

Working Paper 99-79  
Statistics and Econometrics Series 31  
October 1999

Departamento de Estadística y Econometría  
Universidad Carlos III de Madrid  
Calle Madrid, 126  
28903 Getafe (Spain)  
Fax (34-91) 624-9849

COINTEGRATION TESTING IN SINGLE ERROR-CORRECTION EQUATIONS IN  
THE PRESENCE OF LINEAR TIME TRENDS.

Uwe Hassler.\*

Abstract

---

Banerjee, Dolado and Mestre (J. Time Ser. Anal. 19 (1998) 267-283) introduce an error-correction test for the null hypothesis of no cointegration. The present paper supplements their work. They provide critical values for regressions with and without detrending. Here it is shown that the latter are not appropriate if the series display linear trends. This does not mean that detrending is required. Correct percentiles are suggested for the case that series follow linear time trends but tests are based on regressions without detrending. They are readily available from the literature.

---

Keywords: Integrated series with drift; effect of not detrending.

\*Free University of Berlin. Institute of Statistics and Econometrics. Boltzmannstr. 20 D-14195 Berlin Germany, e-mail: [uwe@wiwiss.fu-berlin.de](mailto:uwe@wiwiss.fu-berlin.de). I am grateful to Bruce E. Hansen for comments on an earlier version. This paper was revised while visiting Universidad Carlos III de Madrid. Financial support from the European Commission through the *Training and Mobility of Researchers* programme is gratefully acknowledged.

# 1 Introduction

Engle and Granger (1987) established the common practice of estimating and testing cointegrating relations within the framework of static regressions. They proposed to test for the null hypothesis of no cointegration by means of Dickey-Fuller type residual tests. The asymptotic distributions of these tests have been provided by Phillips and Ouliaris (1990). They consider the cases of integrated variables without linear time trends and of integrated series with drifts. However, in the latter case only detrended regressions are considered. In practice, many time series can be treated as  $I(1)$  with drift, i.e. they are integrated of order one and display a linear trend at the same time. Still, often applied workers are not interested in detrended regressions, see e.g. the discussion in Hassler (1999). Static regressions of  $I(1)$  series with linear time trends without detrending have been studied by Hansen (1992). In particular, he finds that residual-based Dickey-Fuller statistics given  $k$   $I(1)$  regressors with drifts but without detrending have limiting distributions identical to statistics arising from the detrended regression on  $k-1$  integrated regressors. Hansen (1992, p.103) concludes that *the deterministic trends in the data affect the limiting distributions of the test statistics whether or not we detrend the data*. A corresponding result will be obtained in this paper for the single equation error-correction test.

The static cointegration regression has been criticized already by Banerjee, Dolado, Hendry and Smith (1986) on experimental grounds. Consequently, Banerjee, Dolado and Mestre (1998) developed a cointegration test embedded in a single equation error-correction framework. The present paper supplements their work with respect to the effect of linear time trends

on tests based on regressions without detrending. Related work has been published recently by Harbo, Johansen, Nielsen and Rahbek (1998). They consider partial (but not necessarily one-dimensional) subsystems of error-correction models and propose likelihood ratio tests for the cointegration rank. In the presence of linear time trends their test has the disadvantage of depending on nuisance parameters in case of not detrending. Therefore, Harbo et al. (1998) suggest to use a detrended model where the asymptotic distribution is free of nuisance parameters. However, detrending may cost efficiency and may not be desirable from an economic point of view. For this reason it is interesting that it turns out that the t-type test without detrending by Banerjee et al. (1998) does not depend on nuisance parameters asymptotically.

In the next section I briefly review the error-correction test by Banerjee et al. (1998) in its very simplest form and state the main result. Section 3 provides experimental evidence that the percentiles traditionally used in the presence of linear time trends without detrending are not correct in that they reject the null hypothesis of no cointegration more often than the stated nominal level. However, this does not mean that detrending is required to achieve correct inference. The critical values I suggest instead without detrending, which are readily available from the literature, turn out to be appropriate. They are justified asymptotically in Section 4 by deriving the limiting distribution. The final section summarizes and describes in simple words how and where appropriate critical values can be found in the situation investigated here.

## 2 Error-correction test

Consider the simplest prototypical single equation error-correction mechanism

$$\Delta y_t = \gamma(y_{t-1} - \lambda'x_{t-1}) + \beta' \Delta x_t + e_t, \quad t = 1, 2, \dots, T,$$

with the usual difference operator  $\Delta$ . Equivalently one may write

$$\Delta y_t = \gamma y_{t-1} + b'x_{t-1} + \beta' \Delta x_t + e_t, \quad (1)$$

where  $b = -\gamma\lambda$ ,  $e_t$  is white noise, and  $x_t$  is a  $k$ -dimensional vector of  $I(1)$  series with drift,

$$\begin{aligned} x_t &= x_0 + \mu_x t + \sum_{i=1}^t u_i, \quad u_t \sim I(0), \\ &= x_0 + \mu_x t + \xi_t, \quad \mu_x \neq 0, \end{aligned} \quad (2)$$

where  $u_t$  is stationary. Please note that  $\mu_x \neq 0$  does not require that all components are nonzero. It is further assumed that  $\xi_t$  (and hence  $x_t$ ) is not cointegrated. Equation (1) above corresponds to equation (1) or (1'') in Banerjee et al. (1998). If  $\gamma < 0$ , then  $y_t$  and  $x_t$  are cointegrated with cointegrating vector  $(1, -\lambda')$ . If  $\gamma = 0$ , and if  $\Delta x_t$  is not generated by an error-correction mechanism, which is the assumption underlying Banerjee et al. (1998), then  $y_t$  and  $x_t$  are not cointegrated. This is the null hypothesis maintained for the rest of the paper:

$$y_t = c + \beta' x_t + z_t, \quad z_t := \sum_{i=1}^t e_i, \quad (3)$$

so that (1) reduces to

$$\Delta y_t = \beta' \Delta x_t + e_t, \quad t = 1, 2, \dots, T.$$

Banerjee et al. (1998) suggest to test for  $\gamma = 0$  by means of the usual t-statistic  $t_\gamma$  from an ordinary least squares (OLS) regression of (1). The null hypothesis is rejected for too negative values. The authors derive the limiting distribution under  $H_0$  and provide critical values, Banerjee et al. (1998, Table 1A); however, they are simulated under the assumption that  $\mu_x = 0$  in (2), which does not allow for linear time trends.

Apart from the OLS estimation of (1),

$$\Delta y_t = \hat{a} + \hat{\gamma} y_{t-1} + \hat{b}' x_{t-1} + \hat{\beta}' \Delta x_t + \hat{\epsilon}_t, \quad (4)$$

we also have the detrended regression with  $t = 1, 2, \dots, T$ ,

$$\Delta y_t = \hat{a}_\tau + \hat{\gamma}_\tau y_{t-1} + \hat{b}'_\tau x_{t-1} + \hat{\delta} t + \hat{\beta}'_\tau \Delta x_t + \hat{\epsilon}_t. \quad (5)$$

From (5) one may test the null of no cointegration by means of the standard t-ratio testing for  $\gamma_\tau = 0$ . The additional deterministic regressor  $t$  changes the null distribution no matter whether  $\mu_x \neq 0$  or not. Appropriate percentiles are found in Banerjee et al. (1998, Table 1B).

In order to motivate the following result we observe that the linear trend dominates  $x_{t-1}$ :

$$\begin{aligned} x_{t-1} &= x_0 + \mu_x(t-1) + \sum_{i=1}^{t-1} u_i \\ &= O(1) + O(T) + O_p(T^{0.5}). \end{aligned}$$

Let us first consider the case  $k = 1$  and assume that the effect of  $\hat{\beta}' \Delta x_t$  can be ignored (which is by no means self-evident, cf. Hansen, 1995); in this case the regression (4) reduces to a univariate detrended Dickey-Fuller

regression testing  $y_t$  for a unit root. If  $k > 1$  the vector  $x_t$  given in (2) is not cointegrated and hence driven by  $k$  common stochastic trends. At the same time it is driven by exactly one common deterministic trend  $t$ . As the linear trend dominates one stochastic trend,  $x_t$  behaves like  $k-1$  common stochastic and one common linear trend. This is the intuition for the following decision rule.

**Criterion 1 (Rule)** *In the presence of linear trends the test based on  $t_\gamma$  from (4) with one  $x$ -variable ( $k = 1$ ) should be applied with the critical values of the detrended univariate Dickey-Fuller test; if  $k > 1$ ,  $t_\gamma$  from (4) should be applied with critical values from Banerjee et al. (1998, Table 1B) for  $k - 1$  instead of with percentiles from Banerjee et al. (1998, Table 1A) for  $k$ .*

Table 1: Critical values (not) taking into account linear trends

$k =$	1	2	3	4	5
1A: $k$	-2.89	-3.19	-3.42	-3.66	-3.82
1B: $k - 1$	-3.13	-3.39	-3.62	-3.82	-4.00

Asymptotic critical values at the 10% level from Banerjee et al. (1998, Table 1A) for  $k$  (standard practice) and according to Criterion 1 (1B:  $k - 1$ ).

Table 1 compares the asymptotic critical values at the 10% level from Banerjee et al. (1998, Table 1A) with those proposed by the decision rule Criterion 1 (the percentiles for the detrended Dickey-Fuller test are from MacKinnon, 1991). Testing in model (1) for cointegration, it is standard practice to estimate (4) and use the critical values from the first row in Table 1. They differ from those according to the Rule by around 0.2. Remember

that the test rejects for too small values: Following the above Rule instead of standard practice cointegration will be established less often at the 10% level in the presence of linear time trends. The next section will show how severe the overrejection implied by standard practice is.

### 3 Monte Carlo evidence

To verify the usefulness of Criterion 1 a Monte Carlo experiment was performed. Let

$$y_t = x_{1t} + \dots + x_{kt} + r_t, \quad t = 1, 2, \dots, T,$$

with the random walks

$$r_t = r_{t-1} + v_{0t},$$

$$x_{it} = 1 + x_{it-1} + v_{it}, \quad i = 1, 2, \dots, k,$$

where  $v_{it}$  are  $N(0, 1)$  white noise series independent of each other. All experiments were done with GAUSS32 and are based on 5000 replications. I generated  $T + 50$  observations and discarded the first 50 observations to get rid of the influence of starting values.

The estimated regression was (4). The ordinary t-statistics  $t_\gamma$  for the true null hypothesis of no cointegration were computed and compared with the critical values from Banerjee et al. (1998, Table 1A) and with the percentiles according to the Rule in Section 2. Throughout, finite sample critical values were applied (those from the detrended Dickey-Fuller test are from MacKinnon, 1991). The percentages of rejection are summarized in Table 2. From

Table 2: Percentage of rejection

10 % level						
	$T = 100$			$T = 500$		
$k =$	1	2	3	1	2	3
1A: $k$	15.04	15.06	14.72	16.06	18.26	15.48
1B: $k - 1$	9.74	9.78	10.16	9.92	9.76	10.02

  

5 % level						
	$T = 100$			$T = 500$		
$k =$	1	2	3	1	2	3
1A: $k$	7.84	7.20	7.88	8.06	8.02	7.90
1B: $k - 1$	5.12	4.80	5.26	4.92	4.93	5.10

  

1 % level						
	$T = 100$			$T = 500$		
$k =$	1	2	3	1	2	3
1A: $k$	1.78	1.50	1.68	1.64	1.44	1.42
1B: $k - 1$	1.16	0.94	1.14	1.10	0.74	1.02

Given are the percentages of rejection based on  $t_\gamma$  from (4). This t-statistic is compared with the finite sample critical values from Banerjee et al. (1998, Table 1A) for  $k$  and with those according to Criterion 1 (1B:  $k - 1$ ). Further information is given in the text.



Table 2 it is clear that the standard practice of not detrending and using corresponding critical values overrejects if the series are  $I(1)$  with drift. Even if we do not detrend, the fact that the regressors display linear trends affects the critical values. The rejection rates according to Criterion 1, in contrast, are very close to the nominal levels. Fortunately, the overrejection of the standard practice is not very drastic. Nevertheless, the Rule introduced here clearly improves cointegration testing in single error-correction equations in the presence of linear trends.

## 4 Asymptotic justification

In order to justify Criterion 1 in addition to the experimental evidence I shall derive the asymptotic distribution of  $t_\gamma$  from (4) under the null hypothesis and compare it with that of  $t_{\gamma_r}$  from (5). The asymptotic analysis relies on the following assumptions.

Let  $e_t$  be a white noise process independent of the  $I(0)$  process  $u_t$ , and suppose that they satisfy the invariance principle

$$T^{-0.5} \sum_{t=1}^{[rT]} \begin{pmatrix} e_t \\ u_t \end{pmatrix} = T^{-0.5} \begin{pmatrix} z_{[rT]} \\ \xi_{[rT]} \end{pmatrix} \Rightarrow \begin{pmatrix} B(r) \\ P(r) \end{pmatrix} = \begin{pmatrix} \sigma_e W(r) \\ \Omega_u^{0.5} V(r) \end{pmatrix}, \quad (6)$$

where  $\sigma_e^2 = \text{Var}(e_t)$ ,  $\Omega_u$  is the long-run covariance matrix of  $u_t$ , and  $W(r)$  and  $V(r)$  are standard Wiener processes independent of each other. Next it is assumed without loss of generality that the last component (say  $\mu_k$ ) of  $\mu_x$  is different from zero. The variable  $x_t$  and its corresponding Brownian

motion are partitioned accordingly:

$$x_t = \begin{pmatrix} x_{(k-1)t} \\ x_{k,t} \end{pmatrix} = \begin{pmatrix} x_{(k-1)0} \\ x_{k,0} \end{pmatrix} + \begin{pmatrix} \mu_{(k-1)} \\ \mu_k \end{pmatrix} t + \begin{pmatrix} \xi_{(k-1)t} \\ \xi_{k,t} \end{pmatrix}, \quad \mu_k \neq 0,$$

$$P(r) = \begin{pmatrix} P_{(k-1)}(r) \\ P_k(r) \end{pmatrix}.$$

Let the subscript  $(k-1)$  always denote subvectors of length  $k-1$ ; if  $k=1$  the corresponding expressions are understood to be absent so that the Proposition below is derived for  $k=1$  and  $k>1$  at the same time. Correspondingly, the null hypothesis can be rewritten as

$$y_t = c + \beta'_{(k-1)} x_{(k-1)t} + \beta_k x_{k,t} + z_t.$$

It is important to note that  $y_t$  and  $x_{(k-1)t}$  can be expressed in terms of  $x_{k,t}$  (where constant terms are neglected for simplicity):

$$x_{(k-1)t} = \mu_{(k-1)} \frac{x_{k,t}}{\mu_k} + \zeta_t, \quad \zeta_t := \xi_{(k-1)t} - \mu_{(k-1)} \frac{\xi_{k,t}}{\mu_k}, \quad (7)$$

$$y_t = \frac{\beta'_k \mu_x}{\mu_k} x_{k,t} + z_t + \beta'_{(k-1)} \zeta_t, \quad (8)$$

where (6) implies

$$T^{-0.5} \zeta_{[rT]} \Rightarrow Q(r) := P_{(k-1)}(r) - \mu_{(k-1)} \frac{P_k(r)}{\mu_k} = \Omega_{\Delta\zeta}^{0.5} U(r). \quad (9)$$

In (9),  $U(r)$  is a  $(k-1)$ -dimensional standard Wiener process independent of  $W(r)$ . In order to investigate  $t_\gamma$  from (4) I make use of what has been called Frisch-Waugh-Lovell (FWL) theorem by Davidson and MacKinnon (1993). Consider the regressions of the series from (4) on  $x_{k,t-1}$  at first stage,

$$\Delta y_t = \tilde{\psi}_0 + \tilde{\pi}_0 x_{k,t-1} + \tilde{v}_t, \quad (10)$$

$$\Delta x_t = \tilde{\psi}_1 + \tilde{\pi}_1 x_{k,t-1} + \tilde{u}_t, \quad (11)$$

$$y_{t-1} = \tilde{\phi}_0 + \tilde{\theta}_0 x_{k,t-1} + \tilde{z}_{t-1}, \quad (12)$$

$$x_{(k-1)t-1} = \tilde{\phi}_1 + \tilde{\theta}_1 x_{k,t-1} + \tilde{\zeta}_{t-1}, \quad (13)$$

where (12) and (13) are motivated by (8) and (7), respectively. Here, least squares is denoted by means of tildes. The residuals from (10) - (13) are regressed on each other at second stage,

$$\tilde{v}_t = \hat{\gamma} \tilde{z}_{t-1} + \hat{b}'_{(k-1)} \tilde{\zeta}_{t-1} + \hat{\beta}' \tilde{u}_t + \hat{e}_t. \quad (14)$$

According to the FWL theorem the estimators and residuals from (14) are numerically identical to the corresponding values from (4). Further, it can be shown that the difference between the t-ratios from (4) and (14), which is due to different degrees of freedom, vanishes as  $T \rightarrow \infty$ . Hence, the asymptotic distribution of  $t_\gamma$  from (4) can be derived from (14). In the Appendix the following result is established in terms of detrended processes.

**Proposition 2** *Under the null of no cointegration (3) and the assumptions of this section the limiting distribution of  $t_\gamma$  from (4) is*

$$\frac{\int \tilde{W} dW - \int \tilde{W} \tilde{U}' \left( \int \tilde{U} \tilde{U}' \right)^{-1} \int \tilde{U} dW}{\sqrt{\int \tilde{W}^2 - \int \tilde{W} \tilde{U}' \left( \int \tilde{U} \tilde{U}' \right)^{-1} \int \tilde{U} \tilde{W}}}$$

where  $\tilde{W}(r)$  and  $\tilde{U}(r)$  are detrended versions of the Wiener processes from (6) and (9).

Given a Brownian motion  $X(r)$  the detrended process is defined as follows, cf. Park and Phillips (1988),

$$\tilde{X}(r) := X(r) - \int X - 12 \left( \int sX - \frac{1}{2} \int X \right) \left( r - \frac{1}{2} \right).$$

Given a second Brownian motion  $Y(r)$  it follows

$$\int \tilde{X}\tilde{Y}' = \int XY' - \int X \int Y' - 12 \left( \int rX - \frac{1}{2} \int X \right) \left( \int rY - \frac{1}{2} \int Y \right)',$$

$$\int \tilde{X}dY' = \int XdY' - \int X Y(1)' - 12 \left( \int rX - \frac{1}{2} \int X \right) \left( \int rdY - \frac{1}{2} Y(1) \right)'.$$

Integrals such as  $\int_0^1 X(r)dr$ ,  $\int_0^1 rdY(r)$  or  $\int_0^1 X(r)dY(r)'$  and so on are written as  $\int X$ ,  $\int rdY$  and  $\int XdY'$  in order to save space.

The Proposition has the following interpretation. If  $k = 1$ , there is no vector  $U$ , and the asymptotic distribution simplifies to

$$\frac{\int \tilde{W}dW}{\sqrt{\int \tilde{W}^2}} = \frac{\int WdW - W(1)\int W - 12(\int rW - \frac{1}{2}\int W)(\int rdW - \frac{1}{2}W(1))}{\sqrt{\int W^2 - (\int W)^2 - 12(\int rW - \frac{1}{2}\int W)^2}}.$$

With  $W(1) - \int W = \int rdW$  this limiting distribution just equals that one derived by Phillips and Perron (1988, Theorem 1e) for the detrended univariate Dickey-Fuller test. This justifies the Rule in Section 2 asymptotically for  $k = 1$ . If  $k > 1$ , let us consider the limiting behaviour of  $t_{\gamma_r}$  from (5). Its limiting distribution is not explicitly given in Banerjee et al. (1998). But following their arguments or the derivation in the Appendix it turns out: Under the null of no cointegration (3) and the assumptions of this section the limiting distribution of  $t_{\gamma_r}$  from (5) is:

$$\frac{\int \tilde{W}dW - \int \tilde{W}\tilde{V}' \left( \int \tilde{V}\tilde{V}' \right)^{-1} \int \tilde{V}dW}{\sqrt{\int \tilde{W}^2 - \int \tilde{W}\tilde{V}' \left( \int \tilde{V}\tilde{V}' \right)^{-1} \int \tilde{V}\tilde{W}}}.$$

The vector  $V(r)$  defined in (6) is  $k$ -dimensional, while  $U(r)$  defined in (9) arising in the Proposition is of dimension  $k - 1$ . Apart from this the limits evolving from (4) and (5) have identical structures. This justifies Criterion 1 asymptotically for  $k > 1$ .

## 5 Summary

Banerjee et al. (1998) propose an error-correction test for the null hypothesis that  $y_t$  and the  $k$ -dimensional  $I(1)$  vector  $x_t$  are not cointegrated. They provide critical values for two versions of the test, one being based on a detrended regression, the other on a regression without detrending. The percentiles for the latter are simulated under the assumption that the series of interest are integrated without drifts, i.e. they are not designed for variables with linear time trends. Nevertheless those critical values are applied in practice with time series with linear trends entering an error-correction regression without detrending.

The purpose of the present paper is twofold. Firstly, it is shown that the standard practice of using critical values without detrending in the presence of linear trends implies rejection rates higher than the alleged nominal level. Although the overrejection is not very drastic I, secondly, suggest appropriate critical values for the case of linear trends without detrending. Fortunately, they are even available from the literature, and in particular from Banerjee et al. (1998, Table 1B) for  $k > 1$ : With  $k$   $I(1)$  variables  $x_t$  with drift but without detrending the adequate percentiles are from the table for  $k - 1$   $x$ -variables with detrending; if  $k = 1$  the percentiles from the univariate de-

trended Dickey-Fuller test are correct.

## Appendix: Proof of Proposition

In the derivation of the Proposition all sums run from 1 to  $T$  and all integrals are from 0 to 1 if not indicated otherwise. Integrals such as  $\int B(r)dr$  or  $\int r dB(r)$  and so on are written as  $\int B$  or  $\int r dB$  in order to save space. One can obtain a Lemma that implies the Proposition. In order to do so I make use of convergence results implied by the functional central limit theorem (6) and (9) and collected e.g. in Banerjee et al. (1993, Table 3.3). Together with

$$\sum_{t=1}^T \left( t - \frac{T+1}{2} \right)^2 = \frac{T^3 - T}{12}$$

it follows e.g.

$$T^{-3} \sum (x_{k,t-1} - \bar{x}_{k,-1})^2 \rightarrow \frac{\mu_k^2}{12}. \quad (15)$$

This implies for the estimators from (10) - (13):

$$T^{0.5}(\tilde{\theta}_1 - \mu_{(k-1)}/\mu_k) \Rightarrow \frac{12}{\mu_k} \left( \int rQ - \frac{1}{2} \int Q \right) =: \check{\theta}_1, \quad (16)$$

$$T^{0.5}(\tilde{\theta}_0 - \beta' \mu_x / \mu_k) \Rightarrow \beta'_{(k-1)} \check{\theta}_1 + \frac{12}{\mu_k} \left( \int rB - \frac{1}{2} \int B \right) =: \check{\theta}_0, \quad (17)$$

$$T^{1.5} \tilde{\pi}_1 \Rightarrow \frac{12}{\mu_k} \left( \int r dP - \frac{1}{2} P(1) \right) =: \check{\pi}_1, \quad (18)$$

$$T^{1.5} \tilde{\pi}_0 \Rightarrow \beta' \check{\pi}_1 + \frac{12}{\mu_k} \left( \int r dB - \frac{1}{2} B(1) \right) =: \check{\pi}_0. \quad (19)$$

The limiting results (16) - (19) entail for the residuals entering (14):

$$\begin{aligned}
\tilde{\zeta}_{t-1} &= \zeta_{t-1} - \bar{\zeta}_{-1} + O_p(T^{0.5}), \\
\tilde{z}_{t-1} &= z_{t-1} - \bar{z}_{-1} + \beta'_{(k-1)}(\zeta_{t-1} - \bar{\zeta}_{-1}) + O_p(T^{0.5}), \\
\tilde{u}_t &= u_t - \bar{u} + O_p(T^{-0.5}), \\
\tilde{v}_t &= e_t - \bar{e} + \beta'(u_t - \bar{u}) + O_p(T^{-0.5}).
\end{aligned}$$

It follows e.g. with (15) and (16)

$$\begin{aligned}
T^{-2} \sum \tilde{\zeta}_{t-1} \tilde{\zeta}'_{t-1} &\Rightarrow \frac{\mu_k^2 \check{\theta}_1 \check{\theta}'_1 - \mu_k \check{\theta}_1}{12} \left( \int rQ - \frac{1}{2} \int Q \right)' \\
&\quad - \mu_k \left( \int rQ - \frac{1}{2} \int Q \right) \check{\theta}'_1 + \int QQ' - \int Q \int Q' \\
&= \int QQ' - \int Q \int Q' - \mu_k \check{\theta}_1 \left( \int rQ - \frac{1}{2} \int Q \right)' \\
&= \int \tilde{Q} \tilde{Q}'.
\end{aligned}$$

The derivation of the remaining results given in the Lemma is analogously straightforward although sometimes tedious algebra is involved; I omit details.

**Lemma 3** *Under the assumptions of the Proposition it holds for the residuals from (10) - (13) as  $T \rightarrow \infty$ :*

$$T^{-2} \sum \tilde{\zeta}_{t-1} \tilde{\zeta}'_{t-1} \Rightarrow \int \tilde{Q} \tilde{Q}', \quad T^{-2} \sum \tilde{\zeta}_{t-1} \tilde{z}_{t-1} \Rightarrow \int \tilde{Q} \tilde{B} + \int \tilde{Q} \tilde{Q}' \beta_{(k-1)},$$

$$T^{-2} \sum \tilde{z}_{t-1}^2 \Rightarrow \int \tilde{B}^2 + 2\beta'_{(k-1)} \int \tilde{Q} \tilde{B} + \beta'_{(k-1)} \int \tilde{Q} \tilde{Q}' \beta_{(k-1)},$$

$$T^{-1} \sum \tilde{\zeta}_{t-1} \tilde{u}'_t \Rightarrow \int \tilde{Q} dP', \quad T^{-1} \sum \tilde{\zeta}_{t-1} \tilde{v}_t \Rightarrow \int \tilde{Q} dP' \beta + \int \tilde{Q} dB,$$

$$T^{-1} \sum \tilde{z}_{t-1} \tilde{u}'_t \Rightarrow \int \tilde{B} dP' + \beta'_{(k-1)} \int \tilde{Q} dP',$$

$$T^{-1} \sum \tilde{z}_{t-1} \tilde{v}_t \Rightarrow \int \tilde{B} dB + \int \tilde{B} dP' \beta + \beta'_{(k-1)} \int \tilde{Q} dB + \beta'_{(k-1)} \int \tilde{Q} dP' \beta,$$

$$T^{-1} \sum \tilde{u}_t \tilde{u}'_t \rightarrow \text{Cov}(u_t) =: \Gamma_u,$$

$$T^{-1} \sum \tilde{u}_t \tilde{v}_t \rightarrow \Gamma_u \beta, \quad T^{-1} \sum \tilde{v}_t^2 \rightarrow \beta' \Gamma_u \beta + \sigma_e^2.$$

Now we are ready to analyze the OLS estimator from (14),

$$\begin{pmatrix} \hat{\gamma} \\ \hat{b}_{(k-1)} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} F & G \\ G' & H \end{pmatrix} \sum \begin{pmatrix} \tilde{z}_{t-1} \\ \tilde{\zeta}_{t-1} \\ \tilde{u}_t \end{pmatrix} \tilde{v}_t$$

where

$$\tilde{\Xi}'_{t-1} := (\tilde{z}_{t-1}, \tilde{\zeta}'_{t-1}),$$

$$\begin{aligned} F &= \left( \sum \tilde{\Xi}_{t-1} \tilde{\Xi}'_{t-1} - \sum \tilde{\Xi}_{t-1} \tilde{u}'_t (\sum \tilde{u}_t \tilde{u}'_t)^{-1} \sum \tilde{u}_t \tilde{\Xi}'_{t-1} \right)^{-1} \\ G &= -F \sum \tilde{\Xi}_{t-1} \tilde{u}'_{t-1} (\sum \tilde{u}_t \tilde{u}'_t)^{-1} \\ H &= (\sum \tilde{u}_t \tilde{u}'_t)^{-1} \left( I_k - \sum \tilde{u}_t \tilde{\Xi}'_t G \right). \end{aligned}$$

With the Lemma at hand it follows again from the inverse of partitioned matrices that

$$T^2 F = \left( T^{-2} \sum \tilde{\Xi}_{t-1} \tilde{\Xi}'_{t-1} - O_p(T^{-1}) \right)^{-1} \Rightarrow \begin{pmatrix} A & D \\ D' & C \end{pmatrix}$$



where

$$\begin{aligned}
A &= \left( \int \tilde{B}^2 - \int \tilde{B}\tilde{U}' \left( \int \tilde{U}\tilde{U}' \right)^{-1} \int \tilde{U}\tilde{B} \right)^{-1} \\
D &= -A \left( \int \tilde{B}\tilde{Q}' + \beta'_{(k-1)} \int \tilde{Q}\tilde{Q}' \right) \left( \int \tilde{Q}\tilde{Q}' \right)^{-1} \\
C &= \left( \int \tilde{Q}\tilde{Q}' \right)^{-1} \left( I_{k-1} - \int \tilde{Q}\tilde{B}D + \int \tilde{Q}\tilde{Q}'\beta_{(k-1)}D \right).
\end{aligned}$$

Further the Lemma implies

$$\begin{aligned}
T^2G &\Rightarrow - \begin{pmatrix} A & D \\ D' & C \end{pmatrix} \begin{pmatrix} \int \tilde{B}dP' + \beta'_{(k-1)} \int \tilde{Q}dP' \\ \int \tilde{Q}dP' \end{pmatrix} \Gamma_u^{-1}, \\
TH &\rightarrow \Gamma_u^{-1}.
\end{aligned}$$

Collecting those results yields

$$\hat{\beta} \rightarrow \beta, \hat{b}_{(k-1)} = O_p(T^{-1}),$$

$$T\hat{\gamma} \Rightarrow \frac{\int \tilde{W}dW - \int \tilde{W}\tilde{U}' \left( \int \tilde{U}\tilde{U}' \right)^{-1} \int \tilde{U}dW}{\int \tilde{W}^2 - \int \tilde{W}\tilde{U}' \left( \int \tilde{U}\tilde{U}' \right)^{-1} \int \tilde{U}\tilde{W}},$$

and hence

$$s^2 = T^{-1} \sum (\tilde{v}_t - \hat{\gamma}\tilde{z}_{t-1} - \hat{b}'_{(k-1)}\tilde{\zeta}_{t-1} - \hat{\beta}'\tilde{u}_t)^2 \rightarrow \sigma_e^2.$$

The limiting distribution of

$$t_\gamma = \frac{T\hat{\gamma}}{s\sqrt{T^2[F]_{11}}}$$

from (14) turns out to be as stated in the Proposition. The FWL theorem therefore completes the proof.

## References

Banerjee, A., J.J. Dolado and R. Mestre (1998): Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework; *Journal of Time Series Analysis* 19, 267-283.

Banerjee, A., J.J. Dolado, D.F. Hendry 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.

Banerjee, A., J.J. Dolado, J.W. Galbraith, and D.F. Hendry (1993): *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data*; Oxford University Press.

Engle, R.F., and C.W.J. Granger (1987): Co-Integration and Error Correction: Representation, Estimation, and Testing; *Econometrica* 55, 251-276.

Hansen, B.E. (1992): Efficient Estimation and Testing of Cointegrating Vectors in the Presence of Deterministic Trends; *Journal of Econometrics* 53, 87-121.

Hansen, B.E. (1995): Rethinking the Univariate Approach to Unit Root Testing: Using Covariates to Increase Power; *Econometric Theory* 11, 1148-1171.

Harbo, J, S. Johansen, B. Nielsen, and A. Rahbek (1998): Asymptotic Inference on Cointegrating Rank in Partial Systems; *Journal of Business and Economic Statistics* 16, 388-399.

Hassler, U. (1999): (When) Should Cointegrating Regressions be Detrended? The Case of a German Money Demand Function; *Empirical Economics* 24, 155-172.

MacKinnon, J.G. (1991): Critical Values for Co-Integration Tests; in: R.F. Engle und C.W.J. Granger (Hrsg.): *Long-Run Economic Relationships*; Oxford University Press, 267-276.

Park, J.Y., and P.C.B. Phillips (1988): Statistical Inference in Regressions with Integrated Processes: Part I; *Econometric Theory* 4, 468-497.

Phillips, P.C.B., and S. Ouliaris (1990): Asymptotic Properties of Residual Based Tests for Cointegration; *Econometrica* 58, 165-193.

Phillips, P.C.B., and P. Perron (1988): Testing for a Unit Root in Time Series Regression; *Biometrika* 75, 335-346.