Economics Bulletin

Volume 34, Issue 1

What results can we expect from rolling trace tests? A discussion based on the issue of stock market integration

Alexander Ludwig
TU Dresden

Abstract

This paper discusses pitfalls in the application of the rolling trace test. This procedure is based on the iterative calculation of Johansen's (1988) trace test for the rank of a cointegration system in windows of equal length that roll over the sample. Pitfalls lie in the selection of the window length and of the lag order for short-run coefficients as well as in the presence of stationary variables in some sub-periods. We give practical recommendations to solve these issues and demonstrate their implications when assessing the integration of four major European stock markets.

I wish to thank Maike Kronbichler for valuable research support. Financial support by the PhD program of zeb/rolfes.schierenbeck.associates is gratefully acknowledged. The usual disclaimers apply.
Citation: Alexander Ludwig, (2014) ''What results can we expect from rolling trace tests? A discussion based on the issue of stock market integration'', *Economics Bulletin*, Vol. 34 No. 1 pp. 16-24.
Contact: Alexander Ludwig - alexander.ludwig@mailbox.tu-dresden.de.
Submitted: May 15, 2013. **Published:** January 06, 2014.

1. Introduction

Rolling cointegration tests are an often-used econometric technique to examine the time-varying behavior of cointegration relationships.[1] To this end, most researchers apply the rolling trace test, see e.g. Mylonidis and Kollias (2010) and Peri and Baldi (2013) for recent contributions. The trace test was proposed by Johansen (1988) to estimate the cointegration rank in vector error correction models (VECM). Rolling trace statistics enable the empirical researcher to graphically display the evolution of the cointegration rank. In a system of n variables, a cointegration rank r with $0 < r < n$ implies the existence of r long-run relationships between the variables. The larger r, the smaller is the number of common stochastic trends which equals $n - r$, and the more interconnected the variables are. In the case $r = n-1$, there is only one common stochastic trend which influences the development of all series. This is relevant for the practitioner as it implies an absence of diversification possibilities in the long-run. Using daily data of stock market indices (Mylonidis and Kollias, 2010) as well as diesel and oil prices (Perri and Baldi, 2013) and applying rolling windows of 250 days, these authors obtain only few periods with a cointegration rank larger than 0 and smaller than n.

This paper shows that the use of the rolling trace statistics requires clarification of several issues before and during its iterative application. First, we recapitulate that the power of Johansen's (1988) trace test is low when the autocorrelation of regression residuals is high (Section 2). Residuals from regressions of the same stock market indices as in Mylonidis and Kollias (2010) onto each other reveal first order autocorrelation coefficients larger than 0.95. We show that a window length of at least four years should be used in this case.

Second, inference from the trace test depends, in general, on the selection of a lag parameter when setting up the VECM. We demonstrate its impact on the results when using the stock market indices and recommend the derivation of period-specific lag orders which guarantee the absence of autocorrelated errors in each period (Section 3). Third, the rolling trace test is only reliable when the time series are non-stationary in each period. We perform rolling stationary tests and recommend the exclusion of variables if they are found to be stationary in a specific period in order to preclude distortions in the estimation of the cointegration rank (Section 4).

In sum, we provide new evidence for the time-dependent interconnectedness of the main stock market indices of Spain, Italy, France and Germany and give reasons for its evolution (Section 5).

2. Window length and power of the trace test

The trace test for the null of a cointegration rank smaller or equal to r

$$\lambda_{tr} = -T * \sum\nolimits_{i=r+1}^{n} \ln(1 - \hat{\lambda}_i) \tag{1}$$

incorporates the smallest $n - r$ eigenvalues $\hat{\lambda}_i$ from Johansen's (1988) maximum likelihood

[1]Rolling cointegration tests have been applied to various economic issues, e.g. stock market convergence (Awokuse et al., 2009, Pascual, 2003, Yu et al., 2010), real and monetary convergence (Brada et al., 2010), the relationship between money and real income (Swanson, 1998), exchange rate convergence (Rangvid and Sorensen, 2002) and the efficiency of future contracts (Haigh, 2000).

(ML) approach to estimate the parameters of a VECM for an n-dimensional process X_t:

$$\Delta X_t = \mu_t + \Pi X_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta X_{t-j} + u_t \tag{2}$$

Here, Π is an $n \times n$-matrix of rank r, μ_t is a vector of deterministic terms and Γ_i are $n \times n$-matrices. Cointegration is present if and only if $0 < r < n$ holds. If Π has such a reduced rank, it can be written as $\Pi = \alpha\beta'$ where α and β are $n \times r$-matrices. The matrix β contains cointegration coefficients leading to the residuals $e_{t-1} = \beta' X_{t-1}$ of the cointegration equation (deviations from equilibrium). The power of the trace test depends on the persistence (autocorrelation) of these residuals. To illustrate this, we model the same data generating processes as in the seminal works of Banerjee et al. (1986), Engle and Granger (1987) and Gregory and Hansen (1996):

$$\begin{aligned} X_{1t} + X_{2t} &= v_t & v_t(1-L) &= \epsilon_{1t} \\ X_{1t} + \beta_t X_{2t} + \alpha_t &= w_t & w_t(1-\rho L) &= \epsilon_{2t} \end{aligned} \tag{3}$$

where ϵ_{it} are independently $N(0,1)$-distributed random variables and L denotes the lag operator.[2] If the autocorrelation coefficient ρ is smaller than one, there is a linear combination of X_{1t} and X_{2t} such that the two series are cointegrated, i.e. residuals w_t are stationary. If $\rho = 1$, X_{1t} and X_{2t} are not cointegrated. We set $\beta_t = 2$ and $\alpha_t = 1$. The selection of these parameters does not influence the results.

Tables 1 and 2 give the frequencies of acceptance of cointegration, i.e. when the null $r = 0$ is rejected whereas the null $r \leq 1$ is not. Finite sample size-corrected critical values by Doornik (1998) at the $\alpha = 5\%$ level (Table 1) and $\alpha = 10\%$ level (Table 2) are used. As we include the condition that the null $r \leq 1$ has to be accepted, frequencies are exptected to converge to $1 - \alpha$ for growing sample sizes. On average over all replications, information criteria suggest a lag order $p = 1$ in equation (2), whereas three lags are needed in order to eliminate autocorrelation of the errors u_t in equation (2). Since the results are very similar for both choices of the lag order, we only give the results for a lag order of one. All experiments are replicated 10,000 times.

Table 1: Cointegration acceptance frequencies of the trace test ($\alpha = 5\%$)

ρ	$T = 250$	$T = 500$	$T = 750$	$T = 1000$	$T = 1500$	$T = 2000$
0.500	0.949	0.951	0.947	0.947	0.953	0.946
0.900	0.494	0.947	0.952	0.952	0.947	0.949
0.950	0.136	0.474	0.847	0.947	0.943	0.945
0.975	0.051	0.129	0.272	0.477	0.847	0.942
0.990	0.046	0.050	0.068	0.092	0.183	0.311
0.999	0.044	0.042	0.041	0.040	0.038	0.037

[2] Later we consider a system of $n = 4$ stock markets. Without loss of generality, we can assume $n = 2$ in this experiment as one of the series can also represent a non-stationary linear combination of two or more other series.

Table 2: Cointegration acceptance frequencies of the trace test ($\alpha = 10\%$)

ρ	$T = 250$	$T = 500$	$T = 750$	$T = 1000$	$T = 1500$	$T = 2000$
0.500	0.902	0.901	0.897	0.903	0.898	0.904
0.900	0.576	0.899	0.899	0.902	0.901	0.900
0.950	0.206	0.592	0.860	0.897	0.904	0.903
0.975	0.098	0.207	0.389	0.603	0.862	0.897
0.990	0.082	0.094	0.122	0.150	0.277	0.434
0.999	0.088	0.079	0.077	0.076	0.075	0.074

Next, we use daily stock market indices of Spain, Italy, France and Germany for the period 04.01.1999 until 31.12.2012 which includes the data from Mylonidis and Kollias (2010) as a subset. Data are taken from Datastream and transformed to logs. Regressing each stock market index on the remaining three indices via OLS, estimates for the first order autocorrelation coefficient ρ turn out to be 0.977 for Italy, 0.993 for Spain, 0.977 for France, and 0.978 for Germany. Using ML, the estimate for ρ is 0.977 and does not depend on the country since all elements of X_t in equation (2) enter the ML estimation symmetrically. In the presence of such autocorrelation coefficients, the power of the trace test is as low as the nominal size for a window length of 250 days and becomes acceptable for windows of at least 1,000 days (see Tables 1 and 2).

In the following, we apply the rolling trace test to the four stock market indices and select window lengths of 250, 500, 1000 and 1500 days and an increment of 10 trading days between two successive windows. We use the lag order $p = 2$ as suggested by the Schwarz information criterion and include a deterministic trend in the cointegration relation to account for varying trend slopes in the stock indices. Figure 1 shows that for a window of 250 days, there are relatively few periods with cointegration rank of at least one. This is in line with the findings of Mylonidis and Kollias (2010). Increasing the window length, which leads to an increase of the power of the test, we obtain a considerable decrease in the number of periods with cointegration rank 0. For the window length of 1,500 days, there seems to be at least one cointegration relation during the entire sample. We will keep the window length at 1,500 days in the subsequent analysis.[3]

Reasons for the evolution of the cointegration rank and hence for the varying degree of interconnectedness of these four stock markets are discussed in more detail in Section 5.

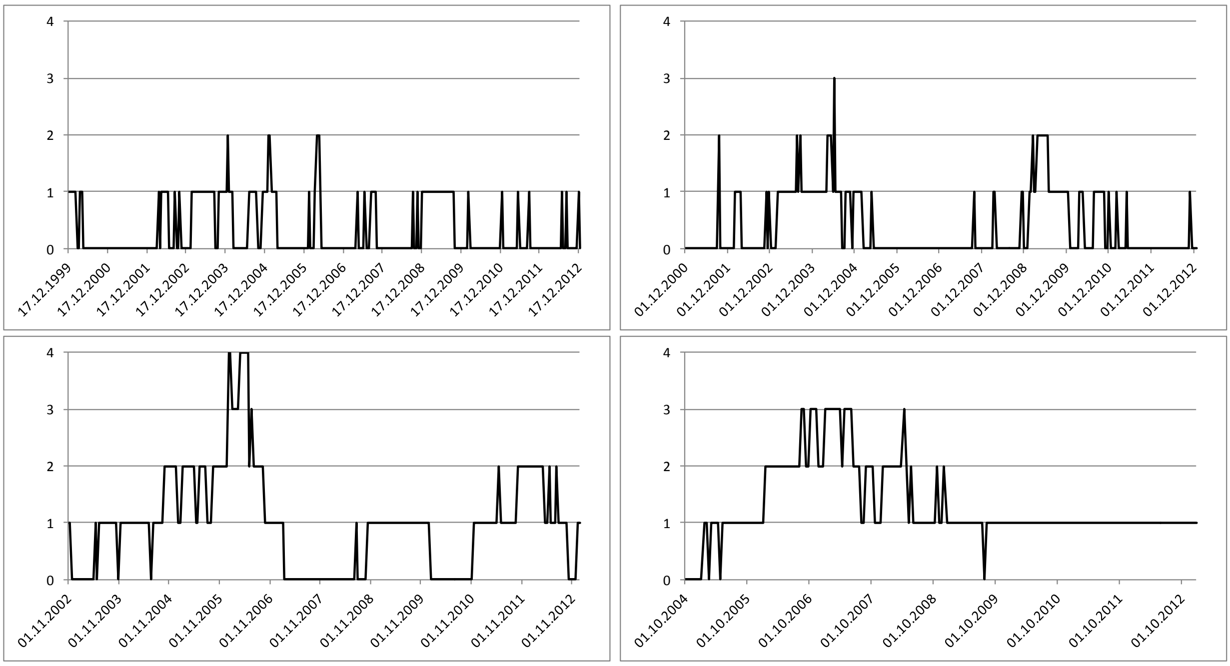
3. Lag order dependence

Information criteria usually inform the decision on the lag parameter p in equation (2). However, using information criteria may be inappropriate. Cheung and Lai (1993) show that standard information criteria do not perform well in finding the correct lag number p and tend to favor under-parameterization. In this case, it often turns out that the errors u_t in equation (2) still exhibit serial correlation, a model mis-specification that leads to incorrect test statistics.

Hence, we opt for a general-to-specific approach in each window period, i.e. we decrease the number of lags sequentially as long as there is no serial correlation in the errors. We apply the Ljung-Box Q statistics with orders 1 to 20 to test for serial correlation. If, for a given lag order, an autocorrelation coefficient of zero is rejected at the 10% level, we stop

[3] Gretl code for all calculations in this paper is provided by the author upon request.

Figure 1: Cointegration rank estimates (dependence on window length)



Note: from the upper left to the lower right panel, the graphs contain the results for the rolling trace test with a window length of T=250, 500, 1,000 and 1,500 days. The lag parameter is set to 2 in all cases. We use the 10% critical value for the trace test.

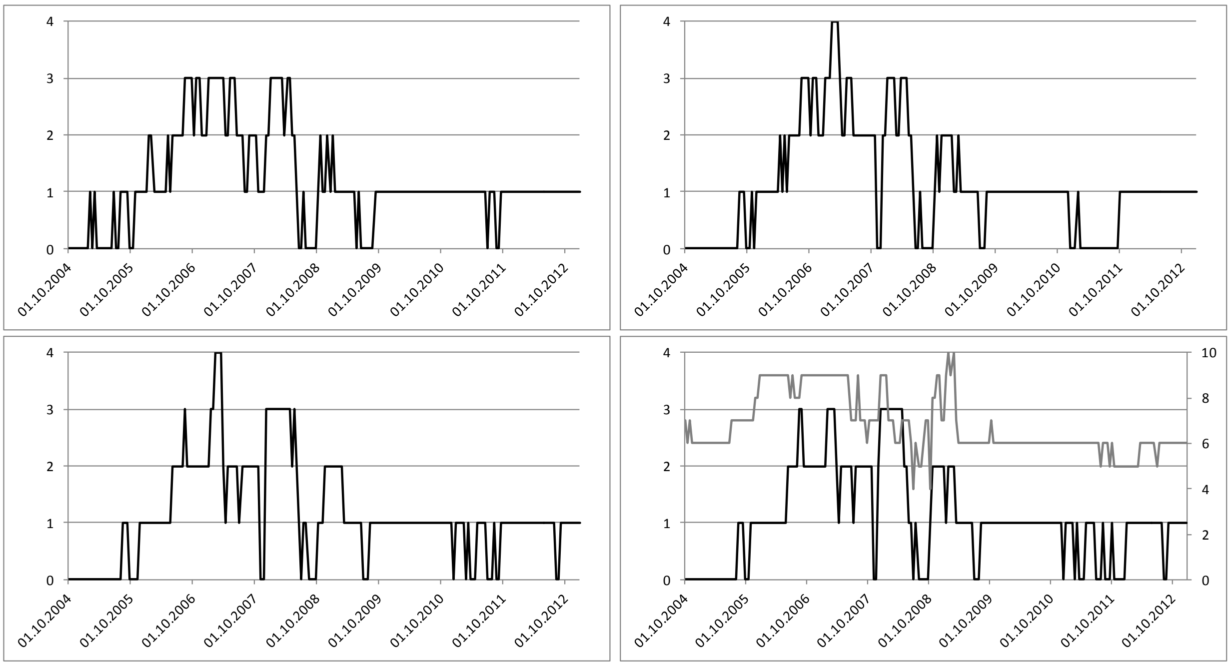
this sequence and set the lag parameter as the current parameter plus one. It turns out that setting a maximum number of 10 lags is sufficient for all periods.

Figure 2 shows that an increase of the lag parameter from 2 to 3 leads to relatively few changes in the graph of rolling trace tests whereas a further increase to 4 and 6 lags does. Moreover, despite the non-stationarity of all series in windows of 1,500 days (see Figure 3), the cointegration rank is estimated to be 4 in periods ending mid-2007 (see the upper right and the lower left panels of Figure 2). These results are conflicting as the latter implies the stationarity of all series and the appropriateness of a VAR model in levels. This contradiction is likely to be a result of mis-specifying the lag parameter. In the lower right panel of Figure 2, the grey graph gives the data-dependent lag parameter estimated for each single window by the general-to-specific approach. Until 2009, lag parameters that guarantee the absence of autocorrelation of errors were usually larger than 6. Taking this into account, the cointegration rank turns out to be always smaller than 4 as expected given the non-stationarity of the four stock market series.

4. Distortions by stationary variables

According to Engle and Granger (1987), two non-stationary variables are cointegrated if a linear combination of them is stationary. Since any linear combination of two stationary variables is again stationary, cointegration tests will reject the null of a cointegration rank of zero in this case as well. Further, if there are already two cointegrated variables in a system of $n > 2$ variables and there is a stationary variable among the remaining variables, any linear combination of the stationary variable and the residuals from the cointegration relationship is again stationary leading to an increase of the cointegration rank. Hence, an increase of the cointegration rank can result solely from the transition of a variable from

Figure 2: Cointegration rank estimates (dependence on lag parameter)



Note: from the upper left to the lower right panel, the graphs contain the results for the rolling trace test with a lag parameter of 3, 4 and 6 and a period-specific parameter that guarantees the elimination of serially correlated errors. In the latter case, the optimal lag parameter is given by the grey line. The window length is set to be 1,500 in all cases.

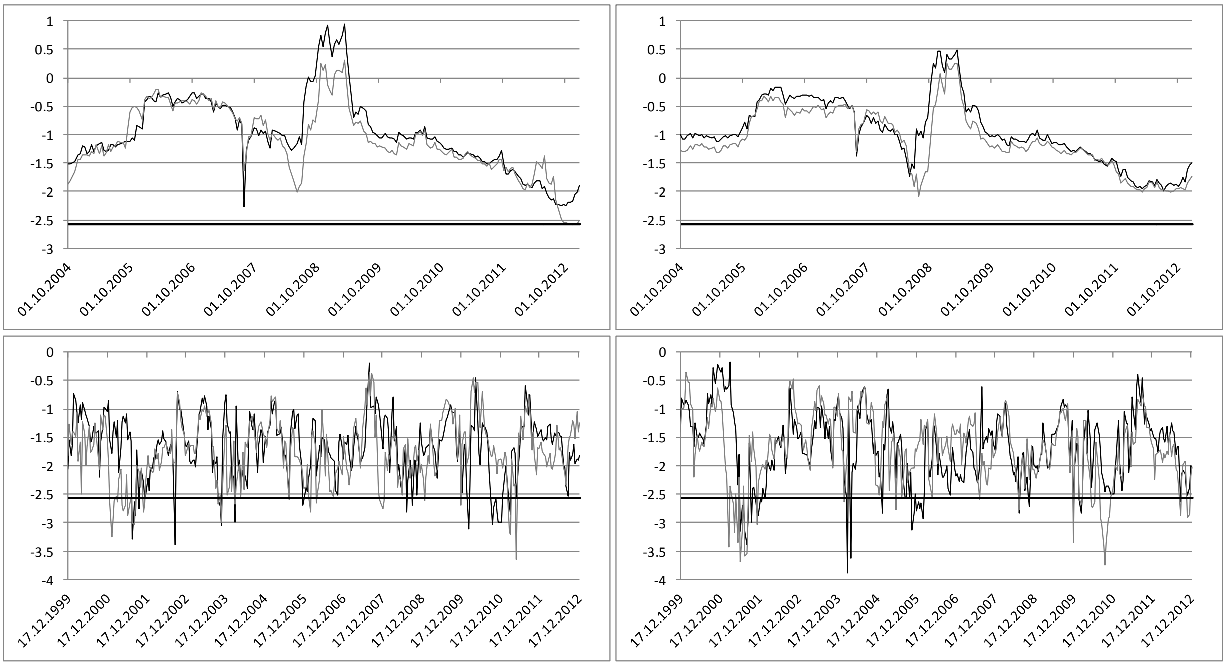
non-stationary to stationary in the course of time. Therefore, we propose to window-wise exclude stationary variables from the estimation of the cointegration rank.

To examine the stationarity of the four stock market series, we apply the ADF-GLS test proposed by Elliot et al. (1996) which has better power properties than the standard ADF test. We allow for a trend in each window and test down from a lag length of 10 to avoid wasting power of the test, i.e. we reduce the lag number until the last lag variable becomes significant at the 10% level. Using windows of 1,500 days, all four series are non-stationary at all conventional levels except for Spain in the window ending 16 November 2012 (see the upper two panels of Figure 3). Using this window size, disturbing effects on the results of rolling trace tests are thus limited to one date. In other words, the left panel of Figure 4 only differs from the lower right panel in Figure 2 in that the cointegration rank is 0 instead of 1 for the window ending 16 November 2012.

For the purpose of illustration, we also performed the stationarity test for windows of 250 days (see the lower two panels of Figure 3). Stationarity of the time series considered emerges more often in this case and the values of the test statistic are more volatile. This yields another argument for the selection of a larger window in addition to the higher power of the trace test for larger windows.

As the results of the test also depend on the critical level chosen, we report cointegration ranks at the 5% level as well (see the right panel in Figure 4). Key differences between the results at the 10% and 5% level only emerge for the two most recent years. Here, the cointegration rank of 1 cannot be supported at the 5% level. In the following, we focus on the results at the 10% level.

Figure 3: Rolling unit root test



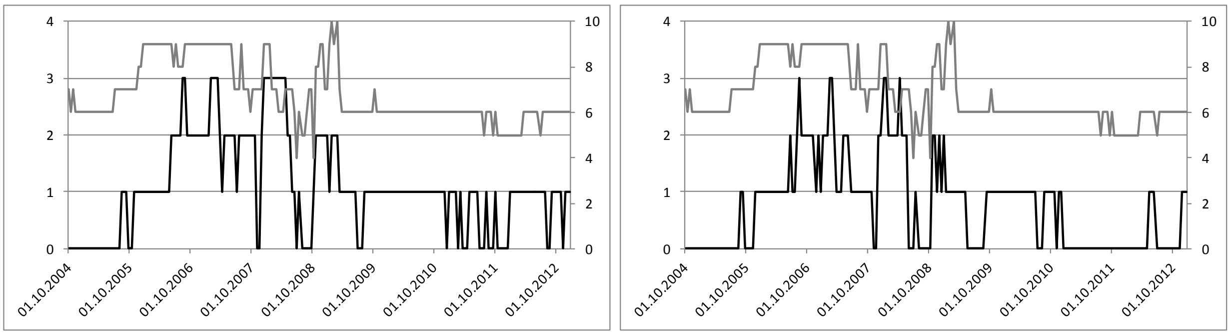
Note: the upper two panels correspond to the window size 1,500, the lower two to the window size 250. In the left panels, the black line refers to Italy and the grey line to Spain. In the right panels, the black line refers to France, the grey line to Germany. The panels give the GLS-detrended ADF-test statistic proposed by Elliot et al. (1996). Critical values are -2.57 (straight line in each panel), -2.89 and -3.48 at the 10%, 5% and 1% level, respectively.

5. Time-dependence of stock market integration

The left panel of Figure 4 indicates that there is usually a cointegration rank of at least 1 at the 10% level since August 2005. For the windows ending between October 2004 and August 2005, no cointegration between stock markets is observable. The joint recovery of equity markets that started in 2003 is reflected first in windows ending in August 2005. A period of economic growth lasting until 2007 was accompanied by an increase of the integration of major European stock markets. Until August 2006, the cointegration rank increased from 0 to 3 implying full integration of these four markets for some periods. Until October 2007, the cointegration rank ranged between 1 and 3 but a sharp break occurred on 9 November 2007 when the cointegration rank fell to 0. As a consequence of the US subprime mortgage crisis, the confidence of banks in each other's solvability decreased sharply leading to the breakdown of the interbank lending market and turmoil on the financial market in the second half of 2007. Large downturns in equity prices followed and the interconnectedness of stock market indices rose again with a cointegration rank equaling 3 as all markets suffered from similarly intensive losses. After April 2008, the cointegration rank fell steadily again to 0 which can be attributed to country-individual crises of the real economies that started in 2008. In Spain, for example, the burst of the house price bubble was at its peak in mid-2008 (see e.g. Eurostat's construction output index for building construction). In Italy, output deterioration already started in the second quarter of 2008 as opposed to the last quarter of 2008 in Germany, Spain and France.

The insolvency of Lehman Brothers in September 2008 supported stock markets' tendencies to move jointly. This led to an increase of the cointegration rank to 2 at the end of 2008 and beginning of 2009. Since February 2009, the cointegration rank has usually been 1

Figure 4: Cointegration rank estimates (adjusted for presence of stationarity) at the 10% and 5% level



Note: the left panel contains the results from a rolling trace test with window length 1,500 at the 10% level, the right panel at the 5% level. Optimal lag parameters are given by the grey lines. In both cases, the cointegration rank was determined from the set of non-stationary variables only.

showing tendencies for co-movement of stock markets. However, there are also periods with no cointegration at all (especially at the 5% critical level) which reflects the relevance of country-individual issues for the behavior of stock prices.

References:

Awokuse, T.O., Chopra, A., and D.A. Bessler (2009) “Structural change and international stock market interdependence: Evidence from Asian emerging markets” *Economic Modelling* **26**, 549-559.

Banerjee, A., Dolado, J.J., Hendry, D.F. and G.W. Smith (1986) “Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence” *Oxford Bulletin of Economics and Statistics* **48**, 253-277.

Brada, J.C., Kutan, A.M., and S. Zhou (2005) “Real and monetary convergence between the European Union’s core and recent member countries: A rolling cointegration approach” *Journal of Banking and Finance* **29**, 249-270.

Cheung, Y., and K.S. Lai (1993) “Finite-sample sizes of Johansen’s likelihood ratio tests for cointegration” *Oxford Bulletin of Economics and Statistics* **55**, 313-328.

Doornik, J. (1998) “Approximations to the asymptotic distribution of cointegration tests” *Journal of Economic Surveys* **12**, 573-593.

Elliott, G., Rothenberg, T.J. and J. H. Stock (1996) “Efficient tests for an autoregressive unit root” *Econometrica* **64**, 813-836.

Engle, R.F. and C.W.J. Granger (1987) “Co-integration and error correction: representation, estimation and testing” *Econometrica* **55**, 251-276.

Gregory, A. and B. Hansen (1996) “Residual-based tests for cointegration in models with regime shifts” *Journal of Econometrics* **70**, 99-126.

Haigh, M.S. (2000) “Cointegration, unbiased expectations, and forecasting in the BIFFEX freight futures market” *The Journal of Futures Markets* **20**, 545-571.

Johansen, S. (1988) “Statistical analysis of cointegrating vectors” *Journal of Economic Dynamics and Control* **12**, 213-254.

Mylonidis, N., and C. Kollias (2010) “Dynamic European stock market convergence: Evidence from rolling cointegration analysis in the first euro-decade” *Journal of Banking and Finance* **34**, 2056-2064.

Pascual, A.G. (2003) “Assessing European stock markets (co)integration” *Economics Letters* **78**, 197-203.

Peri, M., and L. Baldi (2013) "The effect of biofuel policies on feedstock market: Empirical evidence for rapeseed oil prices in EU" *Resource and Energy Economics* **35**, 18-37.

Rangvid, J., and C. Sorensen (2002) "Convergence in the ERM and Declining Numbers of Common Stochastic Trends" *Journal of Emerging Market Finance* **1**, 183-213.

Swanson, N.R. (1998) "Money and output viewed through a rolling window" *Journal of Monetary Economics* **41** 455-473.

Yu, I.-W., Fung, K.-P., and C.-S. Tam (2010) "Assessing financial market integration in Asia – Equity markets" *Journal of Banking and Finance* **34** 2874-2885.