SFB 823

Testing for nonlinear cointegration under heteroskedasticity

Christoph Hanck, Till Massing

Discussion Paper

Nr. 26/2020

Testing for Nonlinear Cointegration under Heteroskedasticity

Christoph Hanck*

Department of Economics, University of Duisburg-Essen.

Till Massing†

Department of Economics, University of Duisburg-Essen.

September 25, 2020

Abstract

This article discusses cointegration tests for nonlinear cointegration in the presence of variance breaks in the errors. We build on approaches of Cavaliere and Taylor (2006, Journal of Time Series Analysis) for heteroskedastic cointegration tests and of Choi and Saikkonen (2010, Econometric Theory) for nonlinear cointegration tests. We propose a bootstrap test and prove its consistency.

A Monte Carlo study shows the approach to have appealing finite sample properties and to work better than an approach using subresiduals. We provide an empirical application to the environmental Kuznets curves (EKC), finding that the cointegration tests do not reject the EKC hypothesis in most cases.

Keywords: Nonlinear cointegration tests; variance breaks; fixed regressor bootstrap.

JEL Classification Numbers: C12, C32, Q2.

*Address: Universitätsstr. 12, 45130 Essen, Germany, e-mail: *christoph.hanck@vwl.uni-due.de.*

†Address: Universitätsstr. 12, 45130 Essen, Germany, e-mail: *till.massing@uni-due.de.*

1 Introduction

In the past decades, a broad literature on cointegration tests has developed, addressing a variety of different possible features of the data like endogeneity, heteroskedasticity, and nonlinearity. For example, the discussion of the environmental Kuznets curve in our application reveals that the data exhibits both a nonlinear cointegrating relation as well as variance breaks.

This paper presents a framework capable to test for cointegration both when the cointegrating relation is nonlinear and in the presence of heteroskedasticity. In order to achieve this, we mainly build on Choi and Saikkonen (2010) and on Cavaliere and Taylor (2006). The nonlinear cointegrating relation can be very general and variance breaks can occur both in the integrated regressor and in the (stationary or integrated) error term.

There are two possibilities for specifying a null hypothesis. Namely, one can formulate the null hypothesis of *no cointegration*. In this field, e.g., Dickey and Fuller (1979), Phillips and Perron (1988) and their numerous extensions test the null of the presence of a unit root for univariate time series. Engle and Granger (1987) extended this to the context of testing for no cointegration. Alternatively, Kwiatkowski et al. (1992) test the null of stationarity against the alternative of a unit root (commonly known as KPSS test). Shin (1994) extended this approach to test the null of *cointegration*, as we do here. The basic idea is to use the ordinary least squares (OLS) residuals of a linear cointegrating regression to build the test statistic.

This theory has been enhanced in several directions. For example, Leybourne and McCabe (1994) and McCabe et al. (1997) proposed extensions of the original framework. Cavaliere (2005) and Cavaliere and Taylor (2006) incorporated variance breaks into the linear cointegration model. Saikkonen and Choi (2004) dropped the linearity assumption of the cointegrating regression and proposed a test for cointegrating smooth transition functions. Choi and Saikkonen (2010) further extended this to general kinds of nonlinear cointegrating regressions. Both employed nonlinear least squares estimation (NLS) and leads-and-lags regression instead of OLS for estimating the cointegrating parameter vector.

The paper is organized as follows. Section 2 describes the nonlinear cointegrating regression model and the maintained assumptions. Section 3 presents the cointegration test and develops its large sample properties. Furthermore, Section 3 discusses a bootstrap approach for practical implementation of the test. Section 4 analyzes the finite sample quality of the test in a Monte Carlo study. Section 5 illustrates the approach with an application to the environmental Kuznets curve. Unless stated otherwise, all proofs are relegated to Appendix A.

Some notational remarks: We denote by $\lfloor x \rfloor$ the largest integer number smaller or equal than $x \in \mathbb{R}$ and $\lceil x \rceil$ the smallest integer number larger or equal than x. $\mathbb{1}(\cdot)$ denotes the indicator function and $\mathcal{D}_{\mathbb{R}^{m\times m}}[0,1]$ denotes the space of $m \times m$ matrices of càdlàg functions on $[0,1]$, endowed with the Skorohod topology. Weak convergence is denoted by $\xrightarrow{w}$, convergence in probability by $\xrightarrow{p}$, weak convergence in probability (see Giné and Zinn, 1990) by $\xrightarrow{w}_p$, and almost sure convergence by $\xrightarrow{\text{a.s.}}$. All limits are taken as $T \to \infty$, unless stated otherwise.

2 The model and assumptions

In this section, we introduce the model and the underlying assumptions. We consider (as in Choi and Saikkonen, 2010) the nonlinear cointegrating regression

$$y_t = g(x_t, \theta) + u_t, \quad t = 1, \ldots, T, \tag{1}$$

where y_t is 1-dimensional and x_t is the k-dimensional regressor vector. Both y_t and x_t are $I(1)$. We assume that $g(x_t, \theta)$ is a known smooth function of x_t up to the unknown k-dimensional parameter vector θ. We furthermore assume that the vector elements of x_t are not cointegrated (see Assumption 3 for a precise statement below). This also means $g(x_t, \theta)$ is not $I(0)$. The error term is taken to be

$$u_t = \zeta_{u,t} + \mu_t,$$

where

$$\mu_t = \mu_{t-1} + \rho_\mu \zeta_{\mu,t}, \quad \mu_0 = 0.$$

The random walk behavior of x_t is specified by

$$x_t = x_{t-1} + \zeta_{x,t}.$$

The following Assumption 1 discusses the $(k+2)$-dimensional vector process $\zeta_t := (\zeta_{u,t}, \zeta'_{x,t}, \zeta_{\mu,t})'$.

Assumption 1 *(i) $\{\zeta_{u,t}\}$ and $\{\zeta_{\mu,t}\}$ are independent.*

(ii) $\zeta_t := (\zeta_{u,t}, \zeta'_{x,t}, \zeta_{\mu,t})' = \Sigma_t^{1/2}\zeta_t^*$, where $\{\zeta_t^*\}$ is a stationary, zero-mean, unit variance, strong-mixing sequence with mixing coefficient of size $-4r/(r-4)$, for some $r > 4$ and $E||\zeta_t^*||^r < \infty$ and

$$\Sigma_t := \begin{pmatrix} \sigma_{u,t}^2 & \sigma'_{ux,t} & 0 \\ \sigma_{ux,t} & \Sigma_{x,t} & 0 \\ 0 & 0' & \sigma_{\mu,t}^2 \end{pmatrix}.$$

The scalars $\sigma_{u,t}^2$ and $\sigma_{\mu,t}^2$ are strictly positive, $\sigma_{ux,t}$ is k-dimensional, $\Sigma_{x,t}$ $(k \times k)$ is positive definite. All entries may depend on t. We assume that Σ_t is positive definite for any t.

This means that u_t has a random walk component unless $\rho_\mu = 0$. Hence the null hypothesis of cointegration is given by $H_0 : \rho_\mu^2 = 0$ against the alternative $H_1 : \rho_\mu^2 > 0$ of no cointegration.

Assumption 1 is similar to Assumption 1 in Cavaliere and Taylor (2006) but additionally permits correlation between $\zeta_{u,t}$ and $\zeta_{x,t}$ to allow for endogeneity. The Monte Carlo experiments in Section 4 will reveal the proposed bootstrap approach effectively handles endogeneity. We conjecture that a correlation between, e.g., $\zeta_{u,t}$ and $\zeta_{\mu,t}$ will not reveal different insights. We, therefore, abstain from considering further non-zero covariance terms in (ii). Moreover, we also generalize Cavaliere and Taylor (2006, Assumption 1) in terms of permitting autocorrelation of the ζ_t's. This is adopted from Assumption 2 of Choi and Saikkonen (2010).

Following Cavaliere (2005) and Cavaliere and Taylor (2006), we allow for general forms of heteroskedastic errors.

Assumption 2 *The sequence $\{\Sigma_t\}_{t=1}^T$ satisfies $\Sigma_T(s) := \Sigma_{\lfloor Ts \rfloor} = \Sigma(s)$, where $\Sigma(\cdot)$ is a non-stochastic function which lies in $\mathcal{D}_{\mathbb{R}^{(k+2)\times(k+2)}}[0,1]$, with i,j-th element $\Sigma_{ij}(\cdot)$.*

Assumption 2 allows for many possible models for the covariance matrix of ζ_t. For simple or multiple variance shifts $\Sigma_{ij}(\cdot)$ is a piecewise constant function. For example, $\Sigma_{ij}(s) := \Sigma_{ij}^0 + (\Sigma_{ij}^1 - \Sigma_{ij}^0)\mathbb{1}(s \geq \lfloor \tau_{ij} \rfloor)$ represents a shift from Σ_{ij}^0 to Σ_{ij}^1 at time $\lfloor \tau_{ij} T \rfloor$ $(0 \leq \tau_{ij} \leq 1)$. Other possibilities are, e.g., affine functions ($\Sigma_{t,ij}$ exhibits a linear trend), piecewise affine functions, or smooth transition functions. The assumption also allows for very general combinations of variance-covariance shifts. For example, the variance of $\zeta_{u,t}$ can have a shift while $\zeta_{x,t}$ is homoskedastic or heteroskedastic with a different shift function $\Sigma_{ij}(s)$. Notice that variance shifts in $\zeta_{\mu,t}$ are only relevant if the alternative H_1 is true. Although we rule out stochastic volatility here, a generalization to stochastic a stochastic $\{\Sigma_t\}$, s.t. $\{\Sigma_t\}$ is strictly exogenous w.r.t. $\{\zeta_t^*\}$, is possible. We refer to Cavaliere and Taylor (2006) for details.

Furthermore, we define $\Omega_t := t^{-1} Var\left(\sum_{i=1}^{t} \zeta_i\right)$, which can be decomposed as

$$\Omega_t = \begin{pmatrix} \omega_{u,t}^2 & \omega_{ux,t}' & 0 \\ \omega_{ux,t} & \Omega_{x,t} & 0 \\ 0 & 0' & \omega_{\mu,t}^2 \end{pmatrix}.$$

Analogously, $\Omega(s) := \Omega_{\lfloor Ts \rfloor}$. Then, the average long-run covariance matrix $\lim_{T\to\infty} \Omega_T$ is given by

$$\bar{\Omega} = \int_0^1 \Omega(s) \mathrm{d}s,$$

which can be partitioned into

$$\bar{\Omega} = \begin{pmatrix} \bar{\omega}_u^2 & \bar{\omega}_{ux}' & 0 \\ \bar{\omega}_{ux} & \bar{\Omega}_x & 0 \\ 0 & 0' & \bar{\omega}_\mu^2 \end{pmatrix}.$$

Assumption 1 & 2 imply a generalized invariance principle as stated in Lemma 1. The standard invariance principle as in Shin (1994) would require a time-constant covariance matrix Σ.

Lemma 1 *Let Assumptions 1 and 2 hold on $\{\zeta_t\}$. Then, as $T \to \infty$,*

$$T^{-1/2} \sum_{t=1}^{\lfloor Ts \rfloor} \zeta_t \xrightarrow{w} B_\Omega(s), \quad s \in [0,1],$$

where

$$B_\Omega(s) := (B_{0,\Omega}(s), B_{1,\Omega}'(s), B_{2,\Omega}(s))' := \int_0^s \Omega^{1/2}(r) \mathrm{d}B(r),$$

with $B = (B_0, B_1', B_2)'$ is a $(k+2)$-dimensional Brownian motion with unit covariance matrix.

Proof. The proof is analogous to the proof of Lemma 1 in Cavaliere and Taylor (2006) and thus is omitted. ■

The next assumption ensures that the components of x_t are not cointegrated. This is given by the special case $\lambda = 0$.

Assumption 3 *The spectral density matrix $f_{\zeta\zeta}(\lambda)$ is bounded away from zero:*

$$f_{\zeta\zeta}(\lambda) \geq \varepsilon I_{k+2}, \quad \varepsilon > 0.$$

Assumption 4 is the usual assumption required for deriving consistency and asymptotic distribution of the NLS estimator.

Assumption 4 *(i) The parameter space Θ of θ is a compact subset of $\mathbb{R}^k$ and the true parameter $\theta_0 \in \Theta^0$, where Θ^0 denotes the interior of Θ.*

(ii) $g(x, \theta)$ is three times continuously differentiable on $\mathbb{R} \times \Theta^*$, where $\Theta^* \supset \Theta$ is open.

The assumptions on x_t theoretically rule out the possibility of deterministic regressors like an intercept or a time trend because they are not $I(1)$. However, we discuss these interesting scenarios in Appendix B and illustrate that the bootstrap generally works well.

3 Test for nonlinear cointegration

3.1 Nonlinear least squares regression

Following Saikkonen and Choi (2004) and Choi and Saikkonen (2010) we use triangular array asymptotics in order to study the large sample behavior of the proposed test statistic (2), presented below. We fix the actual sample size at T_0 and embed the model in a sequence of models dependent on the sample size T, which tends to infinity. We replace the regressor x_t by $x_{tT} := (T_0/T)^{1/2} x_t$. This makes the regressor and regressand dependent on T and we obtain the actual model for $T_0 = T$. If T_0 is large, the triangular asymptotics can be expected to give reasonable approximations to the finite sample distributions of the estimator and test statistics, see Saikkonen and Choi (2004). Choi and Saikkonen (2010) note that conventional asymptotic results on the NLS estimator are not available when the error term u_t is allowed to be serially correlated or x_t is not exogenous. See

Saikkonen and Choi (2004) and Choi and Saikkonen (2010) for a more detailed discussion about triangular asymptotics.

In particular, we embed the model (1) in a sequence of models

$$y_{tT} = g(x_{tT}, \theta) + u_t, \quad t = 1, \ldots, T.$$

In practice, we always choose $T_0 = T$, so that the transformation x_{tT} is not needed. The transformation is made only to apply triangular asymptotics. We define $B^0_{1,\Omega} := T_0^{1/2} B_{1,\Omega}$.

We use NLS regression to estimate θ_0. Let

$$Q(\theta) = \sum_{t=1}^{T} (y_{tT} - g(x_{tT}, \theta))^2$$

be the objective function to be minimized with respect to $\theta \in \Theta$. Since Q is continuous on Θ for each $(y_{1T}, \ldots, y_{TT}, x_{1T}, \ldots, x_{TT})$ and Θ is compact by Assumption 4, the NLS estimator $\hat{\theta}_T$ exists and is Borel measurable (Pötscher and Prucha, 2013).

We need to make additional assumptions about the functions g and K, where $K(x, \theta_0) := \left.\frac{\partial g(x,\theta)}{\partial \theta}\right|_{\theta=\theta_0}$, to show that, under the null, the NLS estimator is consistent and to derive its asymptotic distribution in Proposition 1 below. Assumption 5 guarantees that the limit of the objective function is minimized (a.s.) at the true parameter vector θ_0.

Assumption 5 *For some $s \in [0, 1]$ and all $\theta \neq \theta_0$,*

$$g\left(B^0_{1,\Omega}(s), \theta\right) \neq g\left(B^0_{1,\Omega}(s), \theta_0\right) \quad \text{(a.s.)}.$$

Assumption 6 shall allow to establish the limiting distribution of the NLS estimator.

Assumption 6

$$\int_0^1 K\left(B^0_{1,\Omega}(s), \theta_0\right) K\left(B^0_{1,\Omega}(s), \theta_0\right)' \mathrm{d}s > 0 \quad \text{(a.s.)}.$$

Proposition 1 *Suppose that Assumptions 1–6 hold. Then, under H_0,*

$$\begin{aligned} T^{1/2}\left(\hat{\theta}_T-\theta_0\right) \xrightarrow{w} & \left(\int_0^1 K\left(B_{1,\Omega}^0(s),\theta_0\right) K\left(B_{1,\Omega}^0(s),\theta_0\right)' \mathrm{d}s\right)^{-1} \\ & \cdot\left(\int_0^1 K\left(B_{1,\Omega}^0(s),\theta_0\right) \mathrm{d}B_{0,\Omega}(s)+\int_0^1 K_1\left(B_{1,\Omega}^0(s),\theta_0\right) \mathrm{d}s\kappa\right) \\ =:& \psi\left(B_{1,\Omega}^0,\theta_0,\kappa\right), \end{aligned}$$

where $K_1(x,\theta)=\left.\frac{\partial K(x,\theta)}{\partial x'}\right|_{\theta=\theta_0}$ *and* $\kappa=\sum_{j=0}^{\infty}E(\theta_{1,0}\theta_{0,j})$.

Proof. The proof can be directly adapted from the proof of Theorem 2 in Saikkonen and Choi (2004) and Theorem A.1 in Choi and Saikkonen (2010). ■

3.2 Test statistic and large sample behavior

This subsection introduces the test statistic we work with and establishes its large sample behavior. In order to test for cointegration we test for the stationarity of the error process u_t. The test is residual-based and builds on to the cointegration test of Shin (1994), which, in turn, is based on the KPSS test (Kwiatkowski et al., 1992). We use the test statistic

$$\hat{\eta} := (T^2\hat{\omega}_u^2)^{-1}\sum_{t=1}^{T}\left(\sum_{j=1}^{t}\hat{u}_j\right)^2, \tag{2}$$

where $\hat{u}_t := y_t - g(x_{tT},\hat{\theta}_T)$ and

$$\hat{\omega}_u^2 := \hat{\omega}_u(l)^2 := T^{-1}\sum_{t=1}^{T}\hat{u}_t^2 + 2T^{-1}\sum_{s=1}^{l}w(s,l)\sum_{t=s+1}^{T}\hat{u}_t\hat{u}_{t-s},$$

where w is a kernel which fulfills, e.g., the conditions of Andrews (1991) and the lag truncation parameter $l := l_T$ depends on the sample size. Here, $\hat{\omega}_u^2$ is a consistent estimator of the long-run variance, as long as $T/l \to \infty$ for $T \to \infty$.

The linear case without autocorrelation gives us the model of Cavaliere and Taylor (2006). We may then use the parametric estimator

$$\hat{\sigma}_u^2 := T^{-1}\sum_{t=1}^{T}\hat{u}_t^2 \tag{3}$$

for the variance. In this case one can show that $\hat{\sigma}_u^2$ is consistent similarly as in Cavaliere and Taylor (2006).

Under the null hypothesis, we obtain the following asymptotic behavior of the test statistic.

Theorem 1 *Under the Assumptions 1–6 and the H_0*

$$\hat{\eta} \xrightarrow{w} \bar{\omega}_u^{-2} \int_0^1 \left(B_{0,\Omega}(s) - F(s, B_{1,\Omega}^0, \theta_0)' \psi(B_{1,\Omega}^0, \theta_0, \kappa)\right)^2 \mathrm{d}s, \tag{4}$$

where $F(s, B_{1,\Omega}^0, \theta_0) := \int_0^s K(B_{1,\Omega}^0(r), \theta_0)\mathrm{d}r$ and $\psi(B_{1,\Omega}^0, \theta_0, \kappa)$ is defined in Proposition 1.

As the variance profile $\Sigma(s)$ and thus $\Omega(s)$ is generally unknown, we see that the limiting distribution depends on nuisance parameters, which makes tabulated critical values impractical. The bootstrap, discussed in Section 3.3, is a natural solution.

Under the alternative asymptotic theory becomes even more tedious. Since the NLS estimator $\hat{\theta}_T$ is not consistent anymore a limiting distribution is hard to derive. We may, however, establish the order of magnitude of $\hat{\eta}$ under H_1, which is enough to justify consistency of the cointegration test.

Theorem 2 *Let H_1 be true. Under Assumptions 1–6, $\hat{\eta} = O_p(T/l)$, where l is the lag truncation used in the estimation of $\hat{\omega}_u^2$.*

3.3 Bootstrap procedure

We adopt a bootstrap solution to provide feasible inference building on Cavaliere and Taylor's (2006) bootstrap test for linear cointegration in the presence of variance breaks. They used the *heteroskedastic fixed regressor bootstrap* by Hansen (2000). It treats the regressors as fixed, without imposing strong assumptions on the data generating process (DGP). In Theorem 3 we show that the fixed regressor bootstrap replicates the correct asymptotic distribution of the test statistic. As usual, it does not replicate the finite sample distribution of the test statistic, see Hansen (2000). However, Section 4 will demonstrate that the bootstrap works well in finite samples, as also observed by Cavaliere and Taylor (2006) for testing linear cointegration. Popular other bootstraps, e.g., block resampling (Lahiri, 1999), are not applicable because the regressor is integrated and heteroskedastic and the error term is potentially heteroskedastic under the null hypothesis.

More specifically, the heteroskedastic fixed regressor bootstrap works as follows:

1. Run the original NLS regression, save residuals $\hat{u}_t$ and compute the test statistic $\hat{\eta}$ as given in (2).

2. Construct the bootstrap sample $y_{tT}^b := u_t^b := \hat{u}_t z_t$, $t = 1, \ldots, T$, where $\{z_t\}$ is a sequence of i.i.d. standard normal variates.

3. Estimate $\hat{\theta}_T^b$ via NLS of y_{tT}^b on $g(x_{tT}, \theta)$, save the residuals $\hat{u}_t^b := y_{tT}^b - g(x_{tT}, \hat{\theta}_T^b)$ and compute the bootstrap test statistic as

$$\hat{\eta}^b := (T^2(\hat{\omega}_u^b)^2)^{-1} \sum_{t=1}^{T} \left(\sum_{j=1}^{t} \hat{u}_j^b \right)^2,$$

   where $(\hat{\omega}_u^b)^2$ is the long-run variance estimate using the bootstrap sample.

4. Repeat steps 2 and 3 independently B times and, given that we reject for large values, compute the simulated bootstrap p-value $\tilde{p}_T^b := 1 - \tilde{G}_T^b(\hat{\eta})$, where $\tilde{G}_T^b$ is the empirical cumulative distribution function of the bootstrap test statistics $\{\hat{\eta}^b\}_{b=1}^B$.

The replications, for B sufficiently large, approximate the true bootstrap distribution G_T^b which is the theoretical cumulative distribution function of $\hat{\eta}^b$ and the associated bootstrap p-value is defined as $p_T^b := 1 - G_T^b(\hat{\eta})$. Then, as $B \to \infty$, $\tilde{p}_T^b \xrightarrow{\text{a.s.}} p_T^b$.

The next theorem shows that (i) the bootstrap replicates the correct asymptotic null distribution, and, (ii) that the test based on the bootstrap p-values is consistent.

Theorem 3 *(i) Under Assumptions 1–6 and the H_0, $p_T^b \xrightarrow{w} \mathcal{U}[0, 1]$.*

(ii) Under Assumptions 1–6 and the H_1, $p_T^b \xrightarrow{p} 0$.

3.4 Subresidual tests

Choi and Saikkonen (2010) proposed a KPSS type test for cointegration using subresiduals which we describe below. Its advantage is that the limiting distribution of the test statistic, under homoskedasticity, is nuisance parameter-free and explicitly given although, for nonlinear cointegration, the limiting distribution of the original test statistic was of the form like that in Theorem 1.

However, in the presence of variance breaks the limiting distribution of the subresidual-based statistic depends on nuisance parameters, as we will show in Corollary 1. This makes its direct use impractical. We hence favor the bootstrap approach.

The subresidual-based test statistic is of the same form as $\hat{\eta}$ in (2) but use only a subset of the residuals $\{\hat{u}_t\}_{t=i}^{i+\ell-1}$. We define

$$\hat{\eta}^{i,\ell} = (\ell^2(\hat{\omega}_u^\ell)^2)^{-1} \sum_{t=i}^{i+\ell-1} \left(\sum_{j=i}^{t} \hat{u}_j \right)^2 .$$

The index i is the starting point of the subresiduals and ℓ denotes the size of the set of subresiduals, also called block size. $(\hat{\omega}_u^\ell)^2$ is the long-run variance estimate using the subset of residuals. Then we have the following

Corollary 1 *Suppose that Assumptions 1–6 and H_0 hold. If $\ell \to \infty$ and $\ell/T \to 0$ as $T \to \infty$, we have for any i with $1 \leq i \leq T - \ell$ that*

$$\hat{\eta}^{i,\ell} \xrightarrow{w} \bar{\omega}_u^{-2} \int_0^1 B_{0,\Omega}^2(s) \mathrm{d}s. \tag{5}$$

Choi and Saikkonen (2010) found that, under homoskedasticity, $\hat{\eta}^{i,\ell}$ weakly converges to

$$\int_0^1 W^2(s) \mathrm{d}s, \tag{6}$$

where $W(s)$ is a standard Brownian motion. Moreover, they derived the distribution function of (6) and provided an easy series representation. This makes the residual approach easy to use. However, for heteroskedastic errors the variance terms in (5) do not cancel out in general. Thus, the limiting distribution depends on nuisance parameters.

For comparative purposes, we still use the distribution of (6) for testing the null of nonlinear cointegration in the Monte Carlo experiments in Section 4, ignoring potential heteroskedasticity. This is because we want to investigate the impact of variance breaks for the approach. Moreover, we will compare it with the bootstrap test.

The c.d.f. of (6) is given by

$$F(z) = \sqrt{2}\sum_{n=0}^{\infty} \frac{\Gamma(n+1/2)}{n!\Gamma(1/2)}(-1)^n \left(1 - Erf\left(\frac{\sqrt{2}/2 + 2n\sqrt{2}}{\sqrt{2}z}\right)\right), \; z \geq 0,$$

where $Erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-y^2)\mathrm{d}y$ is the error function. Choi and Saikkonen (2010) demonstrated that truncating the series at $n = 10$ is sufficiently accurate, and we follow their choice.

In order to aggregate subsample tests by using different starting points i Choi and Saikkonen (2010) proposed a Bonferroni procedure. For this, we compute M test statistics $\hat{\eta}^{i_1,\ell}, \ldots, \hat{\eta}^{i_M,\ell}$ and define $\hat{\eta}^{\max,\ell} := \max\left\{\hat{\eta}^{i_1,\ell}, \ldots, \hat{\eta}^{i_M,\ell}\right\}$. Due to the Bonferroni-inequality $\lim_{T\to\infty} P\left(\hat{\eta}^{\max,\ell} \leq c_{\alpha/M}\right) \geq 1 - \alpha$, where $c_{\alpha/M}$ is the α/M-critical value from the distribution of $\int_0^1 W^2(s)\mathrm{d}s$. We choose $M = \lceil T/\ell \rceil$ and ℓ like in Choi and Saikkonen (2010) with the minimum volatility rule proposed by Romano and Wolf (2001).

4 Monte Carlo study

This section provides evidence that the proposed nonlinear cointegration test works well in finite samples. We conduct several simulation studies for different settings. Especially, we study the proposed bootstrap test for linear, polynomial, and smooth transition regression cointegration. We compare the empirical rejection rates with those of the standard Shin (1994) test. Moreover, we compare the bootstrap cointegration test with the subresidual-based approach. For the DGP we extend the example of Cavaliere and Taylor (2006), who generated data with a linear cointegration relation under variance breaks, by also considering nonlinear cointegration. We start with the linear case.

4.1 Linear regression model

We consider the DGP

$$y_t = x_t + u_t, \quad t = 1, \ldots, T, \tag{7}$$

$$u_t = \rho u_{t-1} + \zeta_{u,t} + \mu_t, \quad u_0 = 0 \tag{8}$$

$$\mu_t = \mu_{t-1} + \rho_\mu \zeta_{\mu,t}, \quad \mu_0 = 0, \tag{9}$$

$$x_t = x_{t-1} + \zeta_{x,t}, \quad x_0 = 0, \tag{10}$$

where $\zeta_t := (\zeta_{u,t}, \zeta_{x,t}, \zeta_{\mu,t})' = \Sigma_t^{1/2}\zeta_t^*$, $\zeta_t^* \sim N(0, I_3)$, i.i.d., $|\rho| < 1$ and

$$\Sigma_t := \begin{pmatrix} \sigma_{u,t}^2 & \sigma_{ux,t} & 0 \\ \sigma_{ux,t} & \sigma_{x,t}^2 & 0 \\ 0 & 0' & \sigma_{\mu,t}^2. \end{pmatrix}$$

In particular, here we initially consider the case of a simple linear cointegrating regression with a single non-deterministic integrated regressor.

We consider abrupt variance breaks of the form

$$\begin{aligned} \sigma_{u,t}^2 &= \sigma_{u,0}^2 + (\sigma_{u,1}^2 - \sigma_{u,0}^2)\mathbb{1}\left(t \geq \lfloor \tau_u T \rfloor\right) \\ \sigma_{x,t}^2 &= \sigma_{x,0}^2 + (\sigma_{x,1}^2 - \sigma_{x,0})^2\mathbb{1}\left(t \geq \lfloor \tau_x T \rfloor\right) \\ \sigma_{\mu,t} &= \sigma_{\mu,0}^2 + (\sigma_{\mu,1}^2 - \sigma_{\mu,0}^2)\mathbb{1}\left(t \geq \lfloor \tau_\mu T \rfloor\right). \end{aligned}$$

In all simulations we set $\sigma_{u,0}^2 = \sigma_{x,0}^2 = \sigma_{\mu,0}^2 = 1$.

As Cavaliere and Taylor (2006) noted under the null hypothesis $\rho_\mu^2 = 0$ four cases can occur: (i) if $\tau_u = \tau_x = 0$, then y_t and x_t are both standard $I(1)$ processes with homoskedastic increments and cointegrated; (ii) if $\tau_u \neq 0, \tau_x = 0$ the permanent shocks to the system are homoskedastic (i.e., x_t is integrated with homoskedastic innovations) but there is a variance shift in both the transitory component of y_t and in the cointegrating relation; (iii) if $\tau_u = 0, \tau_x \neq 0$, the permanent shocks to the system are heteroskedastic with changes to both x_t and y_t being heteroskedastic, but there are no variance shifts in the cointegrating relation; (iv) if $\tau_u \neq 0, \tau_x \neq 0$, the permanent shocks to the system are heteroskedastic, changes to both x_t and y_t are heteroskedastic and there is a variance shift both in the transitory component of y_t and in the cointegrating relation. If H_0 is true variance shifts in ζ_μ have no influence. Under the alternative we also allow for variance breaks in ζ_μ which lead to variance breaks in u_t which are similar to cases (ii) and (iv).

Moreover, we consider covariance breaks of the form

$$\sigma_{ux,t} = \sigma^{ux,0} + (\sigma_{ux,1} - \sigma_{ux,0})\mathbb{1}\left(t \geq \lfloor \tau_{ux} T \rfloor\right).$$

In our simulations we only consider the case where all variance shifts occur at the same time, i.e., $\tau := \tau_u = \tau_x = \tau_\mu = \tau_{ux}$. For the results on other possible scenarios see the simulation study of Cavaliere and Taylor (2006).

We investigate the following parameter constellations. Let the sample size be $T \in \{100, 300\}$. We take $\rho_\mu^2 \in \{0, 0.001, 0.01, 0.1\}$. $\rho_\mu^2 = 0$ is to estimate size, the other constellations are for a power analysis. We consider variance breaks at $\tau \in \{0, 0.1, 0.5, 0.9\}$. While the first of the τ-values corresponds to the case of no variance breaks the latter stand for early, middle, and late variance breaks. We also fix the magnitude of the variance breaks by setting $\sigma_1^2 = \sigma_{u,1}^2 = \sigma_{x,1}^2 = \sigma_{\mu,1}^2 \in \{1/16, 16\}$, like in Cavaliere and Taylor (2006). The parameter for the covariance $\sigma_{ux,t}$ are chosen in such a way that the correlation between $\zeta_{u,t}$ and $\zeta_{x,t}$ is fixed over time at $\lambda \in \{0, 0.5\}$, i.e., without or with endogeneity. The $AR(1)$ parameter of u_t is set $\rho \in \{0, 0.5\}$. Empirical rejection rates are based on 10,000 replications (unless stated otherwise) and the number of bootstrap replications is $B = 500$. Finally, the nominal level of significance is $\alpha = 0.05$ for the remainder of this paper.

We perform the test by estimating θ in the linear regression y_t onto $g(x_t, \theta) \equiv \theta x_t$ and using the residuals to compute $\hat{\eta}$.[1] We use the estimator $\hat{\sigma}_u^2$ given in (3) for $\rho = 0$ and, for $\rho = 0.5$, a non-parametric autocorrelation-robust estimator for the long-run variance with a Bartlett kernel and a spectral window of $\lfloor 4(T/100)^{0.25} \rfloor$ as suggested in Kwiatkowski et al. (1992). Table 1 reports empirical rejection rates (as percentages) for the different parameter constellations. Panel (a) shows the rates for the bootstrap approach, panel (b) for the subsample approach and panel (c) for the standard Shin (1994) test. First, the bootstrap generally yields very good empirical sizes and powers. Both time (early or late) and direction (increase or decrease) of a variance break do not

[1]While we formulate the theory for nonlinear cointegrating regressions we for simplicity use the OLS estimator whenever possible to speed up the computations.

have a notable impact on the rejection frequencies. For example, early downward variance breaks yield lower empirical power than early upward variance breaks, and vice versa for late variance breaks. This effect reduces with increasing ρ_μ^2.

The subsample-based test is undersized in the constellation without heteroskedasticity under absence of endogeneity and autocorrelation. Interestingly, it is oversized under endogeneity and autocorrelation, especially in the presence of early downward variance breaks. This effect reduces if the shifts occur later. Moreover, the bootstrap test has higher power for $\rho = 0$, especially if the alternative is close to the null, otherwise the subresidual test has higher power.

Panel (c) shows the result for the test based on critical values tabulated by Shin (1994). We observe that variance breaks are an issue and that the test oversizes or undersizes depending on downward or upward breaks. The empirical power is generally smaller than for the bootstrap test.

4.2 Polynomial cointegrating regression

In this subsection, we consider the case of polynomial cointegrating regression, in particular a quadratic and a cubic relation. We replace the linear model (7) and simulate according to

$$y_t = x_t + x_t^2 + u_t,$$

for the quadratic relation, while (8), (9) & (10) and all further parameter constellations of Subsection 4.1 still hold. We now estimate $\theta = (\theta_1, \theta_2)'$ by regressing y_t on $g(x_t, \theta) = \theta_1 x_t + \theta_2 x_t^2$. In this model, we already cannot use the critical values of Shin (1994) because to consider both x_t and x_t^2 as integrated regressors violates the model assumptions. This is also discussed in Wagner and Hong (2016).

Table 2 shows the tests' rejection frequencies. Similar interpretations like in Subsection 4.1 for the linear case apply here, too. In addition, we observe a decrease of empirical power relative to Table 1, plausibly due to the more complex model to be fitted. The loss is more moderate for the bootstrap test.

Inspired by the application in Section 5, we also consider a cubic cointegrating regression. We simulate from the model

$$y_t = x_t + 2x_t^2 + x_t^3 + u_t,$$

Table 1: The table reports the empirical rejection frequencies for testing the null of cointegration in the *linear* regression model for various parameter constellations. All rejection rates are given as percentages. The nominal size is 5%. Panel (a) is for the bootstrap test, panel (b) for the subresidual-based test and panel (c) for the Shin (1994) test.

				ρ_μ^2:	0				0.001				0.01				0.1			
				ρ:	0		0.5		0		0.5		0		0.5		0		0.5	
	T	τ:	σ^2	λ	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5
(a)	100	0			5.3	4.7	8.3	6.9	15.6	15.4	16.8	16.8	50.1	50.9	37.6	37.2	84.9	84.3	50.4	50.9
		0.1	1/16		6.7	6.4	6.4	6.4	17	17.2	18.7	18.5	43.3	43.9	37.1	37.2	77	78.4	51.4	52.6
			16		5.1	4.4	7.3	6.6	13.6	13.9	14.5	14.1	47.2	47.3	34	33.9	82.3	82.3	46.5	46.6
		0.5	1/16		5.3	4.7	7.2	6.6	13.9	13.9	15.2	14.9	42.3	43.1	31.7	32.8	77.3	78.1	47.5	48
			16		4.8	4.8	5.6	4.9	11	11.1	10.6	10.2	42.4	43.8	29.5	29.8	82.2	82	42.7	42.3
		0.9	1/16		5.2	4.5	7.7	6.4	15.2	14.8	16.3	16	51	51.1	37.9	37.1	84.6	85.4	50.1	50.9
			16		4.7	5.7	6.9	6.2	11.9	12.7	12.7	13.3	40.1	40	31.9	31.8	77.3	77.2	45.6	46.2
	300	0			4.8	4.1	7.1	6.2	48.4	49.5	43.8	43.1	87.1	87	65	65.6	98.5	98.6	70.5	70.4
		0.1	1/16		5.6	5.1	6.4	6.1	38.5	37.8	38.5	37.3	77	76.5	60.3	60	96.3	96	70.3	69.9
			16		5.1	4.3	6.9	6	45.8	46.3	40.4	40.6	85.2	85.5	61.7	62.5	98.4	98.5	68.1	67.9
		0.5	1/16		5.1	5.1	7.2	6.5	41.4	42.5	37.3	38.2	79.9	80.2	60	60.4	96.1	96	68.4	67.7
			16		5	5.2	6.4	6.8	43.1	43.2	38.3	38	85.6	85.8	60.1	60.3	98.9	98.5	64.2	64.3
		0.9	1/16		4.9	4.4	7.1	6.1	50.7	50.5	44.9	44.1	87.1	87.3	65.7	66	98.7	98.6	71.6	70.5
			16		5.4	5.6	6.7	6.8	39.4	38.7	37.5	37	79.9	80.7	58.8	60.3	97.5	97.5	64.6	65.3
(b)	100	0			2	1.6	8.1	7.7	8.5	8.8	18.5	19	40.2	40.8	44.3	44.6	79.4	79.2	62.7	63.1
		0.1	1/16		6.3	6.6	10.8	13	17.1	17.6	26.2	27.9	44.5	45.1	48.4	49.1	78.9	78.2	65.8	65.4
			16		1.3	1.1	5.7	6.6	5.9	6	14.3	14.3	32.6	34.3	35.8	37.3	74.7	74.2	55.2	55.1
		0.5	1/16		4.6	6.8	14.2	19.1	19.7	21.6	31.1	33.4	48.2	49.2	52.9	53.9	80.8	81.2	69.2	69
			16		2.2	1.5	7.1	6.2	4.3	4	12.7	10.9	29.2	28.4	34.8	34.7	75.2	74.8	57.8	56.4
		0.9	1/16		2.2	1.8	8.4	8.1	9.2	9	19.1	19.1	41.7	41.7	44.8	44.6	79.8	79.6	63.4	63.2
			16		5	4.9	11	12	9.1	8.8	19.2	19.7	37.5	37.4	43.9	44.7	78	79.3	64.8	65.7
	300	0			1.8	1.6	4.3	4.4	38.4	37.6	41.4	40.4	83.7	84.3	69.9	70.5	98.1	98.2	77.1	76.9
		0.1	1/16		6.4	6.5	9.5	10.3	43.4	44.8	47.4	48.3	81.9	82.5	71.3	70.6	97.9	97.8	78.1	78
			16		1	1	3.1	3.1	32	31.3	34.2	33.6	78.2	79.2	62.1	62.6	97.1	96.9	71.7	70.5
		0.5	1/16		5.5	8.9	12	17.1	50.1	51.4	55.1	55.4	85.6	85.3	77.4	76.9	98.2	98.1	83	83
			16		3.4	2.4	6.8	5.2	34.7	33.9	36.9	37.5	83	83	69.3	69	97.9	98	78.1	77.4
		0.9	1/16		2.1	1.8	5.1	4.7	40.6	40.3	43.6	43.2	84.5	84.3	71	70.6	98.4	98.2	76.6	76.8
			16		4.5	4.1	7.9	7.2	34.4	34.7	38.9	39.3	80.9	81.4	70.2	70.3	97.5	97.8	77.9	77.9
(c)	100	0			5	3.7	10	8.1	15.5	14.7	19.9	18.8	46.8	45.8	37.4	36.7	79.6	79.3	45.8	46.1
		0.1	1/16		12.3	10.7	12.4	11.7	22.2	20.9	22.8	21.9	45	44.6	35.5	35.2	76.5	75.5	42.6	42.9
			16		4.2	3.3	8.7	7.2	12.4	11.6	16.5	15.5	40.3	41.6	32.6	33.2	75.8	75.7	41.6	41.9
		0.5	1/16		9.5	9.2	15.1	13.4	19.5	19.4	23.8	22.6	46.1	46.3	37.5	36.5	79.3	79.3	47.2	46.3
			16		1.7	1.5	4.4	4.2	4.4	4.3	8.6	8.4	27.9	27.9	25.2	25.5	68.2	67.6	38.1	38
		0.9	1/16		6.1	4.6	11	8.6	16.6	15.7	20.5	19.3	47.7	47.7	37.4	37.5	80.6	80.9	47	47.1
			16		4.2	4.8	9.6	10.4	9.3	10.5	16.8	17.8	35.8	35.8	36.3	36	71.2	72.2	48.1	48.6
	300	0			5	4	7.6	6.6	45.6	44.7	40.4	39	85.2	84.7	56.8	57.4	97.8	98.1	60.9	60.5
		0.1	1/16		12.1	10.9	13.1	11.5	44.8	45	40.4	40.1	79.4	80	54.5	54.1	96.8	96.7	57.7	57.2
			16		4	3.5	6.5	5.6	40.8	40.5	35.9	35.7	80.6	81.1	52.2	52.5	97.2	96.9	55.9	55
		0.5	1/16		9.9	8	12.9	10.7	47	46.3	41.7	41.3	82.7	81.9	57.3	57.5	97.4	97	61	62
			16		1.3	1.3	2.6	2.7	27.7	28	25.2	25.8	74.5	74	47.3	46.7	95.6	95.7	51.8	52.2
		0.9	1/16		6.2	4.4	9.2	6.9	49	48.2	43	41.9	86	85.5	57.7	57.2	98.2	98.2	59.8	60.7
			16		3.9	4.5	6.7	7.4	34.6	35.1	35.7	36	75.9	76.2	54.6	55.4	95.6	95.8	59.5	59.3

Table 2: The table reports the empirical rejection frequencies for testing the null of cointegration in the *quadratic* regression model for various parameter constellations. All rejection rates are given as percentages. The nominal size is 5%. Panel (a) is for the bootstrap test and panel (b) for the subresidual-based test.

				σ_μ^2:	0				0.001				0.01				0.1			
				ρ:	0		0.5		0		0.5		0		0.5		0		0.5	
	T	τ:	σ^2	λ	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5
(a)	100	0			4.9	5.6	7.1	7	12.2	12.6	13.1	13.8	44.2	46.1	30.8	31.7	82.2	81.1	42.1	42.1
		0.1	1/16		6.4	7.3	4.9	6.1	14.5	15.3	14.1	14.7	35.8	36.7	29.1	29.8	68.9	70.4	39.9	40.8
			16		4.9	5.9	6.3	6.7	11.4	11.8	11.6	11.4	40.7	41.3	26.9	26.8	79.4	78.6	37.6	38.1
		0.5	1/16		5	5.6	5.7	6	11.4	11.1	11.5	11.5	35.5	36.6	25.6	26.2	72.5	72.6	37.9	37.7
			16		5.2	5.7	4.2	4.2	9.3	9.8	7.3	7.3	35.1	35.8	21.1	21.2	76.8	76.8	33.8	33.4
		0.9	1/16		5	5.5	6.5	7.1	12.2	12.5	12.5	12.6	44.5	44.7	30.6	30.1	81.8	81.7	41.8	42.1
			16		5.3	5.7	5.1	5.6	10.3	10.7	9.9	10.6	34.5	35.5	26.7	26.9	72.8	73.7	39.8	39.9
	300	0			4.9	5.3	6.9	7.1	42.7	44.3	38.3	38.6	85.1	84.8	60.9	60.3	98.2	98.3	64.4	64.9
		0.1	1/16		5.8	5.4	5.8	6.2	32	30.9	31.3	30.3	68.8	68	51.1	50	94.2	93.8	58.3	58.3
			16		5.1	5	6.9	6.9	40.6	40.5	35.1	35.2	82.5	83	56.6	57.4	97.8	97.8	61.6	61.6
		0.5	1/16		5.2	5.7	6.7	6.9	35.4	36	31.7	32	75.8	75.5	52.7	52.8	95.8	95.4	59.8	59.7
			16		4.8	5.6	6.1	7.1	35.1	35.6	30.3	30	81.3	81.1	53.3	52.5	97.8	97.6	57.2	56.6
		0.9	1/16		4.9	5.2	6.6	6.7	44.5	44.4	38.9	38.2	84.1	84.9	59.9	60.1	98.2	98.3	65.4	65.1
			16		5.2	5.9	5.8	6.5	34.5	34.8	32	32	75.9	76.3	54.7	55.8	96.2	96.8	60.5	60.7
(b)	100	0			1.1	0.9	5.2	5.5	4.6	5.2	12	13	30.2	30.9	34.2	34.9	71	71.2	53.7	54.5
		0.1	1/16		4.5	4.9	8	9.3	11.8	12.4	18.5	19.5	32.6	33.5	36.4	37.1	66.8	66.1	52.9	52.1
			16		0.7	0.7	4	4.4	3.2	3.4	8.9	9.6	23.2	24.8	26.9	28.3	64.8	64.8	46.3	46.4
		0.5	1/16		3.4	4.3	9.7	12.6	12.7	13.7	20.3	22.4	35.1	35.8	39.4	39.3	68.6	68.5	54.7	54
			16		3	2.5	6.9	6.6	4.2	4.2	10.3	9.7	22.3	21.8	27.3	26.9	67.1	67.6	48.7	48.1
		0.9	1/16		1.5	1	5.6	5.8	5.1	5.4	12	12.9	31.1	30.8	34.7	33.8	71.6	71.1	53.9	52.9
			16		4.1	4.8	8.3	8.9	6.6	6.9	13.4	14	28.9	29.4	35.1	35.4	69	70.5	55.4	56.7
	300	0			1.1	0.9	3	3	28.8	28.6	31.9	31	75.3	76.1	61.2	61.3	96.2	96.4	69.8	69.7
		0.1	1/16		4.7	4.9	7	7.7	32.6	33.4	35.3	36.1	70.3	71.4	59	58.8	94.9	94.7	67.9	67.9
			16		0.6	0.7	2	2.1	23.3	22.7	25.7	24.9	69.9	69.5	53.5	53.3	94.7	94.3	64.1	63.5
		0.5	1/16		4	6.1	8.4	11.6	38	37.5	41.2	40.9	75.2	74.5	63.4	62.8	95.7	95.6	69.9	69.9
			16		4.9	5.2	7.8	7.2	28.8	28.6	30	30	77.8	78	61.1	60.4	96.7	96.6	70.7	70.3
		0.9	1/16		1.3	1.2	3.4	3.3	29.9	30.5	33	32.7	75.8	76.3	61.9	61.9	96.3	95.9	69	68.8
			16		4	4.6	5.6	6.6	27.3	27.5	30.6	31.2	73.2	73.9	61.3	61.7	95.3	95.6	70.4	70.7

Table 3: The table reports the empirical rejection frequencies for testing the null of cointegration in the *cubic* regression model for various parameter constellations. All rejection rates are given as percentages. The nominal size is 5%. Panel (a) is for the bootstrap test and panel (b) for the subresidual-based test.

				ρ_μ^2:	0				0.001				0.01				0.1			
				ρ:	0		0.5		0		0.5		0		0.5		0		0.5	
	T	τ:	σ^2	λ	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5
(a)	100	0			4.9	5.8	6.3	6	10.7	11.1	10.4	11.2	40.6	42	26.1	26.5	80.7	79.9	35.9	36
		0.1	1/16		5.9	6.9	4.8	5.8	13.6	14.2	12.7	13.2	34.6	35.5	27.3	27.8	68.8	70.1	38.5	39.8
			16		4.7	5.6	5.3	5.5	9.9	10.6	8.8	9.3	36.9	38.1	21.9	21.7	77.5	77.3	31.5	31.5
		0.5	1/16		4.9	5.1	5	4.8	9.6	10.2	9.3	9.3	31.9	33	21.3	21.6	68.9	69.2	32.3	31.8
			16		4.9	5.3	3.6	3.4	8.1	8.6	5.9	5.4	29.3	30.8	14.9	15.4	73.2	72.5	25.6	24.8
		0.9	1/16		4.8	5.2	5.8	5.8	10.9	11.2	10	10.3	41.6	40.9	25.8	24.9	80.6	80.6	35.7	36
			16		5.2	5.1	4.8	4.7	9	9.5	8.2	8.5	30.6	31.1	22.7	22.5	70.5	71	35.2	35.4
	300	0			5	5.4	7	7	39.1	40.9	34.7	34.3	83.5	83.5	58	56.9	98.2	98.2	62.2	62.4
		0.1	1/16		5.3	5.8	5.5	5.8	30.5	29.7	29.8	28.2	67.7	67.4	50.1	49.3	94	93.6	57.9	58.1
			16		4.8	5.2	6.4	6.6	37	36.4	31.3	30.6	80.9	81.6	53.4	53.8	97.5	97.8	58.2	58.1
		0.5	1/16		5.1	5.2	6.2	6.2	31.4	32.5	27.4	27.7	72.7	72.6	48.6	48.3	95.5	94.8	54.6	55.3
			16		4.9	5.9	5.4	6.5	29.8	31	24.9	24.5	77.2	77.8	46.5	46.3	97.5	96.9	49.7	49.9
		0.9	1/16		5.1	5.4	6.6	6.9	40.5	40.6	34.9	34.7	82.9	83.3	57	57.5	98.1	98.3	63.7	62.3
			16		5.6	5.5	5.7	5.6	30.3	30.2	28.3	27.5	73.4	73.8	51.2	52.2	95.9	96.3	57.6	57.7
(b)	100	0			0.8	0.9	3.8	4.3	3	3.4	9.1	9.7	23.7	24.5	28	28.1	64	64.1	46.4	47.3
		0.1	1/16		3.3	3.6	5.9	7.1	8.4	9.1	14.4	15.4	27.5	27.6	30.7	30.9	61	60.1	47.6	46.8
			16		0.5	0.6	3.1	3.4	2	2.3	6.5	7.1	17.9	19	22.2	22.9	57	57.7	39	39.5
		0.5	1/16		2.9	3	7.6	9.2	9.7	10.3	16.1	17.6	28.5	29.2	32.9	33.1	61.5	61.4	48	47.7
			16		3.1	2.8	5.9	5.8	4.1	3.8	8	7.7	16.8	17	20.9	21	58	57.8	38.2	38
		0.9	1/16		1	0.6	4.4	4.6	3.3	3.6	8.8	10	24.8	24.2	28.5	27.3	63.3	63.4	46.5	46.9
			16		3.6	3.9	6.4	6.9	5.3	5.3	10.5	11.1	23.6	24.3	29.4	29.3	62.3	64.7	49	49.6
	300	0			0.8	0.8	2.4	2.2	23.2	22.7	25.7	25.1	68.8	69.7	52.9	53.8	93.8	93.9	62.1	62.5
		0.1	1/16		3.4	3.8	5	5.6	27.7	28	30.5	30	65.2	65.8	54	53.2	92.2	92.5	63	63.5
			16		0.4	0.5	1.5	1.6	17.5	18.2	20.2	19.4	63	63.2	46.2	45.7	92	92	56.8	55.9
		0.5	1/16		4	4.2	7.3	8.1	30.7	30.9	34.2	34.1	68.4	68.2	56.1	55.5	93.1	93.3	63.8	63.8
			16		5.3	6.1	7.4	7.2	23.2	23.6	23.3	23.8	70.8	71.2	51.8	51	94.7	94.7	62.9	62
		0.9	1/16		0.8	1	2.4	2.7	23.2	24	26.2	25.8	68.9	69.7	54.1	54.3	93.6	93.9	61.3	61.6
			16		3.9	3.6	5	4.4	22.5	22	25.2	24.9	67.8	67.9	54.4	54.8	93.2	93.6	64.3	65.1

where the remaining parameters are specified like in the linear and quadratic case. Table 3 shows the rejection frequencies. Again, size is well controlled for the bootstrap test, and we observe additional loss of power compared to the quadratic model (Table 2). The power loss is higher for the subresidual test.

4.3 Smooth transition regression model

We now discuss an example of a cointegrating regression which is indeed nonlinear in the parameters. Thus, NLS is needed for estimation. We adopt the example of cointegrating smooth transition functions which is also considered in Saikkonen and Choi (2004) and Choi and Saikkonen (2010). We generate data according to

$$y_t = \theta_0 + \theta_1 x_t + \theta_2 \frac{1}{1 + \exp(-(x_t - \theta_3))} + u_t,$$

with the parameter constellation $\theta_0 = 0, \theta_1 = 1, \theta_2 = 1, \theta_3 = 5$. In rare cases, for some generated samples the NLS algorithm does not converge. We thus exclude these cases from the analysis. To save computational time we run 1,000 repetitions for each constellation. Note that while the true parameter $\theta_0 = 0$ we include $\hat{\theta}_0$ in the estimation. This means we are in the setting beyond our model assumptions with an additional deterministic regressor. For a more detailed discussion see Appendix B.

Table 4 panel (a) reports the rejection rates for the bootstrap test and panel (b) for the subresidual-based test. We observe that the bootstrap test works well, again, with some moderate size problems in the presence of either endogeneity or autocorrelation (which can be solved using leads-and-lags as in Appendix B) and somewhat larger size distortions for both endogeneity and autocorrelation. The subresidual based test delivers mixed results, being is undersized and oversized for different scenarios of variance breaks.

5 Application

We now discuss an application of cointegrating polynomial regressions for the environmental Kuznets curve (EKC). It relates per capita GDP and per capita pollution of, e.g., CO_2 emissions.

Table 4: The table reports the empirical rejection frequencies for testing the null of cointegration in the *smooth transition* regression model for various parameter constellations. All rejection rates are given as percentages. The nominal size is 5%. Panel (a) is for the bootstrap test and panel (b) for the subresidual-based test.

				σ_μ^2:	0				0.001				0.01				0.1			
				ρ:	0		0.5		0		0.5		0		0.5		0		0.5	
	T	τ:	σ^2	λ	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5
(a)	100	0			4.5	7.5	5.7	10.9	12.1	14.8	12.6	18.4	43.8	48.2	38.3	38.8	88.1	89.8	53.6	56.8
		0.1	1/16		8.8	9.5	6.7	9.5	8.8	11.6	8.6	11.9	29.7	29.1	26.1	26.4	70.7	70.7	41.9	41.3
			16		4.4	8.8	7.4	13	11.1	15.9	14.8	18.4	47.3	50.8	40	41.2	89.5	91.5	56.2	58.9
		0.5	1/16		4.8	8.8	6.1	8.5	9	9.5	8.9	9.6	32.4	37.7	24.5	28	78.3	77.6	39.5	42.1
			16		6.1	7.3	5	6.7	8.6	12.8	9.3	10.6	35.2	35.8	22.7	23.8	79.9	82.5	34.5	36.6
		0.9	1/16		4.7	9.1	5.7	12.1	12	17.1	14.8	19.6	49.6	50.6	39.4	40.3	88.8	89.2	55.3	55.5
			16		6.4	8	6.5	9.2	9.1	10.1	8.2	11.9	26.3	30.8	23.6	27.4	70.2	70.4	39.2	40.9
	300	0			5.2	8.1	7.2	11.8	43.8	47.7	44.2	46	90.3	91.5	78	77.2	99.7	99.4	79.5	78.8
		0.1	1/16		7	7.5	7	9.4	21.6	27.1	24.1	27.7	69.5	69.8	52.4	54.4	96.5	96.8	63.6	60.4
			16		3.9	8.4	6.2	12.2	44.9	48.8	46.5	48.6	92.2	91.8	79.1	77.9	99.6	99.8	83.5	81.7
		0.5	1/16		5.5	8.9	7.2	10.9	30.9	36.2	30.8	34.7	80.5	82.5	64.5	65.7	98.9	99	70.4	70.7
			16		3.3	7.9	6.2	10.8	33.1	37.2	32.5	34.6	83.3	82.7	64.3	65.3	99.3	98.8	70.9	70.3
		0.9	1/16		6.4	10.1	9	14.1	48.4	47.8	47.5	46.7	90.8	91.2	76.7	75.6	99.8	99.8	81.5	81.6
			16		5.2	6.5	7.5	8.6	24.2	28.1	25.1	27.9	72.7	76.6	57.9	58.4	97.5	97.8	63.5	62.8
(b)	100	0			0.2	0.2	1	1.8	0.7	1.3	3	5.6	13.2	13.6	16.4	18.9	51.2	52.3	31.8	35.8
		0.1	1/16		1.2	1.1	3.2	3.5	1.4	2	3.8	5.6	8.1	7.9	13.3	13	40.6	40.2	30.1	30
			16		0	0.2	0.8	1.3	0.7	1.1	4.2	3.7	13.7	15.1	15.7	19.1	52.2	53.5	34.2	36.7
		0.5	1/16		2.6	3.3	4.5	7.9	3.6	4.8	7.1	10.7	13.6	15.1	18.5	19.5	41.3	43.6	30	33.7
			16		5.3	5.9	6.1	6.3	7.4	7.6	10.1	9.7	22.9	24	19.9	22.9	54.9	56.4	38.5	36.2
		0.9	1/16		0.1	0.1	1.4	0.7	0.5	0.9	2.8	2.7	11.5	11.6	13.7	15.8	46.8	47.7	30.2	32.2
			16		2.5	3.5	3.8	4	3.7	5.4	6.3	6.4	16.7	18.9	16.4	17.7	50.5	50.1	31.6	33.7
	300	0			0.1	0.1	0.2	1	13.2	14.3	16.8	14.9	59.3	58.4	42.5	37.8	91.4	91.6	46.9	48.5
		0.1	1/16		1	1.2	1.8	2.7	8.1	9.1	10.9	13.3	44.3	46.9	32	30.3	85	86.2	43.6	42.8
			16		0	0.2	0.3	1	14.1	14.2	13.1	14.7	59	59.2	41.3	40.6	92	92.1	48.2	50.5
		0.5	1/16		4.5	5.2	7.1	8.5	20.5	22.4	25.2	26.4	56.5	56.1	44.5	42.4	88.2	88.7	51	50.8
			16		9.7	11.8	9.9	11.9	31.4	32.9	32	28.3	69.9	69.8	53.6	50.5	92.8	92.7	57.6	59.1
		0.9	1/16		0.1	0.1	0.2	1	12.6	12.8	15.1	15	56.6	56.6	38.5	41.2	90	90	46.7	49.3
			16		2.6	4.2	3.9	4.8	17.3	20.2	16.9	20.1	58.5	58.6	45.9	41.1	90	90.4	48.5	49.1

The term EKC refers to the inverse U-shape relation of economic development and income inequality postulated by Kuznets (1955). Grossman and Krueger (1995) opened a very active literature with contributions in several directions. See Stern (2004) or Stern (2018) for a more recent survey.

We build on Wagner (2015) and Stypka et al. (2017) who argued that using an ordinary Shin (1994)-type linear cointegration test is inappropriate for polynomial cointegrating regressions (CPR). This is because if we include the k-th power x_t^k of an integrated regressor into the regression this power itself is not $I(1)$ anymore and thus violates the assumptions of the Shin (1994) test. Based on Wagner and Hong (2016) the aforementioned authors applied a fully modified OLS approach for CPRs. However, they did not allow for variance breaks in their approach, which could lead to erroneous inference regarding the EKC hypothesis. We apply the bootstrap discussed above to address this possible issue in the following.

We study data of 19 industrialized countries (see Table 5) over the period from 1870 to 2014.[2] We use per capita GDP data of the Maddison database (`https://www.rug.nl/ggdc/historicaldevelopment/maddison/`). CO_2 data is taken from the homepage of the Carbon Dioxide Information Analysis Center (`https://cdiac.ess-dive.lbl.gov/`) and is expressed as 1,000 tons per capita. We convert all time series to natural logarithms. Among others, Wagner (2015) also examined sulfur dioxide data, but discussion and results are similar. For brevity, we only focus on (the more relevant) CO_2 emissions. Let e_t denote log per capita GDP and y_t denote log CO_2 emissions per capita. We then study the model

$$e_t = c + \delta t + \theta_1 y_t + \theta_2 y_t^2 + \theta_3 y_t^3 + u_t.$$

To assess whether variance breaks are present in the error term we follow Cavaliere and Taylor (2008) and define the *empirical variance profile* as

$$\hat{\rho}(s) := \frac{\sum_{t=1}^{\lfloor Ts \rfloor} \hat{u}_t^2 + (sT - \lfloor Ts \rfloor)\hat{u}_{\lfloor Ts \rfloor + 1}^2}{\sum_{t=1}^{T} \hat{u}_t^2} \tag{11}$$

for $s \in (0, 1)$, with $\hat{\rho}(0) := 0$ and $\hat{\rho}(1) := 1$. In case of homoskedasticity, we should have $\hat{\rho}(s) \approx s$. Figure 1 plots the empirical variance profile for Australia, Austria, Belgium and Canada against

[2]New Zealand is an exception were data is available for 1878-2014.

s. Figures 2–5 for the remaining countries are given in Appendix C. We observe the presence of variance breaks for all countries (except maybe Denmark). For example, there is an early upward variance break for Canada. Thus, the usage of heteroskedasticity-robust tests is advisable.

Next, we run a few univariate tests to characterize the series. In particular, we test for stationarity using a KPSS test (with the null of no unit root) and the test by Phillips and Perron (1988) (with the null of a unit root). Note that heteroskedasticity is an issue for the KPSS test making critical values derived by Kwiatkowski et al. (1992) invalid. A possible remedy is to proceed as in Cavaliere (2005). We use the proposed bootstrap for the series y_t and e_t instead for residuals to test if they have no unit root.

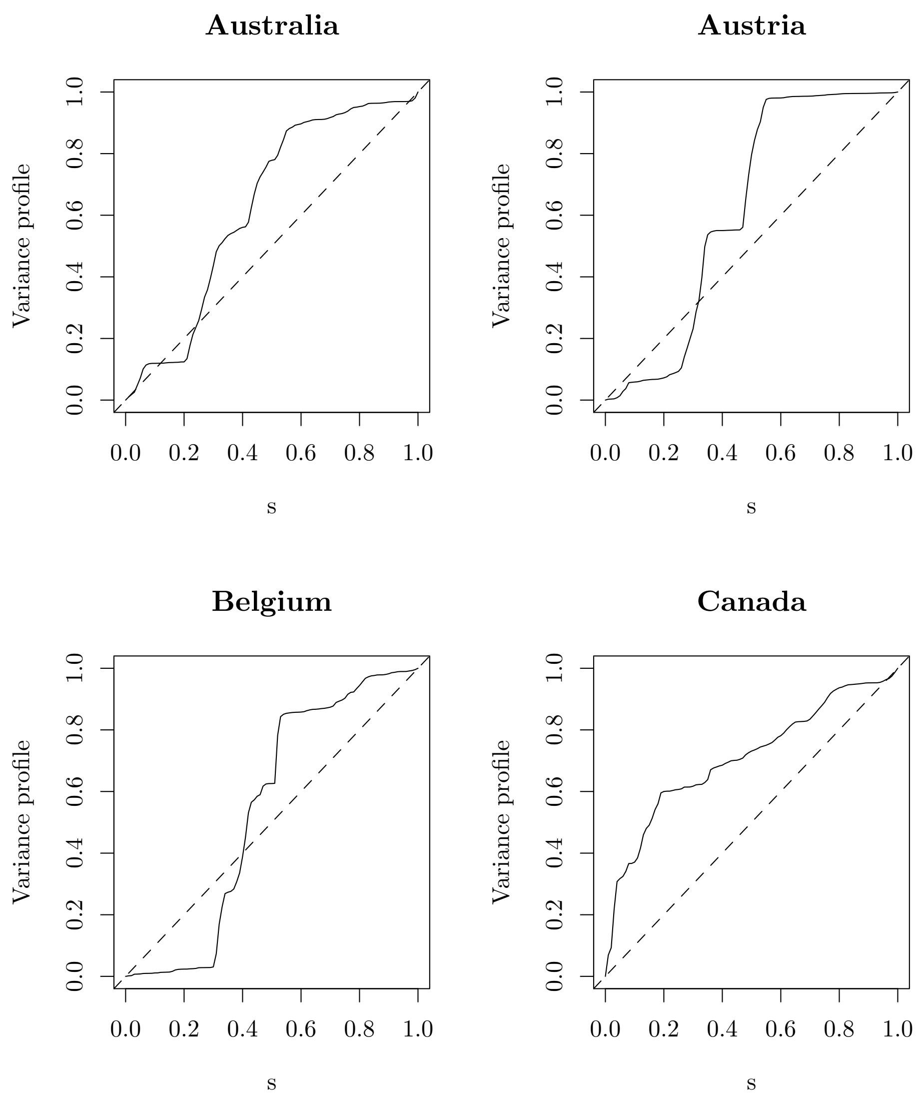
We perform three tests for cointegration, the bootstrap test using NLS residuals, the bootstrap test using leads-and-lags (LL) residuals (see Appendix B) and the subresidual based test. We use a non-parametric autocorrelation-robust estimator for the variance with a Bartlett kernel and a spectral window of $\lfloor 4(T/100)^{0.25} \rfloor$ as suggested in Kwiatkowski et al. (1992).

Table 5 reports the test results for the different countries given in the first column. The second to fourth column are for the cointegration tests with NLS, LL and the subresidual-based test. Columns 5 and 6 give results for the KPSS test for e_t and y_t, and column 7 and 8 for the Phillips-Perron (PP) test, resp. All test results are given by the corresponding p-values where very small p-values are abbreviated with $< .01$.

For the common level of significance of 5% we draw the following conclusions. In almost all cases the KPSS test leads to a rejection of the null of no unit root of both e_t and y_t while the PP test does not reject the null of a unit root. This provides evidence that the regressor and the regressand are both $I(1)$.

The three cointegration tests reveal mixed results. The first observation is that all three lead to acceptance of the null in the majority of the cases. We recall that the subresidual-based test is both in general undersized and second not robust to variance breaks, making it unreliable. Of course, bootstrap tests are dependent on simulation. Moreover, the p-values are all close to the nominal size, so that decisions may hinge on simulation variability. To reduce the effects of randomness we increased the number of bootstrap runs to 2,000. The bootstrap tests come to different test results in the case of Canada, Germany, Japan and Switzerland. Both tests reject only for Australia, New Zealand, Portugal and the United States. In the other cases both tests accept the null, providing

Figure 1: Empirical variance profile (11) for different countries. The dashed line is the reference line for homoskedasticity.



some support for the EKC hypothesis. Wagner (2015) rejected the null for the majority of countries using fully modified OLS for cointegrating polynomial regressions. However, tests which are not robust to variance breaks can lead to over-rejections.[3]

Table 5: p-values for different tests. p^b_{NLS} gives the p-value for the bootstrap NLS-based test and p^b_{LL} for the bootstrap LL version, p_{CS} for the test by Choi and Saikkonen (2010), $p_{KPSS,y}$ for the KPSS test for the CO_2 emissions, $p_{KPSS,y}$ for the KPSS test for the GDP, $p_{PP,y}$ for the PP-test for the CO_2 emissions, $p_{PP,y}$ for the PP-test for the GDP.

Country	p^b_{NLS}	p^b_{LL}	p_{CS}	$p_{KPSS,e}$	$p_{KPSS,y}$	$p_{PP,e}$	$p_{PP,y}$
Australia	.035	.032	.024	$<.01$	$<.01$	.686	.399
Austria	.390	.366	.469	$<.01$	$<.01$	.044	.774
Belgium	.560	.474	.797	.010	$<.01$	.040	.952
Canada	.053	.043	.316	$<.01$	.020	.738	.023
Denmark	.097	.142	.659	$<.01$	$<.01$	$>.99$	.747
Finland	.251	.187	.757	.029	$<.01$	.034	.739
France	.186	.163	.674	$<.01$	$<.01$	.607	.708
Germany	.027	.068	.124	$<.01$	$<.01$	.065	.582
Italy	.134	.143	.584	.045	$<.01$	.152	.900
Japan	.061	.019	.611	$<.01$	$<.01$	.130	.794
Netherlands	.329	.276	.719	.085	$<.01$	.017	.814
New Zealand	.024	.027	.032	.029	$<.01$	.094	.233
Norway	.090	.103	.695	.018	$<.01$	.073	.959
Portugal	.016	.026	.051	$<.01$	$<.01$	$<.01$	.951
Spain	.191	.113	.264	$<.01$	$<.01$	.430	.970
Sweden	.538	.432	.900	$<.01$	$<.01$	.478	.735
Switzerland	.049	.053	.063	.016	$<.01$	.239	.935
United Kingdom	.174	.111	.427	$<.01$	$<.01$	.015	.731
United States	.042	.037	.114	$<.01$	$<.01$	.830	.071

Acknowledgements

Financial support of the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) via the Collaborative Research Center "Statistical modelling of nonlinear dynamic processes" (SFB 823, Teilprojekt A4) is gratefully acknowledged.

The authors thank Janine Langerbein for excellent research assistance.

[3] The results are, in any case, not directly comparable since the Maddison database had a major update since then and also, since polynomials are sensitive to even small changes in the scala.

References

Andrews, D. W. (1991), 'Heteroskedasticity and autocorrelation consistent covariance matrix estimation', *Econometrica* **59**(3), 817–858.

Cavaliere, G. (2005), 'Unit root tests under time-varying variances', *Econometric Reviews* **23**(3), 259–292.

Cavaliere, G., Rahbek, A. and Taylor, A. R. (2010), 'Testing for co-integration in vector autoregressions with non-stationary volatility', *Journal of Econometrics* **158**(1), 7–24.

Cavaliere, G. and Taylor, A. (2006), 'Testing the null of co-integration in the presence of variance breaks', *Journal of Time Series Analysis* **27**(4), 613–636.

Cavaliere, G. and Taylor, A. (2008), 'Time-transformed unit root tests for models with non-stationary volatility', *Journal of Time Series Analysis* **29**(2), 300–330.

Choi, I. and Saikkonen, P. (2010), 'Tests for nonlinear cointegration', *Econometric Theory* **26**(3), 682–709.

Demetrescu, M., Hanck, C. and Kruse, R. (2019), Robust fixed-b inference in the presence of time-varying volatility, Technical report.

Dickey, D. A. and Fuller, W. A. (1979), 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American statistical association* **74**(366a), 427–431.

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

Giné, E. and Zinn, J. (1990), 'Bootstrapping general empirical measures', *The Annals of Probability* **18**(2), 851–869.

Grossman, G. M. and Krueger, A. B. (1995), 'Economic growth and the environment', *The Quarterly Journal of Economics* **110**(2), 353–377.

Hansen, B. E. (1996), 'Inference when a nuisance parameter is not identified under the null hypothesis', *Econometrica* **64**(2), 413–430.

Hansen, B. E. (2000), 'Testing for structural change in conditional models', *Journal of Econometrics* **97**(1), 93–115.

Kiefer, N. M. and Vogelsang, T. J. (2005), 'A new asymptotic theory for heteroskedasticity-autocorrelation robust tests', *Econometric Theory* **21**(6), 1130–1164.

Kuznets, S. (1955), 'Economic growth and income inequality', *The American Economic Review* **45**(1), 1–28.

Kwiatkowski, D., Phillips, P. C., Schmidt, P. and Shin, Y. (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?', *Journal of Econometrics* **54**(1-3), 159–178.

Lahiri, S. N. (1999), 'Theoretical comparisons of block bootstrap methods', *Annals of Statistics* **27**(1), 386–404.

Leybourne, S. J. and McCabe, B. (1994), 'A simple test for cointegration', *Oxford Bulletin of Economics and Statistics* **56**(1), 97–103.

McCabe, B., Leybourne, S. and Shin, Y. (1997), 'A parametric approach to testing the null of cointegration', *Journal of Time Series Analysis* **18**(4), 395–413.

Phillips, P. C. B. and Hansen, B. E. (1990), 'Statistical inference in instrumental variables regression with I(1) processes', *The Review of Economic Studies* **57**(1), 99–125.

Phillips, P. C. B. and Perron, P. (1988), 'Testing for a unit root in time series regression', *Biometrika* **75**(2), 335–346.

Pötscher, B. M. and Prucha, I. R. (2013), *Dynamic nonlinear econometric models: Asymptotic theory*, Springer Berlin Heidelberg.

Romano, J. P. and Wolf, M. (2001), 'Subsampling intervals in autoregressive models with linear time trend', *Econometrica* **69**(5), 1283–1314.

Saikkonen, P. (1991), 'Asymptotically efficient estimation of cointegration regressions', *Econometric Theory* **7**(1), 1–21.

Saikkonen, P. and Choi, I. (2004), 'Cointegrating smooth transition regressions', *Econometric Theory* **20**(2), 301–340.

Shin, Y. (1994), 'A residual-based test of the null of cointegration against the alternative of no cointegration', *Econometric Theory* **10**(1), 91–115.

Stern, D. I. (2004), 'The rise and fall of the environmental Kuznets curve', *World Development* **32**(8), 1419–1439.

Stern, D. I. (2018), The environmental kuznets curve, *in* 'Companion to Environmental Studies', Routledge in association with GSE Research, pp. 49–54.

Stypka, O., Wagner, M., Grabarczyk, P. and Kawka, R. (2017), 'The asymptotic validity of" standard" fully modified OLS estimation and inference in cointegrating polynomial regressions'. Working Paper.

Wagner, M. (2015), 'The environmental Kuznets curve, cointegration and nonlinearity', *Journal of Applied Econometrics* **30**(6), 948–967.

Wagner, M. and Hong, S. H. (2016), 'Cointegrating polynomial regressions: fully modified OLS estimation and inference', *Econometric Theory* **32**(5), 1289–1315.

A Appendix: Proofs

Proof of Theorem 1. Consider $T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\hat{u}_t$. Since $\hat{u}_t = u_t - (g(x_{tT},\hat{\theta}_T) - g(x_{tT},\theta_0))$, a second-order Taylor expansion of $g(x_{tT},\hat{\theta}_T)$ around θ_0 gives

$$\begin{aligned} T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\hat{u}_t =& T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}u_t - T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}K(x_{tT},\theta_0)'(\hat{\theta}_T-\theta_0) \\ &+ T^{1/2}(\hat{\theta}_T-\theta_0)'\left(T^{-1}\sum_{t=1}^{\lfloor Ts\rfloor}\frac{\partial^2 g(x_{tT},\tilde{\theta})}{\partial\theta\partial\theta'}\right)(\hat{\theta}_T-\theta_0), \end{aligned} \tag{12}$$

where $||\tilde{\theta}-\theta_0|| \leq ||\hat{\theta}_T-\theta_0||$.

For the first term in (12) Lemma 1 gives that, under H_0,

$$T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor} u_t = T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor} \zeta_{u,t} \stackrel{w}{\to} B_{0,\Omega}(s).$$

For the second term in (12), recall that $T^{1/2}(\hat\theta_T - \theta_0) \stackrel{w}{\to} \psi\left(B^0_{1,\Omega}, \theta_0, \kappa\right)$ (Proposition 1). By Lemma 1,

$$x_{tT} = (T_0/T)^{1/2} x_t = (T_0/T)^{1/2}\sum_{j=1}^{\lfloor Ts\rfloor} \zeta_{1,j} \stackrel{w}{\to} T_0 B_{1,\Omega}(s) =: B^0_{1,\Omega}(s).$$

This implies that

$$T^{-1}\sum_{t=1}^{\lfloor Ts\rfloor} x_{tT} \stackrel{w}{\to} \int_0^s B^0_{1,\Omega}(r)\mathrm{d}r,$$

and by the continuous mapping theorem,

$$T^{-1}\sum_{t=1}^{\lfloor Ts\rfloor} K(x_{tT}, \theta_0) \stackrel{w}{\to} \int_0^s K(B^0_{1,\Omega}(r), \theta_0)\mathrm{d}r =: F(s, B^0_{1,\Omega}, \theta_0).$$

We conclude that

$$T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor} \hat u_t \stackrel{w}{\to} B_{0,\Omega}(s) - F(s, B^0_{1,\Omega}, \theta_0)'\psi\left(B^0_{1,\Omega}, \theta_0, \kappa\right),$$

since all weak convergences hold jointly. Another application of the continuous mapping theorem yields

$$T^{-2}\sum_{t=1}^{T}\left(\sum_{j=1}^{t} \hat u_j\right)^2 \stackrel{w}{\to} \int_0^1 \left(B_{0,\Omega}(s) - F(s, B^0_{1,\Omega}, \theta_0)'\psi(B^0_{1,\Omega}, \theta_0, \kappa)\right)^2 \mathrm{d}s.$$

Finally, (4) follows by the continuous mapping theorem. ■

Proof of Theorem 2. Under the alternative $H_1 : \rho^2_\mu > 0$

$$T^{-1/2}u_{\lfloor Ts\rfloor} = T^{-1/2}\zeta_{0,\lfloor Ts\rfloor} + T^{-1/2}\rho_\mu\sum_{t=1}^{\lfloor Ts\rfloor} \zeta_{\mu,t} \stackrel{w}{\to} \rho_\mu B_{2,\Omega}(s).$$

This implies that $T^{-3/2}\sum_{t=1}^{\lfloor Ts\rfloor} u_t \stackrel{w}{\to} \rho_\mu\int_0^s B_{2,\Omega}(r)\mathrm{d}r$ and hence $T^{-3/2}\sum_{t=1}^{\lfloor Ts\rfloor} u_t = O_p(1)$.

Like in the proof of Theorem 1 we use a Taylor expansion to obtain

$$T^{-3/2}\sum_{t=1}^{\lfloor Ts\rfloor}\hat{u}_t = T^{-3/2}\sum_{t=1}^{\lfloor Ts\rfloor}u_t - T^{-3/2}\sum_{t=1}^{\lfloor Ts\rfloor}K(x_{tT},\theta_0)'(\hat{\theta}_T-\theta_0)+o_p(1).$$

Next, observe $|\hat{\theta}_T-\theta_0| = O_p(T^{1/2})$. To see this we use a linear approximation

$$g(x_{tT},\hat{\theta}_T)\approx g(x_{tT},\theta_0)+K_t(\hat{\theta}_T-\theta_0),$$

where K is the Jacobian matrix with entries $K_{ti}=\frac{\partial g(x_{tT},\theta)}{\partial\theta_i}$, for $t=1,\ldots,T$, $i=1,\ldots,k$, and K_t is its t-th row. We can use this approximation and the following normal equations of a linear model

$$(K'K)^{-1}(\hat{\theta}_T-\theta_0)=K'\tilde{y},$$

with $\tilde{y}_t = y_{tT}-g(x_{tT},\theta_0)$. We now obtain the asymptotics as for ordinary least squares as in Shin (1994) and McCabe et al. (1997) using that $\sum_{t=1}^{\lfloor Ts\rfloor} g(x_{tT},\theta_0)=O_p(T)$, $\sum_{t=1}^{\lfloor Ts\rfloor}K(x_{tT},\theta_0)=O_p(T)$, and $\sum_{t=1}^{\lfloor Ts\rfloor}u_t = O_p(T^{3/2})$.

Thus, $\sum_{t=1}^{\lfloor Ts\rfloor}\hat{u}_t = O_p(T^{3/2})$, which leads to

$$T^{-2}\sum_{t=1}^{T}\left(\sum_{j=1}^{t}\hat{u}_j\right)^2 = O_p(T^2).$$

Moreover, Kwiatkowski et al. (1992) showed that the long-run variance estimator $\hat{\omega}_u^2 = O(lT)$ which implies $\hat{\eta}=O_p(T/l)$. As long as $T/l\to\infty$ for $T\to\infty$ the test is consistent. ■

Proof of Theorem 3.

(i) Similarly to the proof of Theorem 3 in Cavaliere and Taylor (2006) consider the process M_t^b s.t.

$$M_T^b(s) := T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}u_t^b = T^{-1/2}\sum_{t=1}^{\lfloor Ts\rfloor}\hat{u}_t z_t.$$

Conditionally on $\{\hat{u}_t, x_{tT}\}_{t=1}^T$, this is an exact Gaussian process with kernel

$$\Lambda_T^M(s,s') = T^{-1}\sum_{t=1}^{\lfloor T(s\wedge s')\rfloor}\hat{u}_t^2,$$

where $s \wedge s'$ denotes the minimum of s and s'.

Under the null, $Var(u_t) = \sigma^2_{u,t}$ and $\sigma^2(s) = \sigma^2_{u,\lfloor Ts \rfloor}$ which is the variance profile of the u_t. As in the proof of Lemma A.5 in Cavaliere et al. (2010) we see that

$$T^{-1} \sum_{t=1}^{\lfloor T(s \wedge s') \rfloor} \hat{u}_t^2 = T^{-1} \sum_{t=1}^{\lfloor T(s \wedge s') \rfloor} u_t^2 + o_p(1) \xrightarrow{p} \int_0^{s \wedge s'} \sigma^2(r) \mathrm{d}r,$$

pointwise, where the first equality follows by McCabe et al. (1997). Since $T^{-1} \sum_{t=1}^{\lfloor Ts \rfloor} \hat{u}_t^2$ is monotonically increasing in s and the limit function is continuous in s the convergence in probability is also uniform. The RHS is the kernel of the Gaussian process W_σ s.t. $W_\sigma(s) := \int_0^s \sigma(r) \mathrm{d}W(r)$, where W is a standard Brownian motion. This implies that $M_T^b(s) \xrightarrow{w}_p W_\sigma(s)$, as in Hansen (1996).

Analogously, applying the same mappings as in the proof of Theorem 1,

$$T^{-2} \sum_{t=1}^{T} \left(\sum_{j=1}^{t} \hat{u}_j^b \right)^2 \xrightarrow{w}_p \int_0^1 \left(W_\sigma(s) - F(s, B^0_{1,\Omega}, \theta_0)' \psi(B^0_{1,\Omega}, \theta_0, \kappa) \right)^2 \mathrm{d}s.$$

Now, we derive the large sample behavior of $(\hat{\omega}_u^b)^2$.

$$\begin{aligned}
(\hat{\omega}_u^b)^2 &= T^{-1} \sum_{t=1}^{T} (\hat{u}_t^b)^2 + 2T^{-1} \sum_{s=1}^{l} w(s,l) \sum_{t=s+1}^{T} \hat{u}_t^b \hat{u}_{t-s}^b \\
&= T^{-1} \sum_{t=1}^{T} (u_t^b)^2 + 2T^{-1} \sum_{s=1}^{l} w(s,l) \sum_{t=s+1}^{T} u_t^b u_{t-s}^b + o_p(1) \\
&\xrightarrow{p} \int_0^1 \sigma^2(r) \mathrm{d}r,
\end{aligned}$$

because $E(z_t z_{t-s} | \{\hat{u}_t, x_{tT}\}_{t=1}^T) = 0$ for all $s > 0$ and $= 1$ for $s = 0$, and the same argument as above by McCabe et al. (1997).

This implies that the bootstrap test statistic $\hat{\eta}^b$ samples from a distribution that has the same variance profile as the distribution of $\hat{\eta}$ but with white noise serial correlation. Using the arguments in Demetrescu et al. (2019) which are based upon Kiefer and Vogelsang (2005) the bootstrap (asymptotically) controls size.

(ii) We again consider $M_T^b(s)$ and $\Lambda_T^M(s,s')$ but now it suffices to look at the order of convergence. Recall that under the alternative $\sum_{t=1}^{\lfloor Ts \rfloor} \hat{u}_t = O_p(T^{3/2})$ and $\sum_{t=1}^{\lfloor Ts \rfloor} \hat{u}_t^2 = O_p(T^2)$. This implies that $\Lambda_T^M(s,s') = O_p(T)$ and, like in part (i), $T^{-1/2} M_T^b(s)$ converges weakly in probability to a Gaussian process where the kernel is given by the weak limit of $T^{-1}\Lambda_T^M(s,s')$.

By the continuous mapping theorem it follows that $\sum_{t=1}^{T} \hat{u}_t^b = O_p(T)$ and, hence, that

$$\sum_{t=1}^{T} \left(\sum_{j=1}^{t} \hat{u}_j^b \right)^2 = O_p(T^3).$$

Consider next the long-run variance estimator $(\hat{\omega}_u^b)^2$. Again, as in the proof of Theorem 2, $(\hat{\omega}_u^b)^2$ is consistent under the alternative of order $O(lT)$. All in all, we get $\hat{\eta}^b = O_p(1/l)$. Since $\hat{\eta} = O_p(T/l)$ (Theorem 2) it follows that $p_T^b \xrightarrow{w} 0$, as long as $l \to \infty$ for $T \to \infty$.

■

Proof of Corollary 1. As in Choi and Saikkonen (2010) we modify equation (12) to

$$\ell^{-1/2} \sum_{t=i}^{\lfloor \ell s+i-1 \rfloor} \hat{u}_t = \ell^{-1/2} \sum_{t=i}^{\lfloor \ell s+i-1 \rfloor} u_t - \ell^{-1} \sum_{t=i}^{\lfloor \ell s+i-1 \rfloor} K(x_{tT},\theta_0)' \sqrt{T}(\hat{\theta}_T - \theta_0) \sqrt{\frac{\ell}{T}}$$
$$+ T(\hat{\theta}_T - \theta_0)' \left(\ell^{-3/2} \sum_{t=i}^{\lfloor \ell s+i-1 \rfloor} \frac{\partial^2 g(x_{tT},\tilde{\theta})}{\partial\theta\partial\theta'} \right) (\hat{\theta}_T - \theta_0) \frac{\ell}{T}.$$

We use the arguments from the proof of Theorem 1 and $\frac{\ell}{T} \to 0$ to see

$$\ell^{-1/2} \sum_{t=i}^{\lfloor \ell s+i-1 \rfloor} \hat{u}_t \xrightarrow{w} B_{0,\Omega}(s).$$

The remainder follows by the continuous mapping theorem and $(\hat{\omega}_u^\ell)^2 \to \bar{\omega}_u$. ■

B Appendix: Additional Simulations

This section discusses the case of estimating polynomial regressions with additional deterministic regressors. This is beyond our model assumptions, following the assumption of Choi and Saikkonen (2010) that all regressors are integrated. Deterministic regressors are not integrated. However, deterministic regressors are useful in many applications. Therefore, we extend the simulations of Section 4.2 to study the impact of an intercept or a time trend to the rejection rates for the bootstrap test. More specifically, we discuss the cubic regression model with deterministic because it is the model in Section 5. Unreported results show that the results are qualitatively similar for a linear cointegrating regression model with a deterministic regressor.

First we consider the cubic model including an intercept

$$y_t = 1 + x_t + 2x_t^2 + x_t^3 + u_t.$$

Panel (a) of Table 6 shows, analogously to the previous results, the rejection frequencies with the bootstrap test using NLS. We observe that in the presence of endogeneity the test is somewhat oversized with a rejection rate of about 10%.

We also discuss a version of the cubic polynomial regression with a time trend of the form

$$y_t = 1 + t + x_t + 2x_t^2 + x_t^3 + u_t.$$

We do so mainly because there are some notable differences to the case without deterministic components, and because we use this model for the application in Section 5. Panel (a) of Table 7 shows, analogously to the previous results, the rejection frequencies with the bootstrap test using NLS. We observe that in the presence of endogeneity the test is oversized with a rejection rate of about 10%. This is no surprise as the literature already documented this issue and proposed several solutions. For example, one could use fully modified OLS developed in Phillips and Hansen (1990) as suggested in Wagner and Hong (2016). We here follow Choi and Saikkonen (2010) who use the leads-and-lags (LL) estimator proposed by Saikkonen (1991) (which is also known under the name dynamic (non)-linear least squares). We briefly describe the procedure. We estimate the coefficients in the model

$$y_t = c + \delta t + \theta_1 x_t + \theta_2 x_t^2 + \theta_3 x_t^3 + \sum_{j=-K}^{K} \pi_j \Delta x_{t-j} + e_t,$$

which means that we include $2K$ leads and lags into the regression. As in Choi and Saikkonen (2010) we take $K = 1, 2, 3$. However, panel (b) in Table 7 only reports the case of $K = 1$ as the others have shown similar results. We compute test statistics and bootstrap p-values analogously,

Table 6: The table reports the empirical rejection frequencies for testing the null of cointegration in the *cubic* regression model with *intercept* for various parameter constellations. All rejection rates are given as percentages. The nominal size is 5%. Panel (a) is for the bootstrap test using least squares and panel (b) is for the bootstrap test using leads and lags.

				σ_μ^2:	0				0.001				0.01				0.1			
				ρ:	0		0.5		0		0.5		0		0.5		0		0.5	
	T	τ:	σ^2	λ	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5
(a)	100	0			4,5	7,8	6,7	9,7	11,2	16,3	13,7	16,8	43	49,7	34,7	40,5	85,7	87,7	49,9	50,2
		0,1	0,0625		6,2	6,5	7	8,3	8,7	11,5	8,3	12,3	28,3	29,3	24,9	28,8	71,5	71,3	39,2	43
			16		5,8	8	5,9	10,9	12,3	17	14,5	16,6	45,8	48,2	38,1	35,3	87,5	88,8	51,6	52
		0,5	0,0625		5,2	6,5	4,3	6	9,7	10	6,9	11,5	30,5	35,9	21,9	25,6	76,4	77,9	31,6	36,5
			16		4,8	7,9	6,4	7	9,2	10,8	8,3	8,5	34,9	35,4	24,1	26,1	80,5	81,1	39,8	37,3
		0,9	0,0625		5,6	6,9	6,8	8,1	12,3	14,2	13,2	14,9	47,9	49,1	38,2	37,8	88	89,7	52,5	52,9
			16		4,4	5,1	4,4	7,4	7,3	9,4	8,6	11,1	27,9	29,2	24,9	25,7	70,5	72,7	38,2	41,7
	300	0			5,7	8,3	6,8	11,3	44,9	48,7	43,2	47,2	90,7	89,4	78,1	73,2	99,4	99,8	81	80,1
		0,1	0,0625		5,9	8	7,6	9,4	22,5	25,1	25,8	26,1	73,6	68,9	59,2	54,7	96,5	97,8	60,4	65,9
			16		5,1	8,7	7	11,5	45,9	49,8	46	48,9	90,1	92,4	76,1	78,5	99,7	99,9	81,9	79,8
		0,5	0,0625		6,7	8,5	7,6	9,5	29,4	35,4	29,9	34,3	81,1	81,4	64,4	65	99	98,4	71	69,9
			16		5,3	8,8	6,8	11,4	36	36,9	34,5	35,8	82,4	82,3	65,4	65,6	99,2	98,7	70,6	69,4
		0,9	0,0625		4,4	8,5	7,1	11,7	43,3	49,3	44,3	47,7	89,6	90,7	76,4	75,7	99,3	99,8	78,3	80,8
			16		5,9	6,2	6,7	10	25,6	26,2	28,4	27,8	71,2	71,9	55,5	57,5	96	97,2	64,4	66,3
(b)	100	0			5,2	5,4	9,9	8,5	12,2	14,8	16,1	19,7	41	47,7	41,4	44,9	87,6	90,3	63,9	63,6
		0,1	0,0625		9,2	8	12	11,7	10,5	13,3	14	17,8	28,5	33,1	32,3	35	74,4	77,2	56,5	51,7
			16		4,6	4,7	8,2	10,7	12,2	13,9	16,1	18	44,8	51,6	41,4	47,2	86,9	90,5	62,4	62,1
		0,5	0,0625		6,2	4	7,6	8,4	9,1	10,4	13,2	14,1	30,3	37,8	27,9	31,8	76,4	78,1	51,4	49,6
			16		4,8	5,8	8,1	7,2	10,3	9,7	12,3	11,9	34,5	37,4	30	33,6	79,5	83,8	51,6	52,1
		0,9	0,0625		6	5,1	8,7	12,3	12,5	15,5	17,2	21,2	47	50,2	44,5	45,9	88,3	90,2	64,8	65,7
			16		7,7	6,6	9	11,5	9,2	11,4	13,4	15,9	30,9	34,5	33,9	34,8	71,3	79,5	53,4	55,9
	300	0			5,2	4,7	8,6	8,8	44,6	49,3	44,6	48,9	89,1	92,6	74,6	77,6	99,6	99,8	80,7	79,4
		0,1	0,0625		5	5,7	9,3	9,5	25,9	29,8	27,9	31,9	72	78,2	58,2	57,6	97,7	97,6	67,1	66,2
			16		4,8	4,5	7,6	10,5	44,8	50,1	44,6	47,1	88,8	92,4	75,5	76,9	99,6	100	82,7	82,5
		0,5	0,0625		4,8	6,1	6,5	7,7	30,1	37,6	33,7	38,1	77,9	84,9	63,6	68	99,1	99,1	71,3	73,3
			16		5,2	5,7	5,9	8,3	33,1	38,1	32,5	36	79,7	85,4	64,1	66,2	98,7	99,2	72,1	73,8
		0,9	0,0625		4,9	4,7	6,8	8,2	42,4	51,8	43,3	47,9	90,5	91,3	77,7	77	99,8	99,6	82,1	82,8
			16		6,2	5,1	8	8,1	22,7	26,8	26,4	30,7	71,6	79,4	56,5	62,3	96,9	98,2	67	66,6
(c)	100	0			0	0,1	1,3	1,6	0,5	1	3,1	4,1	10,3	13,6	12,2	16,6	47	45,1	32	28,9
		0,1	0,0625		1,1	0,7	3,1	2,3	1,7	1,5	4,4	5,5	7,3	8,1	13,3	14,3	39,5	39,3	30,9	30,2
			16		0,2	0,3	0,9	1,3	0,6	0,6	2,5	2,9	15,1	13,5	16,3	16,2	45,8	48,4	32,7	30,7
		0,5	0,0625		2,3	2,9	5,5	7	2,5	4,3	6,8	8,6	12,4	14,7	18,8	20	42,8	42,5	31,6	33,3
			16		5,1	6	5,4	6,6	8	10	10,1	10,9	23,7	23,7	24,5	21,1	50,2	49,7	35	32,2
		0,9	0,0625		0	0,1	1,3	2,5	0,7	1	2,9	4,1	12,6	11,1	15,1	14,2	47,3	46,9	31,9	30,8
			16		2,4	3,3	3,5	5,5	3,3	4,1	5,5	7,3	16,3	16,9	18,1	15,9	49,9	48,8	33,4	31,8
	300	0			0,1	0,2	0,7	0,7	12,6	12,7	14,6	12,8	49,8	54,1	33,8	35,3	87	84,9	45,4	44,7
		0,1	0,0625		1,5	0,8	3	2,3	8,2	7,1	10,5	10,1	42,2	47,3	30,7	32,4	86,2	83,4	46,1	42,3
			16		0	0	0,2	0,7	15,1	14	14,7	14,6	52,7	54,5	35,4	36,4	87,8	87,7	45,1	44,1
		0,5	0,0625		3,1	2,8	7	6,6	20,4	19,4	24,7	21,7	51,3	53,7	39,3	42	85,2	85,4	47,7	46,8
			16		7,1	13,4	8,9	13,3	30,1	30,9	30,4	29,6	63,8	66,3	47	46,8	90,6	88,6	50,3	51,3
		0,9	0,0625		0	0,1	0,5	0,8	10,8	12	13,1	14,3	52	54,2	37,2	35,3	86,9	87,3	43,4	47,2
			16		3,1	4	2,9	3,6	17,8	19,8	18,4	18,4	55,7	54,1	41,4	37,1	85,8	86,7	49,8	46,5

now using residuals $\hat{e}_t$. To save computational time we run 1,000 replications in this example for all settings.

Comparing both panels of Table 7 shows that the size problem is corrected. Moreover, the empirical power is of comparable magnitude. We also employ this test based on LL for the application in Section 5.

C Appendix: Plots

Table 7: The table reports the empirical rejection frequencies for testing the null of cointegration in the *cubic* regression model with *time trend* for various parameter constellations. All rejection rates are given as percentages. The nominal size is 5%. Panel (a) is for the bootstrap test using least squares and panel (b) is for the bootstrap test using leads and lags.

				σ_μ^2:	0				0.001				0.01				0.1			
				ρ:	0		0.5		0		0.5		0		0.5		0		0.5	
	T	τ:	σ^2	λ	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5	0	0.5
(a)	100	0			4.9	10.7	3.6	6.3	6.4	14.1	5.2	7.3	24	33.2	12.8	15	78.6	82.4	22.6	25.6
		0.1	1/16		6.9	10.1	4.5	5.2	9.5	11.7	6.3	5.4	15.5	18.8	9	8.6	60	61.7	20.1	19.4
			16		4.8	9.1	3.4	5.6	6.8	13.4	3.8	6.4	24.9	34.1	11.9	16.4	80.2	81.9	22	25
		0.5	1/16		5	11.9	2.1	3.8	7.8	9.8	2.1	4	21.3	27.2	8.3	8.7	69.3	72.2	20.3	20.4
			16		4.3	8.9	3.1	3.1	6.9	10.1	4	3.4	22.3	23.8	8.4	11.6	72.6	73.4	23.6	21.5
		0.9	1/16		6	9.2	3.4	5.1	7.3	11	4.7	4.5	24.5	30	13	14.1	79.9	83.5	24.6	25.6
			16		6.8	7.7	4.1	3.7	6.6	9.4	2.9	5.3	15.3	19	7.5	8.9	61.7	62.5	17.5	16.3
	300	0			4.9	9.2	8.6	14.6	25.2	29.8	25.8	32.7	84	85.1	67.4	67.4	99.9	99.9	75.3	76.6
		0.1	1/16		6.4	8.2	6.9	10.6	12.4	16.9	14.3	19.2	60.4	59.2	47.8	42.6	98	98.3	58.2	62
			16		4.9	10	8.2	14.9	23.8	32.9	25.9	34.7	84.1	85.6	68.6	69.8	99.9	100	76.9	74.2
		0.5	1/16		4.8	9.4	8	12	18.1	28.8	20	29.5	74.4	73.4	58	59	99	99.4	70	68.8
			16		5.1	9.3	7.2	13.7	20.8	26.5	21.5	26.3	76	76.1	60.1	58.3	99.3	99.4	71.7	69.1
		0.9	1/16		5.5	10.8	7.3	16.1	25.4	29.7	27.7	32	82.2	85.8	66.2	68.5	99.7	99.7	78.1	75.4
			16		5.4	8.8	7.5	10.8	14.3	18.7	16.2	20.4	59.6	62.3	45.4	45.7	97.9	98.1	61.6	63.9
(b)	100	0			6.2	4.9	6.4	9.5	6.2	9.5	9.7	10.5	24.6	30.7	24.8	26	78.4	85.2	44.7	48.5
		0.1	1/16		9.5	7.5	11.3	9.9	10.5	10.2	12.3	12.6	19.3	19.6	19.1	18.4	63.1	70.4	36.1	44.8
			16		6.3	4.9	6.9	10.1	7.9	8.7	7.4	10.3	28.3	31	22.7	22.5	80.7	85.7	43.8	45.5
		0.5	1/16		6.2	5.5	5.6	5.5	6.8	8	7.8	8.4	22.7	26.5	17.6	20.5	71.4	73	41.2	37.5
			16		6.3	4.5	5.5	6.7	6.5	8.7	7.2	9.1	23.5	26.2	17.9	19.9	72.7	79.5	38.7	40.7
		0.9	1/16		5.2	5.2	6.5	8.8	7.4	8	8.6	11.8	24.7	29.4	22.4	24	78.6	82.4	44.7	43.2
			16		7.6	8.1	7.2	9.2	9.8	9.5	11.6	11.7	18	21.4	16.3	21	61.3	70.1	35	38.9
	300	0			6.5	4.2	8.6	10.5	22.1	30.6	28.4	31.1	82.2	87	70.7	69.9	99.8	99.8	78.7	77.1
		0.1	1/16		6.3	6.1	7.7	10.4	14.2	17.2	15.1	21.4	59.5	68.9	47.6	52.7	98.2	98.7	65.4	67.1
			16		5.8	5.1	8	11.9	25.8	31.3	27.7	31.7	80.6	89	68.9	70.7	99.8	100	79.8	79.3
		0.5	1/16		5.5	5	8.6	9.3	18.7	28.3	23.5	29	71.4	81.3	59.8	62.7	99.4	99.6	72	69.5
			16		5.9	5.6	7.7	9.2	21.1	26.2	23.9	30.2	74.4	80.2	58.2	62	99.1	99.4	69.1	71.7
		0.9	1/16		5.5	5.9	9	11.8	24.8	29.4	28	31.8	84	86.7	67.5	68.6	99.8	99.9	78.6	79.1
			16		6.5	5.2	7.9	9.1	12.4	17	17.1	21.3	59.1	71.5	46.8	53.2	97.6	98.9	65.1	64.1
(c)	100	0			0	0	0	0	0	0.1	0.1	0	0.4	0.6	0.1	0.1	10.7	10.9	0.9	0.5
		0.1	1/16		0.2	0.3	0	0.3	0.2	0.4	0.3	0.1	1.2	1.6	0	0.4	14.5	14.6	0.7	1.1
			16		0	0	0	0.1	0	0	0	0.1	0.5	0.6	0.2	0.1	9.6	10.5	0.6	1
		0.5	1/16		0.5	0.9	1.2	0.9	0.8	1.3	0.8	0.9	4	5.2	2	2.9	26.2	25.9	6.1	6.7
			16		0.5	0.9	0	0.4	0.7	1	0.1	0.4	2.9	3.3	0.2	0.3	16	16.5	1.1	1.2
		0.9	1/16		0	0	0.1	0.2	0	0	0.1	0	0.5	0.5	0.1	0.2	10.5	11.4	1.2	0.9
			16		0.1	0.1	0	0.1	0.1	0.1	0.1	0.1	0.3	0.4	0.1	0.2	5.1	5.4	0.5	0.4
	300	0			0	0	0.1	0.1	0.3	0.5	0.2	0.1	13.1	13.8	1.8	3.8	58	57.9	9.1	6.9
		0.1	1/16		0.3	0.4	0.1	0.3	1.6	1.3	0.5	0.6	16.7	17.5	4.1	4.7	57.6	58.2	10.1	9.7
			16		0	0	0	0	0.4	0.3	0.1	0.2	13.3	13.7	3.1	3.9	59	58.6	7.9	8.3
		0.5	1/16		0.6	0.7	1.5	1.9	3.5	4.1	5.1	5.5	27.1	28.5	13.7	15.8	69.7	69.2	21.6	23.1
			16		1.6	3.2	1.5	2.1	9.4	10.5	3.3	4	38.6	39.8	7.4	7.9	77	76.7	9.6	9.2
		0.9	1/16		0	0	0.1	0	0.4	0.5	0.3	0.3	13	13.4	3.4	3.7	56	57	9.2	10.8
			16		0.2	0.2	0	0.1	0.5	0.7	0.2	0.2	9.9	10.9	1.7	1.9	56.6	57.4	5.6	5

Figure 2: Empirical variance profile (11) for different countries. The dashed line is the reference line for homoskedasticity.

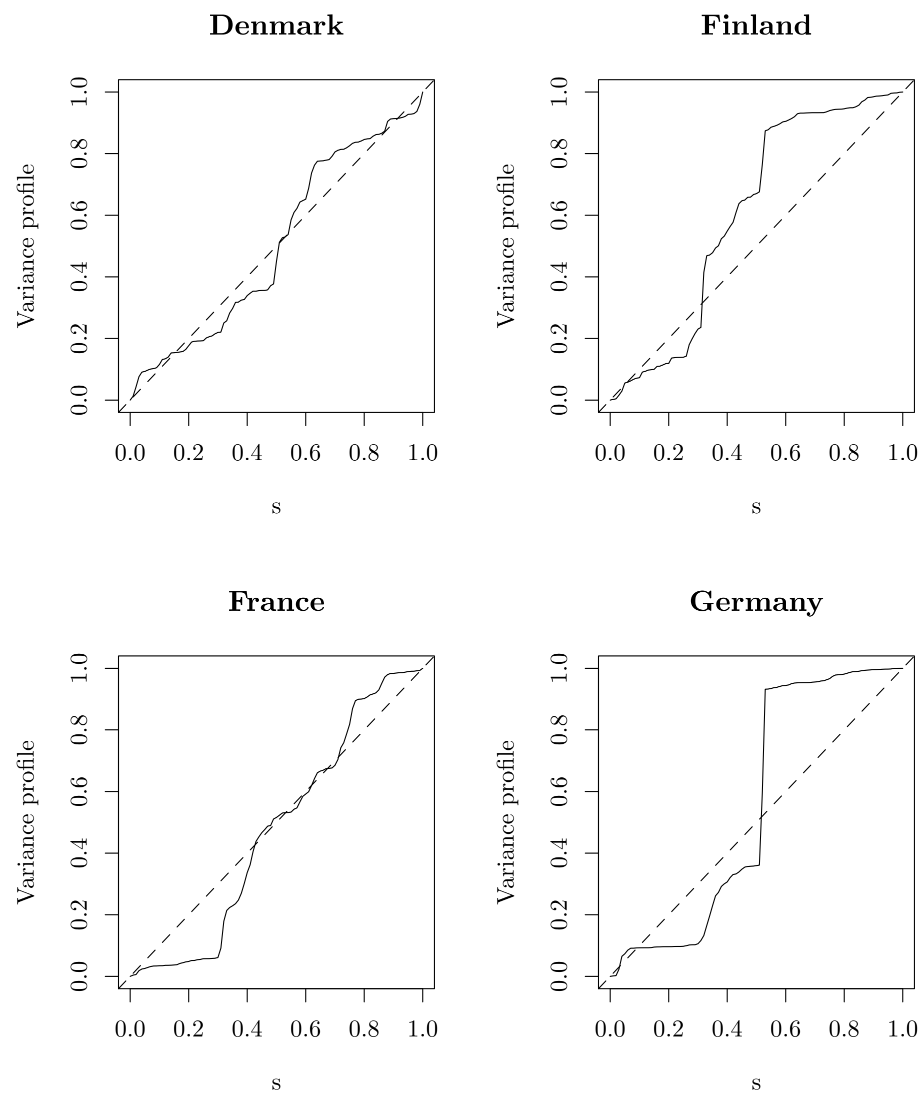


Figure 3: Residuals vs. fitted values for Austria, Belgium, Norway, USA to inspect heteroskedasticity. The red solid lines are LOWESS curves.

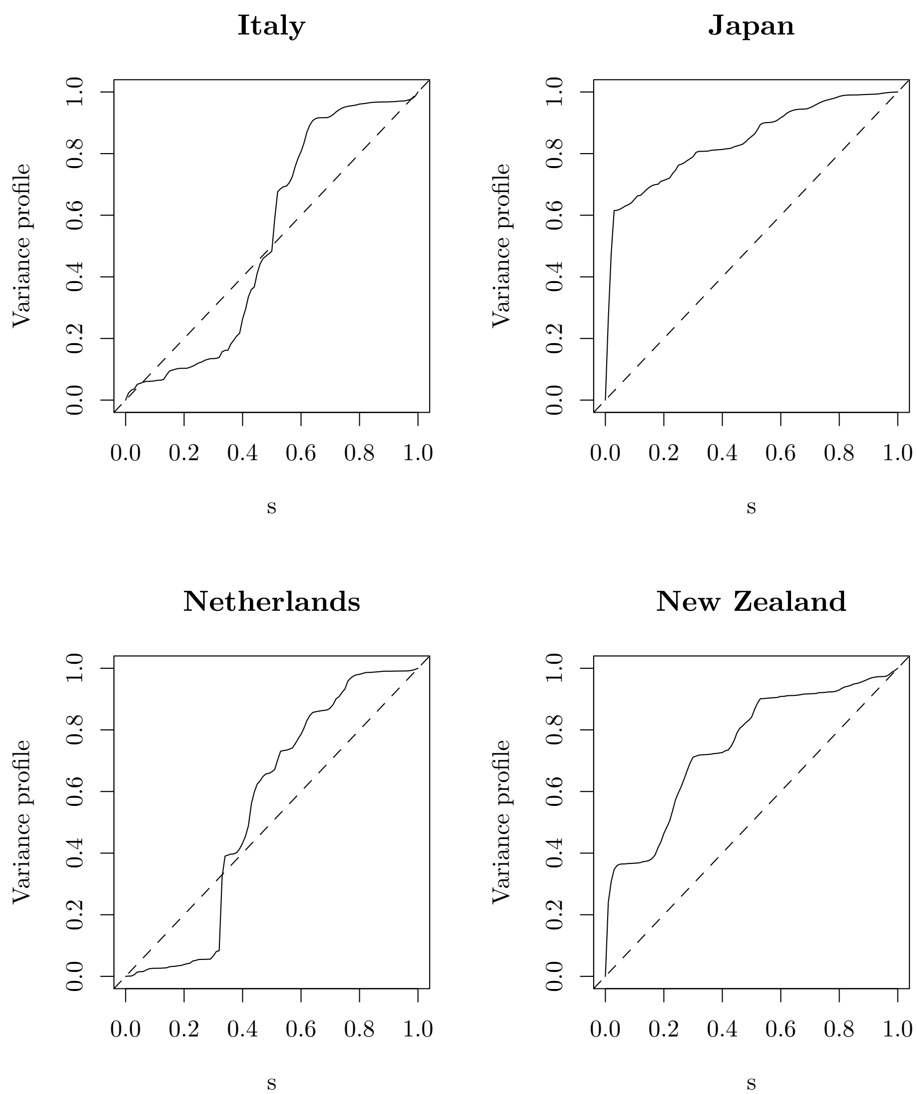


Figure 4: Empirical variance profile (11) for different countries. The dashed line is the reference line for homoskedasticity.

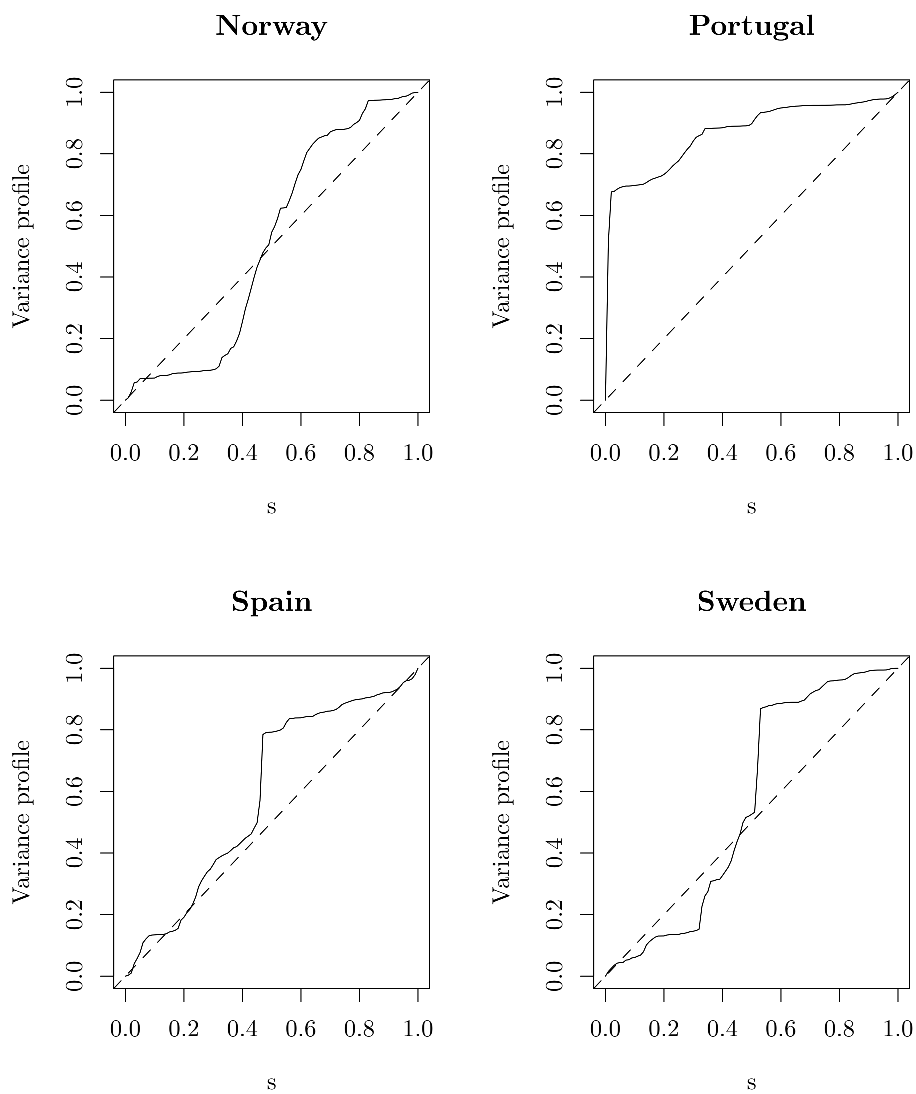


Figure 5: Empirical variance profile (11) for different countries. The dashed line is the reference line for homoskedasticity.