The Power of Residual–Based Tests for Cointegration when Residuals Are Fractionally Integrated[1]

by

Walter Krämer and Francesc Marmol

Fachbereich Statistik, Universität Dortmund, D-44221 Dortmund, Germany

Departamento de Estadistica y Econometria, Universidad Carlos III
38903, Getafe, Madrid, Spain

Version October 1998

1 Introduction

This paper is concerned with testing the null hypothesis of no cointegration among $I(1)$–variables when the cointegration residuals are $I(d)$ with $0 < d < 1$. This possibility is entertained with increasing frequency in many applications (see e.g. Cheung and Lai 1993, Baillie and Bollerslev 1994, Booth and Tse 1995 or Baillie 1996 for examples). We consider the power of various cointegration tests both for the stationary case ($d < 0.5$) and for the nonstationary case ($d \geq 0.5$).

When the potential cointegrating relationship is known, this problem boils down to testing for unit roots against fractional alternatives, as discussed by e.g. Sowell (1990), Diebold and Rudebusch (1991), Hassler and Wolters (1994), Dolado and Marmol (1997) or Krämer(1998). When the potential cointegrating relationship has to be estimated, we encounter the twin problems of nonstandard regression properties due to $I(d)$–disturbances and unobservability of the true residuals. While the second problem has been solved for the case where

[1]Research supported by Deutsche Forschungsgemeinschaft and Volkswagenstiftung; we are grateful to Ingolf Dittmann for helpful criticism and comments, and to Deutsche Finanzdatenbank (DFDB) for providing us with the data for our empirical results

the true residuals are $I(0)$ (Phillips/Ouliaris 1990), an analoguous analysis of the case where the true residuals are $I(d)$ with $d < 1$ is still missing.

A related problem is the power of tests of the null hypothesis that cointegration exists, against the alternative of fractionally integrated residuals. Again, this problem has only been addressed for the case where true residuals are used (Lee and Schmidt 1996, Marmol 1997).

Below we confine ourselves to testing the null hypothesis of <u>no</u> cointegration.

2 The Model and the Tests

Let $\{z_t\}, t = 0, 1, 2, \ldots$ be the m–vector integrated process under test, generated according to

$$z_t = z_{t-1} + \xi_t \quad (t = 1, 2, \ldots). \tag{1}$$

As regards to this and subsequent notation and also as regards assumptions, we follow Phillips and Ouliaris (1990). In particular, let z_0 without loss of generality be zero. The innovations ξ_t in (1) are assumed to have mean zero and to satisfy a multivariate invariance principle

$$X_T(r) := \frac{1}{\sqrt{T}} \sum_{t=1}^{[T_r]} \xi_t \xrightarrow{d} B(r), \tag{2}$$

where $B(r)$ is an m–vector Brownian motion with covariance matrix

$$\Omega := \lim_{T\to\infty} \frac{1}{T} E\left\{ \left(\sum_{t=1}^{T} \xi_t \right) \left(\sum_{t=1}^{T} \xi_t' \right) \right\}. \tag{3}$$

Under the null hypothesis of no cointegration, Ω has full rank m.

Below we consider the alternative that there is exactly one cointegrating relationship, i.e. that the z–vector can be split into

$$z_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix} \begin{matrix} 1 \\ n \end{matrix}, \quad m = n + 1 \tag{4}$$

such that

$$y_t = \beta' x_t + u_t, \tag{5}$$

where u_t is $I(d)$ with $d < 1$. This generalizes Phillips and Ouliaris (1990), who consider the case where under the alternative u_t is $I(0)$.

Given that the data follow (5), we consider the following tests of the null hypothesis of no cointegration (i.e. $u_t \sim I(1)$):

<u>Augmented Dickey Fuller (ADF):</u>
The t–statistic for $\alpha = 0$ in the regression

$$\Delta \hat{u}_t = \alpha \hat{u}_{t-1} + \varphi_1 \Delta \hat{u}_{t-1} + \ldots + \varphi_p \Delta \hat{u}_{t-p} + v_{tp}, \tag{6}$$

where the $\hat{u}_t$ are OLS–residuals from (5).

<u>Phillips' $\hat{Z}_\alpha$:</u>

$$\hat{Z}_\alpha = T(\hat{\alpha} - 1) - \frac{1}{2} \frac{S_{T\ell}^2 - S_k^2}{\frac{1}{T^2} \sum_{t=2}^{T} \hat{u}_{t-1}^2}, \tag{7}$$

where $\hat{\alpha}$ is from the regression $\hat{u}_t = \hat{\alpha} \hat{u}_{t-1} + \hat{k}_t$ and where

$$S_k^2 = \frac{1}{T} \sum_{t=1}^{T} \hat{k}_t^2, \tag{8}$$

$$S_{T\ell}^2 = \frac{1}{T} \sum_{t=1}^{T} \hat{k}_t^2 + \frac{2}{T} \sum_{s=1}^{\ell} w_{s\ell} \sum_{t=s+\ell}^{T} \hat{k}_t \hat{k}_{t-s}, \tag{9}$$

$$w_{s\ell} = 1 - \frac{s}{\ell + 1}. \tag{10}$$

The extra term on the right in (7) takes care of nuisance parameters that would otherwise affect the limiting rejection probability under H_0: The limiting distribution of the standard Dickey–Fuller statistic $T(\hat{\alpha} - 1)$ depends on the correlation structure of the residuals, and this dependency is thereby (asymptotically) removed (see Hamilton 1994, chapter 17.6 for a didactical exposition of this issue).

Phillips' $\hat{Z}_t$:

$$\hat{Z}_t := \frac{\hat{\alpha}-1}{\sqrt{\frac{S^2_{T\ell}}{\sum_{t=2}^{T}\hat{u}^2_{t-1}}}} - \frac{1}{2}\frac{S^2_{T\ell}-S^2_k}{\sqrt{S^2_{T\ell}\times\frac{1}{T^2}\times\sum_{t=2}^{T}\hat{u}^2_{t-1}}}. \tag{11}$$

Again, the second term is added to remove the dependency of the limiting null distribution on the correlation structure of the residuals.

The Phillips/Ouliaris variance ratio test:

$$\hat{P}_u := \frac{T\hat{\omega}_{11.2}}{\frac{1}{T}\sum_{t=1}^{T}\hat{u}^2_t}, \tag{12}$$

where $\hat{\omega}_{11.2} = \hat{\omega}_{11} - \hat{\omega}'_{21}\hat{\Omega}^{-1}_{22}\hat{\omega}_{21}$ and

$$\hat{\Omega} = \frac{1}{T}\sum_{t=1}^{T}\hat{\xi}_t\hat{\xi}'_t + \frac{1}{T}\sum_{s=1}^{\ell}w_{s\ell}\left(\sum_{t=s+1}^{T}\hat{\xi}_t\hat{\xi}'_{t-s}+\hat{\xi}_{t-s}\hat{\xi}'_t\right) \tag{13}$$

and where the $\hat{\xi}_t$ are the residuals from the least squares regression

$$z_t = \hat{\Pi}z_{t-1} + \hat{\xi}_t. \tag{14}$$

The Phillips/Ouliaris multivariate trace statistics:

$$\hat{P}_z = Ttr\left(\hat{\Omega}M^{-1}_{zz}\right), \tag{15}$$

where $M_{zz} = \frac{1}{T}\sum_{t=1}^{T} z_t z'_t$.

The $\hat{Z}_\alpha$- and $\hat{Z}_t$-tests can be viewed as generalizations of the standard Dickey–Fuller tests based on either $T(\hat{\alpha}-1)$ or on the standard t–statistic for $H_0 : \alpha = 1$ in the regression

$$u_t = \alpha u_{t-1} + k_t. \tag{16}$$

The variance-ratio and multivariate trace-statistic tests explore the relationship between direct and indirect estimates of the conditional variance of y_t given x_t, along the lines of Hausman (1978): Under the null hypothesis of no cointegration, both estimates are close together, but they diverge when there is cointegration.

Sowell (1990) has shown that $\hat{\alpha}-1$ from the regression (16) is $O_p(T^{1-2d})$ when the u_t's are $I(d)$, $(0.5 < d < 1)$ so

$$T(\hat{\alpha}-1) = O_p(T^{2-2d}), \tag{17}$$

i.e. the Dickey–Fuller–test is consistent as it diverges under the alternative, albeit much slower than when the disturbances in (16) are $I(0)$. Similarly, the Dickey–Fuller–t–test diverges, again much slower than under the $I(0)$–alternative.

Below we extend these results to the case where estimated rather than true residuals are used, and where the tests account for the fact that there is autocorrelation among the k_t's from (16) under the null hypothesis.

3 Divergence Rates under Nonstationary Alternatives

THEOREM 1: *Under the assumptions specified in Section 2, when the residuals u_t in $y_t = \beta' x_t + u_t$ are $I(d)$ with $0.5 < d < 1$, we have the following divergence rates of the various test–statistics:*[2]

(i) $\hat{Z}_\alpha = O_p(T^{2-2d})$,

(ii) $\hat{Z}_t = O_p(T^{1-d})$,

(iii) $ADF = O_p(T^{1-d})$,

(iv) $\hat{P}_u = O_p(T^{2-2d})$,

(v) $\hat{P}_z = O_p(T^{2-2d)}$,

REMARK: The theorem shows that the tests remain consistent, but the divergence rates are smaller in the context of fractional cointegration, so conventional tests will often fail to pick it up. Also, the relative differences established by Phillips and Ouliaris (1990, Theorem 5.1 and 5.2) remain: the $\hat{Z}_t$ and ADF-tests diverge still slower than the rest, which explains why there tests are particularly poor in detecting fractional cointegration in the empirical example discussed in section 5.

PROOF OF THEOREM: For ease of exposition and notation, we initially confine ourselves to bivariate systems, i.e. to the case $n = 1$ and $m = 2$. Then β and x_t in (5) are scalars, and we have in obvious notation:

$$\hat{\beta} - \beta \quad = \quad \frac{x'u}{x'x} \quad \text{and} \tag{18}$$

[2]Here and elsewhere, "$O_p g(T)$" is taken to imply that $g(T)$ is the largest function of T such that the respective expressions divided by $g(T)$, remain stochastically bounded, but do not tend to zero in probability either. $f(T)/g(T) \to \infty$ implies that the respective expressions are $o_p(f(T))$.

$$\begin{aligned} \hat{u} &:= y - \hat{\beta}x = u - (\hat{\beta} - \beta)x = u - \frac{x'u}{x'x}x \quad \text{where} && (19) \\ \frac{x'u}{x'x} &= O_p(T^{d-1}). && (20) \end{aligned}$$

(see Cheung and Lai 1993, p. 106). This latter relationship implies that, with nonstationary fractional alternatives, we no longer have

$$\hat{u}_{t,T} = u_t + o_p(1), \tag{21}$$

since

$$\hat{u}_{T,T} = u_T - \frac{x'u}{x'x}x_T = u_T + O_p\left(T^{d-\frac{1}{2}}\right). \tag{22}$$

This makes various subsequent derivations rather complicated, since estimated residuals do no longer tend in probability to the true residuals <u>uniformly in t</u>.[3]

In the regression

$$\hat{u}_{t,T} = \hat{\alpha}\hat{u}_{t-1,T} + \hat{k}_t \tag{23}$$

we have

$$\begin{aligned} \hat{\alpha} - 1 &= \frac{\hat{u}'_{-1}(\hat{u} - \hat{u}_{-1})}{\hat{u}'_{-1}\hat{u}_{-1}} \\ &= \frac{\frac{1}{2}\left(\hat{u}^2_{T,T} - \sum_1^T \left(\Delta\hat{u}_{t,T}\right)^2\right)}{\hat{u}'_{-1}\hat{u}_{-1}}, && (24) \end{aligned}$$

[3]Unlike Phillips/Ouliaris (1990), we have added a second subscript to $\hat{u}_t$ in (21) and (22), to highlight the fact that OLS residuals depend on sample size. This dependence on sample size is inconsequential in the case of ARMA–residuals, as then

$$\hat{u}_{t,T} = u_t + O_p(T^{-\frac{1}{2}}) \ \underline{\text{uniformly in t}},$$

(see Phillips/Ouliaris 1990, p. 184), so the subscript T can without danger be omitted. With fractionally integrated residuals, we still have $\hat{u}_{t,T} \xrightarrow{p} u_t$ <u>for any given</u> t, as for given t, $\hat{u}_{t,T} - u_t = O_p(T^{d-1})$), but this convergence is no longer uniform in t.

where

$$\hat{u}_{T,T}^2 = \left(u_T - \frac{x'u}{x'x}x_T\right)^2 = O_p\left(T^{2d-1}\right), \tag{25}$$

$$\sum_1^T(\Delta\hat{u}_{t,T})^2 = O_p(T) \tag{26}$$

and

$$\hat{u}_{-1}'\hat{u}_{-1} = O_p(T^{2d}). \tag{27}$$

Taken together, (24) – (27) imply that under the alternative

$$\hat{\alpha} - 1 = O_p(T^{1-2d}), \tag{28}$$

which is the same convergence rate that obtains when true residuals are used (Sowell 1990).

In the general case where $m > 2$, the simple formula (19) for the cointegration residuals $\hat{u}$ is replaced by $\hat{u} = y - X\hat{\beta} = Z\hat{b}$ ($\hat{b} = [1, -\hat{\beta}']'$), where

$$\hat{b} = b + O_p(T^{d-1}) \quad \text{and} \quad \hat{u}_t = b'z_t + O_p(T^{d-\frac{1}{2}}) \tag{29}$$

and where it can again be shown, using the fact that

$$T^{-d}\sum_{t=1}^{[Tr]} b'z_t \quad \xrightarrow{d} \quad \text{fractional Brownian Motion}, \tag{30}$$

that $\hat{\alpha} - 1 = O_p(T^{1-2d})$.

Now consider $\hat{Z}_\alpha$ from (7). From (28), we have

$$T(\hat{\alpha} - 1) = O_p(T^{2-2d}), \tag{31}$$

and as the second term in $\hat{Z}_\alpha$ does not diverge any faster, this gives at the same time the divergence rate of the $\hat{Z}_\alpha$–test.

As to $\hat{Z}_t$, we have from (27) and (28) that

$$\sqrt{\hat{u}'_{-1}\hat{u}_{-1}}(\hat{\alpha}-1) = O_p(T^{1-d}), \tag{32}$$

which again is equal to the divergence rate of the complete test statistic.

As to ADF, the t–statistic for $H_0 : \alpha = 0$ in the regression (6) can be written as

$$ADF = (\hat{u}'_{-1}Q_{X_p}\hat{u}_{-1})^{\frac{1}{2}}\frac{\hat{\alpha}}{S_v}, \tag{33}$$

where $Q_{X_p} = I - X_p(X_p'X_p)^{-1}X_p$ and X_p is the matrix of observations on the p regressors $(\Delta\hat{u}_{-1}, \Delta\hat{u}_{-2}, \ldots, \Delta\hat{u}_{-p})$ in (6). We have

$$\hat{u}'_{-1}Q_{X_p}\hat{u}_{-1} = \hat{u}'_{-1}\hat{u}_{-1} - \hat{u}'_{-1}X_p(X_p'X_p)^{-1}X_p\hat{u}_{-1},$$

where

$$\begin{aligned} \hat{u}'_{-1}\hat{u}_{-1} &= O_p(T^{2d}) \quad \text{and} \\ \hat{u}'_{-1}X_p(X_p'X_p)^{-1}X_p\hat{u}_{-1} &= O_p(T^{2d}) \end{aligned}$$

if p does not tend to infinity too fast (see Krämer 1998), implying

$$\hat{u}^{-1}Q_{X_p}\hat{u}_{-1} = O_p(T^{2d}). \tag{34}$$

In the same vein, if p does not tend to infinity too fast, we have

$$\begin{aligned} \hat{\alpha} &= O_p(T^{1-2d}) && (35) \\ S_v &\xrightarrow{p} \text{constant} > 0, \quad \text{so} && (36) \\ ADF &= O_p(T^{1-d}). && (37) \end{aligned}$$

As to $\hat{P}_u$, one first verifies that the steps in the proof of Theorem 5.2 in Phillips/Ouliaris (1990, p. 186) that lead to

$$\hat{w}_{11.2} \xrightarrow{p} \text{constant} > 0 \tag{38}$$

are still valid in the present context. The divergence rate under fractional cointegration of $\hat{P}_u$ then follows from $\hat{u}'\hat{u} = O_p(T^{2d})$ (see (27)).

As to $\hat{P}_z$, we decompose $M_{zz} = \frac{1}{T}\sum_{t=1}^{T} z_t z_t'$ in into

$$M_{zz} = \frac{1}{T}\begin{bmatrix} y'y & y'X \\ X'y & X'X \end{bmatrix}, \tag{39}$$

where

$$M_{zz}^{-1} = T\begin{bmatrix} (\hat{u}'\hat{u})^{-1} & \cdot \\ \cdot & (X'X - X'yy'X/y'y)^{-1} \end{bmatrix} = O_p(T^{1-2d}). \tag{40}$$

As $\hat{\Omega}$ remains $O_p(1)$ under fractional cointegration, the theorem follows from the definition (15).

4 Divergence rates under stationary alternatives

THEOREM 2: *Under the assumptions specified in section 2, where the residuals u_t in $y_t = \beta' x_t + u_t$ are $I(d)$ with $-0 < d < \frac{1}{2}$, we have*

(i) $\hat{Z}_\alpha = O_p(T)$

(ii) $\hat{Z}_t = O_p(T^{\frac{1}{2}})$

(iii) $ADF = O_p(T^{\frac{1}{2}})$

(iv) $\hat{P}_u = O_p(T)$

(v) $\hat{P}_z = O_p(T)$

REMARK: The theorem shows that the divergence rates under stationary long memory alternatives are identical to divergence rates under stationary short memory alternatives, as given by Phillips/Ouliaris (1990, Theorem 5.2). In particular, they no longer depend on d. Also, the relative differences in divergence speeds from the nonstationary case are retained.

PROOF OF THEOREM: By assumption,

$$u_t = y_t - \beta' z_t =: q_t$$

is stationary, and from

$$\begin{aligned} \hat{b} &= b + O_p(T^{d-1}) \quad \text{we have} \\ \hat{u}_t &= q_t + O_p(T^{d-\frac{1}{2}}) \quad \text{and} \end{aligned} \tag{41}$$

$$\hat{\alpha} = \frac{\frac{1}{T}\Sigma q_t q_{t-q}}{\frac{1}{T}\Sigma q_{t-1}^2} + O_p(1) \quad \xrightarrow{p} \quad \frac{E(q_t q_{t-1})}{E(q_{t-1}^2)} =: \alpha < 1. \tag{42}$$

Therefore

$$T(\hat{\alpha} - 1) = O_p(T), \tag{43}$$

which is also the divergence rate of the $\hat{Z}_\alpha$–test.

In the same vein, the divergence rate of the $\hat{Z}_t$–test follows from

$$\sqrt{\hat{u}'_{-1}\hat{u}_{-1}}(\hat{\alpha} - 1) = O_p(T^{\frac{1}{2}}), \tag{44}$$

and the divergence rates of the $\hat{P}_u$ and $\hat{P}_z$–tests are obtained by replicating the proof of theorem 5.2 in Phillips/Ouliaris (1990, p. 186) (this proof establishes the divergence rates under stationary short memory alternatives, but goes through with stationary long memory alternatives as well).

It is more difficult to establish the divergence rate of the Augmented Dickey Fuller test. We have

$$\hat{u}'_{-1}\hat{u}_{-1} = O_p(T) \quad \text{and} \tag{45}$$

$$\hat{u}'_{-1}X_p(X_p'X_p)^{-1}X_p'\hat{u}'_{-1} = O_P(1), \quad \text{so} \tag{46}$$

$$\hat{u}'_{-1}Q_{X_p}\hat{u}_{-1} = O_p(T). \tag{47}$$

Also, if p does not tend to infinity too fast (see Krämer 1998), we have

$$\hat{\alpha} = O_p(1)$$

and

$$S_V^2 \xrightarrow{p} \sigma_\varepsilon^2, \tag{48}$$

where the ε_t's are the innovations in the infinite AR–representation of $q_t = b' z_t$ (see Fuller 1996, p. 374), and the divergence rate of the test statistic follows.

5 An Empirical Illustration

Next we apply the tests discussed so far to three time series of German common stocks (logarithms, daily, from Jan. 4, 1960 to Dec. 30, 1991, comprising $T = 7928$ observations adjusted for dividends, stock splits etc.): Chemical companies Bayer, BASF and Hoechst. In an efficient market, stock prices cannot be cointegrated (since returns would otherwise be predictable, using the Granger representation theorem), but as Figure 1 seems to imply, there is certainly cointegration among the stocks above (we show only log prices of Bayer of Hoechst in order not to overload the picture).

Figure 1:
log prices of Bayer and Hoechst plotted against time

Figure 2, a three–dimensional scatterplot of all three stocks against each other, corroborates this visual impression of cointegration: Prices seem to stick closely to a line in $\mathbb{R}^3$ (implying two cointegrating relationships).

Figure 2:
log prices of Bayer, Hoechst and BASF plotted against each other

However, applying formal tests of the null hypothesis of no cointegration to the residuals $\hat{u}$ of the regression

$$\ell n\ (\text{Bayer})\ = \hat{\beta}_1\ \ell n\ (\text{Hoechst})\ + \hat{\beta}_2\ \ell n\ (\text{BASF})\ + \hat{u} \tag{49}$$

(and similarly to the residuals of alternative regressions where the roles of dependent and independent variables are reversed), one scarcely can reject: Table 1 gives the test statistics and the respective 5%–values of the tests discussed above — more often than not, the null hypothesis of no cointegration cannot be rejected.

Table 1: Residual based cointegration test applied to ℓn (Bayer), ℓn (BASF) and ℓn (Hoechst)

	test statistic	critical values (5 %)	rejection
ADF(p=3)	-2,97	-3,76	no
ADF(p=7)	-2,82	-3,76	no
$\hat{Z}_\alpha$	-7,94	-22,27	no
$\hat{Z}_t$	-1,87	-3,33	no
$\hat{P}_u$	20,73	53,97	no
$\hat{P}_z$	100,46	89,87	yes

The reason for this apparant failure to recognize a cointegrating relationship when visual inspection strongly suggests that one exists appears to be the long memory in the cointegrating residuals: Figure 3 gives the first 100 empirical autocorrelations of the residuals from the regression (37), and Figure 4 the estimated spectral density: Both figures strongly suggest that the cointegrating residuals are best modelled as an $I(d)$–process, and that the conventional cointegration theory with ARMA–residuals does not apply.

Figure 3:

Empirical autocorrelations of estimated cointegration residuals

Figure 4:

Estimated spectral density of estimated cointegration residuals

We also estimated the d–parameter by both the Geweke–Porter–Hudak method and by Range–Scale analysis, with estimated values clustering around 0.5. The lesson from this empirical application therefore seems to be that even blatant cointegration (in the sense that trending variables stick very close to each other) is easily overlooked by standard tests when the cointegrating residuals are fractionally integrated.

References

Baillie, R.T. (1996): "Long memory processes and fractional integration in econometrics." *Journal of Econometrics* 73, 5 – 59.

Booth, G. and Tse, Y. (1995): "Long memory in interest rate future markets, a fractional cointegration analysis." *Journal-of-Futures-Markets* 15(5), 573 – 587.

Cheung, Y.W. (1993): "Tests for fractional integration: A Monte Carlo investigation." *Journal of Time Series Analysis* 14, 331 – 345.

Cheung, Y.W. and Lai, K.S. (1993): "A fractional cointegration analysis of purchasing power parity." *Journal of Business and Economic Statistics* 11, 103 – 112.

Diebold, F.X. and Rudebusch, G.D. (1989): "Long memory and persistence in aggregate output." *Journal of Monetary Economics* 24, 189 – 209.

Dolado, J. and Marmol, F. (1997): "On the properties of the Dickey–Pantula test against fractional alternatives." Mimeo, Universidad Carlos III, Madrid.

Fuller, W. (1996): *Statistical Analysis of Time Series.* Wiley.

Hamilton, J.D. (1994): *Time series analysis.* Princeton University Press.

Hassler, U. and Wolters, J. (1994): "On the power of unit root tests against fractional alternatives." *Economics Letters* 45, 1 – 5.

Hausman, J. (1978): "Specification tests in econometrics." *Econometrica* 46, 1251 – 1271.

Krämer, W. (1998): "Fractional integration and the Augmented Dickey Fuller test." To appear in *Economics Letters.*

Lee, D. and Schmidt, P. (1996): "On the power of the KPSS test of stationarity against fractionally integrated alternatives." *Journal of Econometrics* 73, 285 – 302.

Marmol, F. (1997): "Searching for fractional evidence using unit root tests." Mimeo, Universidad Carlos III, Madrid.

Phillips, P.C.B. and Ouliaris, S. (1990): "Asymptotic properties of residual based tests for cointegration." *Econometrica* 58, 165 – 193.

Sowell, F. (1990): "The fractional unit root distribution." *Econometrica* 58, 495 – 505.

Figure 1:
log prices of Bayer and Hoechst plotted against time

Figure 2:
log prices of Bayer, Hoechst and BASF plotted against each other

Figure 3:
Empirical autocorrelations of estimated cointegration residuals

Figure 4:
Estimated spectral density of estimated cointegration residuals