $\mathbf{x_t}$ regressed on only lagged values of both itself and all the other variables in the system ; therefore, the right-hand side of each equation in the system comprises a common set of predetermined regressors.

Equation (c.1) can be reformulated into a vector error-correction (VEC) form as

$$\mathbf{\Delta x_t} = \mathbf{\Gamma_1 \Delta x_{t-1}} + \ldots + \mathbf{\Gamma_{k-1} \Delta x_{t-k+1}} + \mathbf{\Pi x_{t-k}} + \boldsymbol{\varepsilon}_\mathbf{t} \qquad \text{(c.2)}$$

$$\text{where } \begin{matrix} \Gamma_i = -(I - A_1 - \ldots - A_i) \\ \Pi = -(I - A_1 - \ldots - A_k) \end{matrix} \quad ;i = 1,\ldots,k-1$$

The system contains information on both the short and long-run adjustment to changes in $\mathbf{x_t}$, via the estimates of Γ_i and Π respectively. As will be seen, $\Pi = \alpha\beta'$ when $\boldsymbol{\alpha}$ represents the speed of adjustment to disequilibrium while $\boldsymbol{\beta}$ is a matrix of long-run coefficients. The term $\boldsymbol{\beta}'\mathbf{x}_{t-k}$ in (c.2) represents up to (n-1) cointegration relationships in the multivariate model which ensure that $\mathbf{x}_t$ converges to their long-run steady-state solutions.

Hence, the rank of matrix $\boldsymbol{\Pi}$ is equal to the number of independent cointegrating vectors. If rank of $\boldsymbol{\Pi}$ is zero, the matrix is null, and thus equation (c.2) is the VAR model in first differences and contains no long-run elements. Instead, if $\boldsymbol{\Pi}$ is of rank n, the vector process is stationary. In intermediate case which $1 \leq \text{rank of } \Pi \text{ (or r)} \leq n-1$, there exists (r) cointegration vector in $\boldsymbol{\beta}$ together with (n-r) non-stationary vectors. That is, r columns of $\boldsymbol{\beta}'$ attribute r linearly independent combinations of the variables in $\mathbf{x}_t$ that are stationary while n-r columns of $\boldsymbol{\beta}'$ are I(1). As a result, testing for cointegration is equivalent to the test for the rank of **Π.** Johansen (1998) proposed an estimate of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ using the procedure known as 'reduced rank regression'.

The maximum likelihood estimate of $\boldsymbol{\beta}$ is obtained as the eigenvectors corresponding to their largest eigenvalues from solving the characteristic equation.

$$\left|\lambda S_{kk} - S_{k0}S_{00}^{-1}S_{0k}\right| = 0 \qquad \text{(c.3)}$$

Equation (c.3) gives the eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > ... > \hat{\lambda}_n$ and the corresponding eigenvector $\hat{\mathbf{V}} = (\hat{V}_1 ... \hat{V}_n)$. Those r elements in $\mathbf{V}$ which determine the linear combinations of stationary relationships can be denoted $\hat{\boldsymbol{\beta}} = (\hat{v}_1, ..., \hat{v}_r)$, i.e., these are the cointegration vectors.

Therefore, a test procedure for reduced rank is to test the null hypothesis that there are at most r cointegration vector as follows:

$$H_0 : \lambda_i = 0 \qquad ; i = r+1, ..., n$$

where only the first r eigenvalues are non-zero.

This restriction can be imposed for different value of r, and then the log of the maximized likelihood function for the restricted model is compared to the log of the maximized likelihood function of the unrestricted model and standard likelihood ratio test computed. This means that it is possible to test the null hypothesis using the 'trace statistic' with the null hypothesis of at most r cointegrating vectors against the alternative of more than r cointegrating vectors;

$$\lambda_{trace}(r) = -2\log(Q)$$

$$= -T\sum_{i=r=1}^{n}\log(1-\hat{\lambda}) \qquad ;r = 0,1,2,....n-2,n-1$$

where Q = (the ratio of restricted maximized likelihood to unrestricted maximized likelihood).

Another test of the significance of the largest eigenvalue is known as 'mazimal eigenvalue' or λ_{max} statistic with the alternative hypothesis of r+1 cointegrating vectors.

$$\lambda_{max}(r, r+1) = -T\log(1-\lambda_{r+1}) \qquad ;r = 0,1,2,...,n-2,n-1$$

It should be emphasized that neither of the likelihood ratio (LR) statistic, i.e., λ_{trace} and λ_{max} have standard distribution. Thus, Johansen and Juselius (1990) have provided appropriated critical values for these two tests.

Appropriate lag length in Johansen analysis in the VAR system is chosen by utilizing the multivariate generalizations of the AIC as follows:

$$AIC = T\log|\Sigma| + 2N$$

where T = the number of usable observations

$|\Sigma|$ = determinant of the variance / covariance matrix of the residuals

N = total number of performance estimated in all equations.