IHS Economics Series
Working Paper 224
July 2008

Nonlinear Cointegration Analysis and the Environmental Kuznets Curve

Seung Hyun Hong
Martin Wagner

Institut für Höhere Studien
Institute for Advanced Studies
Vienna

224

Reihe Ökonomie

Economics Series

Nonlinear Cointegration Analysis and the Environmental Kuznets Curve

Seung Hyun Hong, Martin Wagner

224

Reihe Ökonomie

Economics Series

Nonlinear Cointegration Analysis and the Environmental Kuznets Curve

Seung Hyun Hong, Martin Wagner

July 2008

Institut für Höhere Studien (IHS), Wien
Institute for Advanced Studies, Vienna

Contact:

Seung Hyun Hong
Department of Economics
Concordia University
1455 de Maisonneuve Blvd. West
Montreal, Quebec H3G 1M8
email: shhong@alcor.concordia.ca

Martin Wagner
Department of Economics and Finance
Institute for Advanced Studies
Stumpergasse 56
1060 Vienna, Austria
☎: +43/1/599 91-150
email: Martin.Wagner@ihs.ac.at

Abstract

Recent years have seen a growing literature on the environmental Kuznets curve (EKC) that resorts in a large part to cointegration techniques. The EKC literature has failed to acknowledge that such regressions involve unit root nonstationary regressors and their integer powers (e.g. GDP and GDP squared), which behave differently from linear cointegrating regressions. Here we provide the necessary tools for EKC analysis by deriving estimation and testing theory for cointegrating equations including stationary regressors, deterministic regressors, unit root nonstationary regressors and their integer powers. We consider fully modified OLS estimation, specification tests based on augmented and auxiliary regressions, as well as a sub-sample KPSS type cointegration test. We present simulation results illustrating the performance of the estimators and tests. In the empirical application for CO_2 and SO_2 emissions for 19 early industrialized countries over the period 1870-2000 we find evidence for an EKC in roughly half of the countries.

Keywords

Integrated Process, Nonlinear Transformation, Fully Modified Estimation, Nonlinear Cointegration Analysis, Environmental Kuznets Curve

JEL Classification

C12, C13, Q20

Contents

1 Introduction

Since the seminal work of Grossmann and Krueger (1993, 1995) many econometric studies of the relationship between measures of economic development (typically proxied by per capita GDP) and pollution, respectively emissions, have been conducted. Survey articles like Stern (2004) or Yandle, Bjattarai, and Vijayaraghavan (2004) count more than one-hundred refereed publications. Most of the papers focus on a specific conjecture, the so called 'environmental Kuznets curve' (EKC) hypothesis, which postulates an inverted U-shaped relationship between the level of economic development and the degree of income inequality. The term EKC refers by analogy to the inverted U-shaped relationship between the level of economic development and the degree of income inequality postulated by Kuznets (1955) in his 1954 presidential address to the American Economic Association.

The largest part of the empirical EKC literature estimates parametric EKCs, however, also other estimation strategies have been followed in the empirical EKC literature: non-parametric EKCs (see e.g. Millimet, List, and Stengos, 2003), semi-parametric EKCs (see e.g. Bertinelli and Strobl, 2005) or EKCs using spline interpolations (see e.g. Schmalensee, Stoker, and Judson, 1998). Within the parametric EKC literature many studies rely upon unit root and cointegration analysis given the widespread non-rejection of the unit root hypothesis for GDP. With the exception of very few papers who note and bypass in one way or another the associated problems (see Bradford, Fender, Shore, and Wagner, 2005; Müller-Fürstenberger and Wagner, 2007; Wagner, 2008), the empirical EKC literature fails to acknowledge the implications of the presence of nonlinear transformations of unit root processes. In a typical EKC, compare (19) below, emissions are regressed on GDP and GDP squared. Since (log per capita) GDP is often well characterized as being a unit root process, GDP squared is a nonlinear transformation of an integrated process and regressions involving such processes require different asymptotic theory than the usual 'linear' unit root and cointegration analysis.

In this paper we derive the asymptotic distributions of both the OLS estimator as well as of a *fully modified* OLS (FM-OLS) estimator of equations containing deterministic variables, stationary regressors, integrated regressors and integer powers of integrated regressors. The results we obtain resemble in several respects that of linear cointegration analysis as derived in Phillips and Hansen (1990) and rely upon the important contributions of Chang, Park, and Phillips (2001) and Park and Phillips (1999, 2001). First, the OLS estimator is consistent, but its limiting distribution is contaminated by so called second order bias terms, rendering valid inference infeasible. Second, the proposed FM-OLS estimator has a limiting distribution that is free of second order biases and thus forms the basis for asymptotically valid χ^2-inference for certain hypotheses. Third, this property of the limiting distribution of the FM-OLS estimator also forms the basis for specification testing based on augmented respectively auxiliary regressions including higher order polynomial powers of the integrated regressors and/or polynomial powers of additional integrated regressors. In this respect we consider tests based on both the Wald and Lagrange Multiplier testing principles. Fourth, we consider a KPSS type (compare Kwiatkowski, Phillips, Schmidt, and Shin, 1992) cointegration test to *directly* test the null hypothesis of nonlinear cointegration for a given specification. Since the asymptotic distribution of this test is contaminated by nuisance parameters we follow Choi and Saikkonen (2005) and present a sub-sample version of the test that has an asymptotic distribution free of nuisance parameters. Since the sub-sample test can be used in conjunction with the Bonferroni bound we investigate the potential performance gains that can be realized by using adjusted Bonferroni bound test procedures that are less conservative, such as those proposed in Simes (1986) or Rom (1990).

We conduct a small simulation study to assess the performance of the proposed methods. The findings show that the FM-OLS correction has only modest effects on the biases of the coefficients estimated by FM-OLS compared to the – also consistent – OLS estimates. The effects are much larger on the tests, whose performance hinges crucially on applying appropriate FM-OLS corrections. The Wald and Lagrange Multiplier tests behave very similarly and perform very well in terms of size. Their power performance depends, as expected, quite strongly upon the auxiliary regressors as well as the alternative considered. With respect to the sub-sample KPSS type tests we note that all considered modifications of the Bonferroni bound lead to very similar performance in the simulations. The KPSS type tests tend to be undersized for small samples and their power increases quite slowly with the sample size. However, their power is quite similar for all considered alternatives, which is consistent with the fact that the KPSS type tests are not specified against any particular alternative.

After the simulations we turn to the empirical analysis where we study the relationship between CO_2 respectively SO_2 emissions and GDP for a panel of 19 early developed countries over the period 1870–2000. When considering a quadratic formulation of the EKC we find support, based on the LM specification test, for eight countries for CO_2 emissions and for five countries SO_2 emissions. Allowing for smooth asymmetries by modeling a cubic EKC leads to non-rejections of a nonlinear cointegrating relationship in three more countries for both CO_2 and SO_2 emissions. The turning points implied by the FM-OLS coefficient estimates are with very few exceptions reasonable in-sample values.

The paper is organized as follows. In Section 2 we derive the asymptotic results for the estimators and tests. Section 3 contains a small simulation study to assess the finite sample performance of the proposed methods and Section 4 contains the results of the EKC analysis. Section 5 briefly summarizes and concludes. Two appendices follow the main text: Appendix A contains the proofs of all propositions results and Appendix B collects additional material related to the empirical EKC analysis.

We use the following notation: As usual, the symbols $\Rightarrow$ and $\to_p$ signify weak convergence and convergence in probability, respectively. Definitional equality is signified by $:=$. Further, $a_T = O(T^n)$ (respectively $a_T = O_p(T^n)$) denotes that $\{a_T\}$ is at most of order T^n (in probability). Standard Brownian motions are denoted as $W(r)$ or in short W, whereas Brownian motions with non-identity covariance matrices (specified in the context) are denoted with $B(r)$ or B. For integrals of the form $\int_0^1 B(s)ds$ and $\int_0^1 B(s)dB(s)$ we use short-hand notation $\int B$ and $\int BdB$. For notational simplicity we also often drop function arguments. With $\lfloor x \rfloor$ we denote the integer part of $x \in \mathbb{R}$ and $diag(\cdot)$ denotes a diagonal matrix with the entries specified throughout. $\mathbb{E}$ denotes the expected value and L denotes the backward-shift operator, i.e. $L\{x_t\}_{t\in\mathbb{Z}} = \{x_{t-1}\}_{t\in\mathbb{Z}}$.

2 Econometric Theory

2.1 Setup and Assumptions

We consider the following equation including stationary regressors w_{it}, $i = 1, \ldots, n$, a constant, polynomial time trends up to power q and integer powers of integrated regressors $x_{jt}, j = 1, \ldots, m$ up to degrees p_j

$$y_t = w_t'\theta_w + D_t'\theta_D + \sum_{j=1}^{m} X_{jt}'\theta_{X_j} + u_t \quad , \quad \text{for } t = 1, \ldots, T \tag{1}$$

with $w_t := [w_{1t}, \ldots, w_{nt}]'$, $D_t := [1, t, t^2, \ldots, t^q]'$, $x_t := [x_{1t}, \ldots, x_{mt}]'$, $X_{jt} := [x_{jt}, x_{jt}^2, \ldots, x_{jt}^{p_j}]'$ and the parameter vectors $\theta_D \in \mathbb{R}^{q+1}$, $\theta_w \in \mathbb{R}^n$ and $\theta_{X_j} \in \mathbb{R}^{p_j}$. Furthermore define for later use $X_t := [X_{1t}', \ldots, X_{mt}']'$, $Z_t := [w_t', D_t', X_t']'$ and $p := \sum_{j=1}^{m} p_j$.

In a more compact way,

$$\begin{aligned} y &= w\theta_w + D\theta_D + X\theta_X + u \qquad (2) \\ &= Z\theta + u, \end{aligned}$$

with $y := [y_1, \ldots, y_T]'$, $u := [u_1, \ldots, u_T]'$, $Z := [w \;\; D \;\; X]$ and $\theta = [\theta_w' \;\; \theta_D' \;\; \theta_X']' \in \mathbb{R}^{(q+1)+n+p}$ and

$$w := \begin{bmatrix} w_1' \\ \vdots \\ w_T' \end{bmatrix} \in \mathbb{R}^{T\times n}, \; D := \begin{bmatrix} D_1' \\ \vdots \\ D_T' \end{bmatrix} \in \mathbb{R}^{T\times(q+1)}, \; X := \begin{bmatrix} X_1' \\ \vdots \\ X_T' \end{bmatrix} \in \mathbb{R}^{T\times p}.$$

Let us next state the assumptions concerning the regressors and the error processes:

Assumption 1 *The processes* $\{\Delta x_t\}_{t\in\mathbb{Z}}$, $\{w_t\}_{t\in\mathbb{Z}}$ *and* $\{u_t\}_{t\in\mathbb{Z}}$ *are generated as*

$$\begin{aligned} \Delta x_t = v_t &= C_v(L)\varepsilon_t = \sum_{j=0}^{\infty} c_{vj}\varepsilon_{t-j} \\ w_t &= C_w(L)\eta_t = \sum_{j=0}^{\infty} c_{wj}\eta_{t-j} \\ u_t &= C_u(L)\zeta_t = \sum_{j=0}^{\infty} c_{uj}\zeta_{t-j}, \end{aligned}$$

with the summability conditions

$$C_v(1) \neq 0, \;\; \sum_{j=0}^{\infty} j||c_{vj}|| < \infty \;, \;\; \sum_{j=0}^{\infty} j^{1/2}||c_{wj}|| < \infty \;, \;\; \sum_{j=0}^{\infty} j^{1/2}|c_{uj}| < \infty.$$

Furthermore we assume that the regressors are predetermined, i.e. we assume that the process $\{\xi_t^0\}_{t\in\mathbb{Z}} = \{[\varepsilon_{t+1}', \eta_{t+1}', \zeta_t]'\}_{t\in\mathbb{Z}}$ *is a stationary and ergodic martingale difference sequence with natural filtration* $\mathcal{F}_t = \sigma\left(\{\xi_s^0\}_{-\infty}^t\right)$ *and denote the (conditional) covariance matrix by*

$$\Sigma^0 = \begin{bmatrix} \Sigma_{\varepsilon\varepsilon} & \Sigma_{\varepsilon\eta} & \Sigma_{\varepsilon\zeta} \\ \Sigma_{\eta\varepsilon} & \Sigma_{\eta\eta} & \Sigma_{\eta\zeta} \\ \Sigma_{\zeta\varepsilon} & \Sigma_{\zeta\eta} & \sigma_\zeta^2 \end{bmatrix} := \mathbb{E}(\xi_t^0(\xi_t^0)'|\mathcal{F}_{t-1}).$$

We also assume that $\Sigma_{ww} := \mathbb{E}w_t w_t' > 0$.

The above assumptions allow to draw on the asymptotic results of Chang, Park, and Phillips (2001) and Park and Phillips (1999, 2001). The assumption that the regressors are predetermined implies that the conditional expectation $\mathbb{E}(y_t|\mathcal{F}_{t-1}) = 0$, which is usual in nonlinear regression theory. Considering the stationary regressors to have zero mean is only done for convenience and is not a restriction since typically an intercept will be included in a regression. The assumption $C_v(1) \neq 0$ implies that x_t is indeed an integrated process whereas $\Sigma_{ww} > 0$ is mainly put in place for convenience as it is likely to be fulfilled in all practical applications.

Further additionally required moment assumptions given below are also similar to those formulated in Chang, Park, and Phillips (2001).

Assumption 2 *For the process $\{\xi_t^0\}_{t\in\mathbb{Z}}$ the following conditions hold:*

1. $\sup_{t\geq 1} \mathbb{E}(\|\xi_t^0\|^r|\mathcal{F}_{t-1}) < \infty$ *a.s. for some* $r > 4$.
2. $\mathbb{E}\left((\xi_{i,t}^0)^2\xi_{j,t-l}^0\right) = 0$ *for all* i, j *and for all* $l \geq 1$.
3. ζ_t *is i.i.d. with* $\mathbb{E}(|\zeta|^r) < \infty$ *for some* $r > 8$ *and its distribution function is absolutely continuous with respect to the Lebesgue measure and for the characteristic function* φ *it holds that* $\varphi(\lambda) = o(|\lambda|^{-\delta})$ *as* $\lambda \to \infty$ *for some* $\delta > 0$.

The above assumptions are sufficient for the following invariance principle to hold for $\{\xi_t\}_{t\in\mathbb{Z}} = \{[v_{t+1}', w_{t+1}', u_t]'\}_{t\in\mathbb{Z}}$ using the Beveridge-Nelson decomposition (compare Phillips and Solo, 1992)

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]} \xi_t \Rightarrow B(r) = \begin{bmatrix} B_v(r) \\ B_w(r) \\ B_u(r) \end{bmatrix}. \tag{3}$$

Note here that it holds that $B(r) = \Omega^{1/2}W(r)$ with the long-run covariance matrix $\Omega := \sum_{h=-\infty}^{\infty} \mathbb{E}\left(\xi_0\xi_h'\right)$. We also define the one-sided long-run covariance $\Lambda := \sum_{h=0}^{\infty} \mathbb{E}\left(\xi_0\xi_h'\right)$ and both covariance matrices are partitioned according to the partitioning of ξ_t, i.e.:

$$\Omega = \begin{bmatrix} \Omega_{vv} & \Omega_{vw} & \Omega_{vu} \\ \Omega_{wv} & \Omega_{ww} & \Omega_{wu} \\ \Omega_{uv} & \Omega_{uw} & \omega_{uu} \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \Lambda_{vv} & \Lambda_{vw} & \Lambda_{vu} \\ \Lambda_{wv} & \Lambda_{ww} & \Lambda_{wu} \\ \Lambda_{uv} & \Lambda_{uw} & \lambda_{uu} \end{bmatrix}.$$

When referring to quantities corresponding to only one of the nonstationary regressors and its powers, e.g. X_{jt}, we use the according notation, e.g. $B_{v_j}(r)$ or $\Lambda_{v_j u}$.

To study the asymptotic behavior of the estimators, we next introduce appropriate weighting matrices, whose entries reflect the divergence rates of the corresponding variables. Thus, denote with $G(T) = diag\{G_w(T), G_D(T), G_X(T)\}$, where for notational brevity we often use $G := G(T)$.

The three diagonal sub-matrices are given by:

$$G_w(T) := \begin{pmatrix} T^{-1/2} & & \\ & \ddots & \\ & & T^{-1/2} \end{pmatrix} \in \mathbb{R}^{n\times n}, G_D(T) := \begin{pmatrix} T^{-1/2} & & \\ & \ddots & \\ & & T^{-(q+1/2)} \end{pmatrix} \in \mathbb{R}^{(q+1)\times(q+1)},$$

$$G_X(T) := \begin{pmatrix} G_{X_1} & & \\ & \ddots & \\ & & G_{X_m} \end{pmatrix} \in \mathbb{R}^{p\times p} \text{ with } G_{X_j} := \begin{pmatrix} T^{-1} & & \\ & \ddots & \\ & & T^{-\frac{p_j+1}{2}} \end{pmatrix} \in \mathbb{R}^{p_j\times p_j}.$$

Using these weighting matrices, we can define the following limits of the major building blocks. For t such that $\lim_{T\to\infty} t/T = r$ the following results hold:

$$\begin{aligned} \lim_{T\to\infty} \sqrt{T} G_D(T) D_t &= \lim_{T\to\infty} \begin{pmatrix} 1 & & \\ & \ddots & \\ & & T^{-q} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ t^q \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ r^q \end{pmatrix} =: D(r) \\ \lim_{T\to\infty} \sqrt{T} G_{X_j}(T) X_{jt} &= \lim_{T\to\infty} \begin{pmatrix} T^{-1/2} & & \\ & \ddots & \\ & & T^{-p_j/2} \end{pmatrix} \begin{pmatrix} x_{jt} \\ \vdots \\ x_{jt}^{p_j} \end{pmatrix} = \begin{pmatrix} B_{v_j} \\ \vdots \\ B_{v_j}^{p_j} \end{pmatrix} =: \mathbf{B}_{v_j}(r), \end{aligned}$$

separating here the coordinates of $v_t = [v_{1t}, \ldots, v_{mt}]'$ corresponding to the different variables x_{jt}. The first result is immediate and the second follows from Chang, Park, and Phillips (2001, Lemma 5). The stacked vector of the scaled polynomial transformations of the integrated processes is denoted as $\mathbf{B}_v(r) := [\mathbf{B}_{v_1}(r)', \ldots, \mathbf{B}_{v_m}(r)']'$. We are confident that D as defined in (2) is not confused with $D(r)$ defined above even when the latter is used in abbreviated form D in integrals.

Remark 1 *More general deterministic components can be included with the necessary condition being that the correspondingly defined limit quantity satisfies $\int DD' > 0$, i.e. that the considered functions are linearly independent in $L^2[0,1]$. This allows in addition to the polynomial trends on which we focus in this paper e.g. also to include time dummies, broken trends or trigonometric functions of time (compare the discussion in Park, 1992).*

The relationship postulated in (1) is restrictive in the sense that e.g. no cross-products of the form $x_{it}^m x_{jt}^n$ or $t^m x_{jt}^n$ are included. Considering such cross-terms increases not only the flexibility of the functional form but also immediately allows for an interpretation of the estimated relationship as a Taylor expansion of an unknown nonlinear function. The theory developed in this paper, based on the underlying results of Park and Phillips (1999, 2001), can be extended to include these cross-terms. However, the curse of dimensionality will often limit the practical usefulness of specifications including all cross-terms. Also for the empirical application in this paper we consider only one integrated regressor, namely per capita GDP.

2.2 OLS Estimation

We first study the asymptotic behavior of the OLS estimator. As in the linear cointegration case, its limiting distribution is contaminated by nuisance parameters due to serial correlation in the error process $\{u_t\}_{t\in\mathbb{Z}}$ and endogeneity of $\{\Delta X_t\}_{t\in\mathbb{Z}}$. Both of these aspects are very similar to those in the linear case as in Phillips and Hansen (1990) and are summarized in Proposition 1.

Proposition 1 *Let y_t be generated from (1) with the regressors Z_t and errors u_t satisfying Assumptions 1 and 2. Then the asymptotic distribution of the OLS estimator $\hat{\theta} := (Z'Z)^{-1}Z'y$ is given by*

$$G^{-1}(\hat{\theta}-\theta) = \begin{bmatrix} G_w^{-1}(\hat{\theta}_w-\theta_w) \\ G_D^{-1}(\hat{\theta}_D-\theta_D) \\ G_X^{-1}(\hat{\theta}_X-\theta_X) \end{bmatrix} \Rightarrow \begin{bmatrix} \Sigma_{ww}^{-1}N_{wu} \\ \left[\int \tilde{D}\tilde{D}'\right]^{-1}\left\{\int \tilde{D}dB_u - \int D\mathbf{B}_v'\left[\int \mathbf{B}_v\mathbf{B}_v'\right]^{-1}M\right\} \\ \left[\int \tilde{\mathbf{B}}_v\tilde{\mathbf{B}}_v'\right]^{-1}\left\{\int \tilde{\mathbf{B}}_v dB_{u.v} - \int \tilde{\mathbf{B}}_v dB_v'\Omega_{vv}^{-1}\Omega_{vu} + M\right\} \end{bmatrix} \tag{4}$$

where $B_{u.v}(r) := B_u(r) - \Omega_{uv}\Omega_{vv}^{-1}B_v(r)$ with corresponding variance $\omega_{u.v} := \omega_u - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}$ and $N_{wu} := \lim_{T\to\infty}\frac{1}{\sqrt{T}}\sum_{t=1}^T w_t u_t$. The random variable N_{wu} is normally distributed with mean $\Sigma_{wu} := \mathbb{E}(w_t u_t)$ and variance depending upon the coefficients $c_{w,j}$, $c_{u,j}$, $\Sigma_{\eta\eta}$ and σ_ζ^2 given in Assumption 1. Furthermore

$$\tilde{D} := D - \int D\mathbf{B}_v'\left(\int \mathbf{B}_v\mathbf{B}_v'\right)^{-1}\mathbf{B}_v\,,$$
$$\tilde{\mathbf{B}}_v := \mathbf{B}_v - \int \mathbf{B}_v D'\left(\int DD'\right)^{-1}D,$$

and

$$M := \begin{bmatrix} M_1 \\ \vdots \\ M_m \end{bmatrix} \quad \textit{where } M_j := \Lambda_{v_j u}\begin{bmatrix} 1 \\ 2\int B_{v_j}(r)dr \\ \vdots \\ p_j\int B_{v_j}(r)^{p_j-1}dr \end{bmatrix}. \tag{5}$$

The limiting distribution of the consistent OLS estimator displayed in (4) is contaminated by so-called *second order* bias terms: the serial correlation bias and the endogeneity bias, using the same names in our nonlinear setup as used in the limiting distribution of the OLS estimator in the linear cointegration case (see Phillips and Hansen, 1990). Note that when X_t is strictly exogenous, these bias terms vanish with $\Lambda_{vu} = \Omega_{vu} = 0$. If this is not the case, standard inference on the parameters becomes invalid due to the presence of these bias terms.

Remark 2 *Note that the serial correlation bias term M, which is due to correlation between u_t and v_t, appears not only in the limiting distribution of $\hat{\theta}_X$, but also in that of $\hat{\theta}_D$, reflecting the asymptotic correlation between deterministic and stochastic trends. Thus, putting these two blocks of the coefficient vector θ together in $\theta_N := \begin{bmatrix} \theta_D' & \theta_X' \end{bmatrix}'$, we can explicitly identify the source of the serial correlation bias by writing the limiting distribution of the OLS estimator of θ_D as*

$$G_N^{-1}\left(\hat{\theta}_N - \theta_N\right) \Rightarrow \left(\int JJ'\right)^{-1}\left\{\int JdBu + \begin{pmatrix} 0_{(q+1)\times 1} \\ M \end{pmatrix}\right\},$$

for $J(r) := \begin{bmatrix} D(r)' & \mathbf{B}_v(r)' \end{bmatrix}'$ and $G_N := diag(G_D, G_X)$.

2.3 Fully Modified OLS Estimation

Two ways to remove the bias terms present in the OLS limiting distributions have been proposed in the cointegration literature. These are *fully modified* OLS (FM-OLS) estimation (see Phillips and Hansen, 1990) based on a direct non-parametric correction and *dynamic* OLS (D-OLS) estimation (see Saikkonen, 1991) where the correction is achieved by running lead and lag augmented regressions. In this paper we consider FM-OLS estimation which requires consistent estimators of the bias terms. In this respect define

$$M^* := \begin{bmatrix} M_1^* \\ \vdots \\ M_m^* \end{bmatrix}, \quad M_j^* := \hat{\Lambda}_{v_j u}^+ \begin{bmatrix} T \\ 2\sum x_{jt} \\ \vdots \\ p_j \sum x_{jt}^{p_j - 1} \end{bmatrix}, \tag{6}$$

with a consistent estimator $\hat{\Lambda}_{v_j u}^+ := \hat{\Lambda}_{v_j u} - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\hat{\Lambda}_{vv_j}$. Once appropriately scaled the quantity M^* in (6) converges to M as given in (5), which in conjunction with using the transformed dependent variable[1] $y_t^+ := y_t - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_t$, $y^+ := [y_1^+, \ldots, y_T^+]'$ leads to an asymptotic distribution that is free of bias terms as summarized in the following Proposition 2.

[1]For notational simplicity we ignore the dependence of y^+ upon the specific consistent long-run covariance estimator chosen.

Proposition 2 *Let y_t be generated by* (1) *with the regressors Z_t and errors u_t satisfying Assumptions 1 and 2. Define the FM-OLS estimator of θ as*

$$\hat{\theta}^{+} := (Z'Z)^{-1} \left(Z'y^{+} - A^{*} \right),$$

with

$$A^{*} := \left[\begin{array}{l} \hat{\Sigma}_{wu}^{+} \\ 0_{(q+1)\times 1} \\ M^{*} \end{array} \right]$$

with $\hat{\Sigma}_{wu}^{+}$ a consistent estimator of $\Sigma_{wu}^{+} := \Sigma_{wu} - \Sigma_{wv}\Omega_{vv}^{-1}\Omega_{vu}$ and M^{*} as given in (6) *with consistent estimators of the required long-run (co)variances. Then the asymptotic distribution of $\hat{\theta}^{+}$ is given by*

$$G^{-1}\left(\hat{\theta}^{+} - \theta\right) = \left[\begin{array}{c} G_w^{-1}(\hat{\theta}_w^{+} - \theta_w) \\ G_D^{-1}(\hat{\theta}_D^{+} - \theta_D) \\ G_X^{-1}(\hat{\theta}_X^{+} - \theta_X) \end{array} \right] \Rightarrow \left[\begin{array}{c} \Sigma_{ww}^{-1} N_{wu.v} \\ \left[\int \tilde{D}\tilde{D}'\right]^{-1} \int \tilde{D} dB_{u.v} \\ \left[\int \tilde{\mathbf{B}}_v \tilde{\mathbf{B}}_v'\right]^{-1} \int \tilde{\mathbf{B}}_v dB_{u.v} \end{array} \right], \tag{7}$$

with a normally distributed mean zero random variable $N_{wu.v} := \lim_{T\to\infty} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_t u_t^{+}$, where $u_t^{+} := u_t - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1} v_t$.

Using the quantities defined in Remark 2 it holds more compactly written that $G_N^{-1}\left(\hat{\theta}_N^{+} - \theta_N\right) \Rightarrow \left(\int JJ'\right)^{-1} \int J dB_{u.v}$. The limiting distribution of $G_N^{-1}(\hat{\theta}_N^{+} - \theta_N)$ is free of second order bias terms and mixed normal with mean zero. This stems from the fact that the vector $\tilde{\mathbf{B}}_v$ is, by construction, independent of $B_{u.v}$.

As shown in Phillips and Hansen (1990, Theorem 5.1) the special form of the FM-OLS limiting distribution allows for asymptotic χ^2-inference for testing *certain* linear hypothesis on the coefficients by using the Wald test. From the discussion in Phillips and Hansen (1990, p. 106), in particular from the corresponding proofs in their paper, it becomes clear that only certain linear hypotheses can be tested with asymptotic χ^2-inference. A similar result that allows to test for certain hypotheses, which we formulate for notational convenience for θ_N as defined in Remark 2, can be established in our setup.

Proposition 3 *Let y_t be generated by* (1) *with the regressors Z_t and errors u_t satisfying Assumptions 1 and 2. Consider s linearly independent restrictions collected in*

$$H_0 : R\theta_N = r,$$

with $R \in \mathbb{R}^{s \times q+1+p}$ with full rank s and $r \in \mathbb{R}^s$. Furthermore let $\hat{\omega}_{u.v}$ denote a consistent estimator of $\omega_{u.v}$. Then it holds with $Z_N = [D\ X]$ that the Wald statistic

$$W := \left(R\hat{\theta}_N^+ - r\right)' \left[\hat{\omega}_{u.v} R \left(Z_N' Z_N\right)^{-1} R'\right]^{-1} \left(R\hat{\theta}_N^+ - r\right) \tag{8}$$

is under the null hypothesis asymptotically distributed as χ_s^2 under one of the following conditions:

(i) H_0 only involves coefficients with the same convergence rate, or

(ii) each of the restrictions in H_0 involves only one coefficient, i.e. the off-diagonal elements of R are all equal to 0.

The above result implies that for instance the appropriate t-statistic for coefficient θ_i, with θ_i a component of θ_N, given by $t_{\theta_i} := \frac{\hat{\theta}_i^+}{\sqrt{\hat{\omega}_{u.v}(Z'Z)^{-1}_{[i,i]}}}$, is asymptotically standard normally distributed. Note furthermore that hypothesis testing for the coefficients θ_w can simply be based on their asymptotic normal distribution.

2.4 Specification Testing based on Augmented and Auxiliary Regressions

Testing the correct specification of equation (1) is clearly an important issue. In this respect we are particularly interested in the prevalence of cointegration, i.e. stationarity of u_t. Absence of cointegration can be due to several reasons. First, there is no cointegrating relationship of any functional form between y_t and x_t. Second, y_t and x_t are nonlinearly cointegrated but the functional relationship is different than postulated by equation (1). This case covers the possibilities of missing higher order polynomial terms or cointegration with a different functional form of the relationship. Third, the absence of cointegration is due to missing explanatory variables in equation (1).

In a general formulation all the above possibilities can be cast into a testing problem within the augmented regression

$$y_t = Z_t'\theta + F(x_t, q_t, \theta_F) + \phi_t, \tag{9}$$

where F is such that $F(x_t, g_t, 0) = 0$ and q_t denotes additional integrated regressors. If cointegration prevails in (1) then $\theta_F = 0$ and $\phi_t = u_t$.

In many cases the researcher will not have a specific parametric formulation in mind for the function $F(\cdot)$, which implies that typically the unknown $F(\cdot)$ is replaced by a partial sum approximation. This approach has a long tradition in specification testing in a stationary setup, see

Ramsey (1969), Phillips (1983), Lee, White, and Granger (1993) or de Benedictis and Giles (1998). Given our FM-OLS results it appears convenient to replace the unknown $F(\cdot)$ by using polynomial powers of the integrated regressors, which will include higher order powers larger than p_j for the components x_{jt} of x_t and powers larger equal than 1 for the additional integrated regressors q_{it}.

Of course this simple approach is also subject to the discussion in the introduction in that no multivariate expansion is considered. However, for specification analysis the advantage of a parsimonious setup may outweigh the potential disadvantages of considering only univariate polynomials since a test based on such a formulation will also have power against alternatives where e.g. products terms are present. Clearly, the power properties of tests based on univariate polynomials depend upon the unknown alternative $F(\cdot)$ and will be the more favorable the more $F(\cdot)$ 'resembles' univariate polynomials. This trade-off is exactly the same as in the stationary case, as also discussed in Hong and Phillips (2008).

Denote with $\bar{X}_{jt} := [x_{jt}^{p_j+1}, x_{jt}^{p_j+2}, \ldots, x_{jt}^{p_j+r_j}]'$ for $j = 1, \ldots, m$, $Q_{it} := [q_{it}^1, q_{it}^2, \ldots, q_{it}^{s_i}]'$ for $i = 1, \ldots, k$, $F_t := [\bar{X}_{1t}', \ldots, \bar{X}_{mt}', Q_{1t}', \ldots, Q_{kt}']'$ and $F := [F_1', \ldots, F_T']'$. Using this notation the augmented polynomial regression including higher order polynomial powers of the regressors x_{jt} and polynomial powers of additional integrated regressors q_{it} can be written as

$$y = Z\theta + F\theta_F + \phi, \tag{10}$$

with $\phi := [\phi_1, \ldots, \phi_T]'$. If equation (10) is well specified the parameters can be estimated consistently by FM-OLS according to Proposition 2 if the additional regressors q_{it} fulfill the necessary assumptions stated in Section 2.1 which are now modified to accommodate additional regressors.

Assumption 3 *When considering additional regressors q_{it} and their polynomial powers define $\tilde{v}_t := [v_t', (v_t^*)']' = [\Delta x_t', \Delta q_t']'$, with $v_t^* = \Delta q_t$ and $q_t = [q_{1t}, \ldots, q_{kt}]'$. Assumptions 1 and 2 are extended such that they are fulfilled for the extended process $\tilde{v}_t$ generated by $C_{\tilde{v}}(L)\tilde{\varepsilon}_t$, with $C_{\tilde{v}}(L)$ and $\tilde{\varepsilon}_t$ also extended accordingly.*

Note that equation (10) can be well-specified for different reasons. The first is that (1) is a cointegrating relationship, in which case consistently estimated coefficients $\hat{\theta}_F^+$ will converge to their true value equal to 0. The second possibility is that (1) is misspecified, but the extended equation (10) is well-specified. In this case at least some entries of $\hat{\theta}_F^+$ will converge to their non-zero true values. In case that (10) and consequently also (1) are misspecified and ϕ_t is not stationary, spurious regression results similar to the linear case that lead to non-zero limit coefficients apply. Consequently, a specification test based on $H_0 : \theta_F = 0$ is consistent against the three discussed forms of misspecification of (1) discussed in the beginning of the sub-section.

Testing the restriction $\theta_F = 0$ in (10) can be done in several ways. One is given by FM-OLS estimation of the augmented regression (10) and performing a Wald test on the estimated coefficients using Proposition 3. Another possibility is to use the FM-OLS residuals of the original equation (2) and to perform a Lagrange Multiplier RESET type test in an auxiliary regression. These two possibilities are discussed in turn.

Proposition 4 *Let y_t be generated by* (1) *with the regressors Z_t, Q_t and errors u_t satisfying Assumptions 1, 2 and 3. Denote with $\hat{\theta}_F^+$ the FM-OLS estimator of θ_F in equation* (10)*, with $\tilde{F}_N = F - Z_N(Z_N'Z_N)^{-1}Z_N'F$, and as above $Z_N = [D\ X]$ and let $\hat{\omega}_{u.\tilde{v}}$ be a consistent estimator of $\omega_{u.\tilde{v}}$. Then it holds that the Wald test statistic for the null hypothesis $H_0 : \theta_F = 0$ in equation* (10)*, given by*

$$T_W := \frac{\left(\hat{\theta}_F^+\right)' \left(\tilde{F}_N'\tilde{F}_N\right) \hat{\theta}_F^+}{\hat{\omega}_{u.\tilde{v}}}, \tag{11}$$

is under the null hypothesis asymptotically distributed as χ_b^2, with $b := \sum_{j=1}^m r_j + \sum_{j=1}^n s_j$.

Note that the required variance and covariance estimates in Proposition 4 are all based on the $(m+k)$-dimensional process $\tilde{v}_t$. The result given in Proposition 4 follows straightforwardly from Propositions 2 and 3.

The basis of the Lagrange Multiplier (LM) test are the FM-OLS residuals $\hat{u}_t^+$ of (2), which are regressed on $\tilde{F}$ in the auxiliary regression

$$\hat{u}^+ = \tilde{F}\theta_{\tilde{F}} + \psi_t. \tag{12}$$

with $\hat{u}^+ = [\hat{u}_1^+, \ldots, \hat{u}_T^+]'$. Clearly, to allow for asymptotic standard inference the coefficients $\theta_{\tilde{F}}$ have to be estimated by FM-OLS to achieve a second order bias free limiting distribution, since the limiting distribution of the OLS estimator of $\theta_{\tilde{F}}$ in (12) also depends upon second order bias terms (see the proof of Proposition 5 in Appendix A for details). The FM-OLS estimator as well as the test statistic for testing the hypothesis $\theta_{\tilde{F}} = 0$ are presented in the following proposition for the case that (1) is well specified. Consistency of the tests against the above-discussed forms of misspecification of (1) follows from the same arguments as for the Wald test.

Proposition 5 *Let y_t be generated by* (1) *with the regressors Z_t, Q_t and errors u_t satisfying Assumptions 1, 2 and 3. Define the fully modified OLS estimator of $\theta_{\tilde{F}}$ in equation* (12) *as*

$$\hat{\theta}^+_{\tilde{F}} := \left(\tilde{F}'\tilde{F}\right)^{-1}\left(\tilde{F}'\hat{u}^+ - O^{F*} - M^{F*} + k^{F*}M^*\right), \tag{13}$$

with

$$O^{F*} := \hat{\Omega}_{u\tilde{v}}\hat{\Omega}^{-1}_{\tilde{v}\tilde{v}}\sum F_t\tilde{v}_t - \hat{\Omega}_{uv}\hat{\Omega}^{-1}_{vv}\sum F_t v_t$$

and $M^{F*} := [M'_{\bar{X}_1}, \ldots, M'_{\bar{X}_m}, M'_{Q_1}, \ldots, M'_{Q_k}]'$, where

$$M_{\bar{X}_j} = \hat{\Lambda}^+_{v_j u}\begin{bmatrix} (p_j+1)\sum x_{jt}^{p_j} \\ \vdots \\ (p_j+r_j)\sum x_{jt}^{p_j+r_j-1}\end{bmatrix}, \quad M_{Q_i} = \hat{\Lambda}^+_{v_i^* u}\begin{bmatrix} T \\ 2\sum q_{it} \\ \vdots \\ s_i \sum q_{it}^{s_i-1}\end{bmatrix},$$

$k^{F*} = F'\tilde{X}(\tilde{X}'\tilde{X})^{-1}$, $\tilde{X} = X - D\left(D'D\right)^{-1}D'X$, $\hat{\Lambda}^+_{v_j u}$ as defined above Proposition 2 and $\hat{\Lambda}^+_{v_i^* u} := \hat{\Lambda}_{v_i^* u} - \hat{\Omega}_{uv}\hat{\Omega}^{-1}_{vv}\hat{\Lambda}_{vv_i^*}$. Let $\hat{\omega}_{u,\tilde{v}}$ denote a consistent estimator of $\omega_{u,\tilde{v}}$. Then it holds that the LM test statistic for the null hypothesis $H_0 : \theta_{\tilde{F}} = 0$ in (12)

$$T_{LM} := \frac{\left(\hat{\theta}^+_{\tilde{F}}\right)'\left(\tilde{F}'\tilde{F}\right)\hat{\theta}^+_{\tilde{F}}}{\hat{\omega}_{u.\tilde{v}}}, \tag{14}$$

is under the null hypothesis asymptotically distributed as χ^2_b, with $b = \sum_{j=1}^m r_j + \sum_{j=1}^k s_j$.

Remark 3 *Proposition 5 is as a generalization of the modified RESET test considered in Hong and Phillips (2008, Theorem 3), who consider a related test in a bivariate linear cointegrating relationship with only one I(1) regressor and without deterministic and stationary variables, i.e. they consider the case $q = 0$, $n = 0$, $m = 1$ and $p = 1$. A second difference to our result is that Hong and Phillips use the OLS residuals $\hat{u}_t$ of the linear cointegrating relationship in the auxiliary regression, which leads to different bias correction terms than ours based on the FM-OLS residuals $\hat{u}_t^+$.*

Remark 4 *In the misspecification analysis as discussed here we do not consider the deterministic and stationary regressors. With obvious modifications of the test statistics completely analogously also higher order deterministic components can be used in F. If one considers only higher order deterministic terms in F one arrives at tests similar to those of Park and Choi (1988) and Park (1990). These authors propose cointegration tests based on adding superfluous higher order deterministic trend terms. This approach is nested within ours. With respect to the stationary regressors* *the issue is different, since omission of stationary regressors with mean zero in* (1) *does not change that the corresponding error term is still stationary with mean zero and thus does not invalidate the presence of cointegration in* (1)*.*

Remark 5 *Note also that any selection of higher order polynomial terms can be chosen as additional regressors and one need not choose, as done for simplicity, a set of consecutive powers ranging from e.g.* $p_j + 1$ *to* $p_j + r_j$*. Again both propositions continue to hold with obvious modifications.*

2.5 KPSS Type Tests for Cointegration

In this section we discuss a residual based 'direct' test for nonlinear cointegration which prevails in (1) if the error process $\{u_t\}_{t\in\mathbb{Z}}$ is stationary. To test this null hypothesis directly we present a Kwiatkowski, Phillips, Schmidt, and Shin (1992), in short KPSS, type test statistic based on the FM-OLS residuals $\hat{u}_t^+$ of (1). The KPSS test is a variance-ratio test, comparing estimated short- and long-run variances, that converges toward a well defined distribution under stationarity but diverges under the unit root alternative. Note that this as well as other related tests can be interpreted to a certain extent as specification test as well, since integrated errors also prevail if e.g. relevant I(1) regressors are omitted in (1). The test statistic is given by

$$CT := \frac{1}{T\hat{\omega}_{u.v}} \sum_{t=1}^{T} \left(\frac{1}{\sqrt{T}} \sum_{j=1}^{t} \hat{u}_j^+ \right)^2, \tag{15}$$

with $\hat{\omega}_{u.v}$ a consistent estimator of the long-run variance $\omega_{u.v}$ of $\hat{u}_t^+$. The asymptotic distribution of this test statistic is considered in the following proposition.

Proposition 6 *Let* y_t *be generated by* (1) *with the regressors* Z_t *and errors* u_t *satisfying Assumptions 1 and 2 and let* $\hat{\omega}_{u.v}$ *be a consistent estimator of* $\omega_{u.v}$*, then the asymptotic distribution of the statistic* (15) *defined above is*

$$CT \Rightarrow \frac{1}{\omega_{u.v}} \int (B_{u.v}^*)^2,$$

with

$$B_{u.v}^*(r) := B_{u.v}(r) - \int_0^r D' \left[\int \tilde{D}\tilde{D}' \right]^{-1} \int \tilde{D} dB_{u.v} - \int_0^r \mathbf{B}_v' \left[\int \tilde{\mathbf{B}}_v \tilde{\mathbf{B}}_v' \right]^{-1} \int \tilde{\mathbf{B}}_v dB_{u.v}. \tag{16}$$

The above limiting distribution (16) depends upon the specification of the deterministic component, the number and the polynomial degrees of the integrated regressors as well as upon the correlation structure between $\{u_t\}_{t\in\mathbb{Z}}$ and $\{v_t\}_{t\in\mathbb{Z}}$. Albeit critical values can be simulated for any given constellation, basing tests upon the result in Proposition 6 appears to be impractical.

Like Choi and Saikkonen (2005), who consider a similar testing problem in a dynamic OLS estimation framework, we therefore propose to use a sub-sample based test statistic whose limiting distribution is free of nuisance parameters.

Proposition 7 *Under the same assumptions as in Proposition 6 it holds that*

$$CT_{b,i} := \frac{1}{b\hat{\omega}_{u.v}} \sum_{t=i}^{i+b-1} \left(\sum_{j=i}^{t} \frac{1}{\sqrt{b}} \hat{u}_j^+ \right)^2 \Rightarrow \int W^2$$

with b such that for $T \to \infty$ it holds that $b \to \infty$ and $\sqrt{b/T} \to 0$.

Note that for a given block size b there are $M := \lfloor T/b \rfloor$ sub-sample based test statistics, $\{CT_{b,i_1}, \ldots, CT_{b,i_M}\}$, that all lead to asymptotically valid statistics for the same null hypothesis. Basing a test on all these statistics might lead to reduced power and increased size (compare again Choi and Saikkonen, 2005). Therefore we consider using this set of statistics in combination with the Bonferroni inequality to modify the critical values using

$$\lim_{T\to\infty} \mathbb{P}\left(CT_{max} \leq c_{\alpha/M}\right) \geq 1 - \alpha,$$

where $CT_{max} := \max(CT_{b,i_1}, \ldots, CT_{b,i_M})$, suppressing the dependence of CT_{max} on b for notational brevity, and $c_{\alpha/M}$ denotes the α/M-percent critical value of the distribution of $\int W^2$. For the computation of the critical values from the distribution function, F say, of $\int W^2$ Choi and Saikkonen (2005) obtain the interesting result that

$$F(z) = \sqrt{2} \sum_{n=0}^{\infty} \frac{\Gamma(n+1/2)}{n!\Gamma(1/2)} (-1)^n \left(1 - f\left(\frac{g}{2\sqrt{z}}\right)\right), \quad z \geq 0, \tag{17}$$

with $f(x) = \frac{2}{\sqrt{\pi}} \int_0^x exp(-y^2)dy$ and $g = \sqrt{2}/2 + 2n\sqrt{2}$. Using this series representation and truncating the series at $n = 30$ we obtain the critical values for the required distribution used in the simulations and the empirical study. We present critical values based on $n = 30$ and for comparison also for $n = 10$ (as used in Choi and Saikkonen, 2005) in Table 1. We refer to the standard Bonferroni bound test procedure, where the null hypothesis is rejected if $CT_{max} \geq c_{\alpha/M}$, as Choi and Saikkonen (2005) test.

By construction a test based on the Bonferroni bound is conservative and is known to be particularly conservative when the test statistics used are highly correlated (see Hommel, 1986). In

Table 1: Critical values $c_{\frac{\alpha}{M}}$ from $\mathbb{P}\left[\int W^2 \geq c_{\frac{\alpha}{M}}\right] = \frac{\alpha}{M}$ for $\alpha = 5\%$ and 10%

M	5%	10%	M	5%	10%	M	5%	10%
Sum in (17) truncated at 30								
2	2.135	1.656	15	3.588	3.076	28	4.034	3.538
3	2.421	1.934	16	3.635	3.121	29	4.058	3.563
4	2.627	2.135	17	3.680	3.164	30	4.081	3.588
5	2.787	2.292	18	3.721	3.203	31	4.103	3.612
6	2.917	2.421	19	3.760	3.241	32	4.124	3.635
7	3.027	2.531	20	3.797	3.276	33	4.145	3.658
8	3.121	2.627	21	3.832	3.309	34	4.165	3.680
9	3.203	2.711	22	3.865	3.340	35	4.184	3.700
10	3.276	2.787	23	3.897	3.370	36	4.202	3.721
11	3.340	2.855	24	3.927	3.398	37	4.220	3.741
12	3.398	2.917	25	3.955	3.424	38	4.237	3.760
13	3.484	2.974	26	3.983	3.484	39	4.253	3.779
14	3.538	3.027	27	4.009	3.511	40	4.269	3.797
Sum in (17) truncated at 10								
2	2.135	1.656	15	3.582	3.081	28	3.997	3.533
3	2.421	1.934	16	3.627	3.128	29	4.018	3.558
4	2.626	2.135	17	3.669	3.172	30	4.038	3.582
5	2.785	2.292	18	3.709	3.214	31	4.058	3.605
6	2.912	2.421	19	3.746	3.253	32	4.076	3.627
7	3.031	2.531	20	3.781	3.291	33	4.094	3.649
8	3.128	2.626	21	3.813	3.326	34	4.111	3.669
9	3.214	2.710	22	3.844	3.360	35	4.127	3.689
10	3.291	2.785	23	3.873	3.392	36	4.143	3.709
11	3.360	2.852	24	3.900	3.422	37	4.158	3.728
12	3.422	2.912	25	3.926	3.452	38	4.172	3.746
13	3.480	2.977	26	3.951	3.480	39	4.186	3.763
14	3.533	3.031	27	3.974	3.507	40	4.199	3.781

the literature several less conservative modified Bonferroni bound type test procedures have been presented. Some of them are developed in Hommel (1988), Simes (1986) and Rom (1990). Denote the test statistics ordered in magnitude by $CT_b^{(1)} \geq \cdots \geq CT_b^{(M)}$. The modification of Hommel (1988) amounts to rejecting the null hypothesis if at least one of the test statistics $CT_b^{(j)} \geq c_{\alpha^H(j)}$ with $\alpha^H(j) = \frac{j}{C_M}\frac{\alpha}{M}$ and $C_M = 1 + 1/2 + \cdots + 1/M$. The modification of Simes (1986) is very similar and almost coincides with the procedure of Hommel with the only difference being that the additional adjustment factor C_M is not included, i.e. $\alpha^S(j) = j\frac{\alpha}{M}$. A further modification of the computation of the levels used in the sequential test procedure has been proposed in Rom (1990). For this modification the levels $\alpha^R(j)$ are computed recursively via $\alpha^R(M) = \alpha$, $\alpha^R(M-1) = \frac{\alpha}{2}$ and for $k = 3, \ldots, M$

$$\alpha^R(M-k+1) = \frac{1}{k}\left[\sum_{j=1}^{k-1}\alpha^j - \sum_{j=1}^{k-1}\binom{k}{j}(\alpha^R(M-j))^{k-j}\right]$$

The null hypothesis is rejected if all test statistics $CT_b^{(j)} \geq c_{\alpha^R(j)}$.

Another important practical problem when using the sub-sample based test is the choice of the block-length b. As Choi and Saikkonen (2005) we apply the so called *minimum volatility rule* proposed by Romano and Wolf (2001, p. 1297) in the simulations and empirical study. To be precise, we choose $b_{min} = 0.5\sqrt{T}$ and $b_{max} = 2.5\sqrt{T}$. Let us start the discussion with the Choi and Saikkonen (2005) test. For all $b \in [b_{min}, b_{max}]$ we compute the standard deviations of the test statistics over the five neighboring block sizes, i.e. for a block size b^*, we use the test statistics $CT_{b,max}$ for $b = b^*-2, b^*-1, b^*, b^*+1, b^*+2$ to compute the standard deviation of $CT_{b,max}$ as a function of b. The optimal block-length is then given by the value $b_{opt} \in [b_{min}+2, b_{max}-2]$ that leads to the smallest standard deviation. For the modified tests that involve all M test statistics we base the block-length selection on the following procedure. For each block-length $b_i \in [b_{min}, b_{max}]$ we compute the mean and standard deviation of the empirical distribution of the test statistics $\{CT_{b_i,i_1}, \ldots, CT_{b_i,i_M}\}$, which we denote by m_{b_i} and sd_{b_i}. The idea of the minimum volatility principle is now implemented by minimizing (again over five neighboring values of b) the change of the empirical distribution in terms of the first two moments. Hence we choose the block-length to minimize $vm_{b_i} = std(m_{b_i-2}, m_{b_i-1}, m_{b_i}, m_{b_i+1}, m_{b_i+2}) + std(sd_{b_i-2}, sd_{b_i-1}, sd_{b_i}, sd_{b_i+1}, sd_{b_i+2})$, with $std(\cdot)$ denoting the standard deviation.

3 Simulation Performance

In this section we present some simulation results to investigate the finite sample performance of the proposed estimators and tests. For assessing the performance of the estimators and size of the tests we use

$$y_t = c + \delta t + \beta_1 x_t + \beta_2 x_t^2 + u_t \ , \quad t = 1, \ldots, T \tag{18}$$

to generate $\{(y_t, x_t)\}_{t=1}^{T}$ for five different sample sizes $T \in \{50,\ 100,\ 200,\ 500,\ 1000\}$ and parameter values $c = \delta = 1$, $\beta_1 = 5$ and $\beta_2 = -0.3$. $\Delta X_t = v_t$ and u_t are generated as

$$\begin{aligned} (1 - \rho_1 L)u_t &= e_{1t} + \rho_2 e_{2t} \\ v_t &= e_{2t-1} + 0.5 e_{2t-2} \end{aligned}$$

with $(e_{1t}, e_{2t})' \sim \mathcal{N}(0, I_2)$. The two parameters ρ_1 and ρ_2 control the level of serial correlation in the error term and the level of endogeneity of the regressor, respectively, and they are set to take four different values $\rho_1,\ \rho_2 \in \{0.2,\ 0.4,\ 0.6,\ 0.8\}$.

To construct the bias correction terms, we need consistent estimators of the required long-run variances. The present results are based on the estimator proposed in Newey and West (1987) with bandwidth equal to $\lfloor 4(T/100)^{1/4} \rfloor$.[2] All simulation results are based on 5,000 replications and all computations have been performed in `MATLAB`. All tests results reported in this section are for a nominal level of $\alpha = 5\%$.

3.1 Performance of the Estimators

Tables 2 and 3 show the Monte Carlo means and standard deviations of the absolute values of the biases $|\hat{\beta}_1 - \beta_1|$ and $|\hat{\beta}_2 - \beta_2|$ from (18) for the OLS and FM-OLS estimators as a function of the sample size T and ρ_1 and ρ_2. The simulation results confirm the expectations concerning the relative performance of the OLS and FM-OLS estimators. The relative performance of the FM-OLS estimator compared to the OLS estimator improves for increasing serial correlation (i.e. increasing ρ_1) and increasing sample size. Increasing endogeneity (via increasing ρ_2) implies that the sample size required for which FM-OLS outperforms OLS is larger, e.g. for $\rho_2 = 0.7$ and β_1 the sample size should be about 100 or larger to result in smaller biases of FM-OLS. Generally, for β_1 corresponding

[2]Other kernels like the Parzen kernel and other bandwidth choices have also been investigated but do not lead to qualitatively different results.

to the integrated regressor, FM-OLS outperforms OLS already for many constellations for the smaller sample sizes. For the coefficient β_2 corresponding to the squared integrated process the sample size at which FM-OLS begins to outperform OLS has to be larger. Additional simulation results available upon request show that the discussed findings are qualitatively very robust with respect to the variance of u_t.

The sensitivity of the results with respect to the sample size T reflects the fact that the computation of the FM-OLS estimator requires non-parametric estimates of long-run covariances. Consequently, the finite sample performance of the FM-OLS estimator is dependent upon the properties of the long-run variance estimators, which sometimes perform poorly in small samples. However, the removal of the second order bias terms in the distribution is of prime importance for performing valid inference and the (potentially only small) modification to the point estimates is therefore not the only relevant aspect.

3.2 Performance of the Augmented and Auxiliary Regression Tests

In the discussion of the specification tests we consider here only the LM test and merely note that very similar results are also obtained with the Wald test. We compare two test versions to assess the importance of bias correction via FM-OLS estimation. One test statistic corresponds to the result given in Proposition 5 based on appropriate FM-OLS estimation of $\hat{u}_t^+$ on F_t. A second test statistic is computed by simply performing an OLS regression of $\hat{u}_t^+$ on F_t, with this test statistic suffering from biases even asymptotically.

The results in Table 4 are based on the FM-OLS residuals of (18) with $F_t = [x_t^3, x_t^4, q_t]'$ where q_t is generated as follows. First, a random walk $\tilde{q}_t = \sum_{j=1}^{t} \varepsilon_t$ with $\varepsilon_t \sim \mathcal{N}(0,1)$ and ε_t independent of e_{1t} and e_{2t} is generated. Then, this variable is orthogonalized with respect both y_t and all four regressors in (18) by taking the OLS residuals of the regression of $\tilde{q}_t$ on all these variables. These residuals are denoted q_t. In a variety of preliminary experiments this orthogonalization has improved the finite sample performance of the tests.

Bias correction has huge and important effects for the performance of the LM test, as can be seen in Table 4. The test based on OLS estimation of the auxiliary regression leads to rejections almost throughout, whereas the LM test based on FM-OLS estimation of the auxiliary regression shows very good performance already for the smallest sample size.

The large effect of bias correction on the LM test statistic is graphically displayed in Figure 1,

Table 2: Mean and standard deviation of $|\hat{\beta}_1 - \beta_1|$

	OLS				FM-OLS			
	$\rho_1 = 0.2$	$\rho_1 = 0.4$	$\rho_1 = 0.6$	$\rho_1 = 0.8$	$\rho_1 = 0.2$	$\rho_1 = 0.4$	$\rho_1 = 0.6$	$\rho_1 = 0.8$
$\rho_2 = 0.2$								
T= 50	0.119	0.167	0.330	1.045	0.162	0.207	0.369	1.165
	(0.112)	(0.153)	(0.288)	(0.904)	(0.181)	(0.209)	(0.331)	(0.990)
T= 100	0.057	0.085	0.185	0.717	0.070	0.098	0.198	0.757
	(0.053)	(0.079)	(0.169)	(0.629)	(0.075)	(0.096)	(0.182)	(0.663)
T= 200	0.028	0.044	0.103	0.444	0.032	0.048	0.106	0.457
	(0.027)	(0.041)	(0.093)	(0.386)	(0.032)	(0.045)	(0.096)	(0.395)
T= 500	0.011	0.018	0.044	0.207	0.012	0.019	0.044	0.207
	(0.010)	(0.017)	(0.040)	(0.180)	(0.012)	(0.017)	(0.040)	(0.181)
T=1000	0.006	0.009	0.023	0.111	0.006	0.009	0.022	0.109
	(0.005)	(0.008)	(0.020)	(0.098)	(0.005)	(0.008)	(0.020)	(0.097)
$\rho_2 = 0.4$								
T= 50	0.127	0.168	0.334	1.128	0.164	0.206	0.368	1.225
	(0.116)	(0.154)	(0.292)	(0.909)	(0.179)	(0.206)	(0.329)	(0.995)
T= 100	0.061	0.086	0.191	0.785	0.072	0.098	0.199	0.803
	(0.057)	(0.080)	(0.167)	(0.634)	(0.080)	(0.098)	(0.177)	(0.658)
T= 200	0.031	0.045	0.106	0.492	0.032	0.047	0.106	0.485
	(0.028)	(0.041)	(0.096)	(0.409)	(0.032)	(0.045)	(0.096)	(0.408)
T= 500	0.012	0.018	0.046	0.237	0.012	0.019	0.045	0.225
	(0.011)	(0.018)	(0.044)	(0.208)	(0.012)	(0.018)	(0.043)	(0.202)
T=1000	0.006	0.009	0.023	0.125	0.006	0.009	0.022	0.115
	(0.005)	(0.008)	(0.021)	(0.109)	(0.005)	(0.008)	(0.021)	(0.104)
$\rho_2 = 0.6$								
T= 50	0.140	0.171	0.345	1.281	0.175	0.214	0.376	1.353
	(0.127)	(0.156)	(0.295)	(0.995)	(0.200)	(0.226)	(0.344)	(1.071)
T= 100	0.070	0.090	0.202	0.893	0.075	0.101	0.203	0.881
	(0.063)	(0.082)	(0.175)	(0.682)	(0.079)	(0.097)	(0.180)	(0.704)
T= 200	0.035	0.047	0.115	0.568	0.033	0.048	0.110	0.535
	(0.031)	(0.043)	(0.101)	(0.449)	(0.033)	(0.046)	(0.098)	(0.439)
T= 500	0.014	0.019	0.049	0.275	0.012	0.019	0.045	0.244
	(0.012)	(0.017)	(0.043)	(0.221)	(0.011)	(0.017)	(0.041)	(0.206)
T=1000	0.007	0.009	0.025	0.143	0.006	0.009	0.022	0.122
	(0.006)	(0.008)	(0.022)	(0.120)	(0.005)	(0.008)	(0.020)	(0.108)
$\rho_2 = 0.8$								
T= 50	0.158	0.180	0.373	1.481	0.177	0.211	0.389	1.528
	(0.135)	(0.163)	(0.311)	(1.063)	(0.188)	(0.213)	(0.338)	(1.145)
T= 100	0.077	0.092	0.224	1.067	0.077	0.101	0.213	1.026
	(0.065)	(0.083)	(0.187)	(0.758)	(0.080)	(0.098)	(0.186)	(0.767)
T= 200	0.039	0.047	0.122	0.669	0.034	0.048	0.111	0.602
	(0.034)	(0.043)	(0.103)	(0.492)	(0.034)	(0.046)	(0.097)	(0.472)
T= 500	0.016	0.019	0.054	0.323	0.012	0.019	0.046	0.271
	(0.014)	(0.018)	(0.047)	(0.250)	(0.012)	(0.018)	(0.042)	(0.225)
T=1000	0.008	0.010	0.027	0.170	0.006	0.009	0.023	0.138
	(0.007)	(0.009)	(0.024)	(0.135)	(0.006)	(0.009)	(0.021)	(0.116)

Table 3: Mean and standard deviation of $|\hat{\beta}_2 - \beta_2|$

	OLS				FM-OLS			
	$\rho_1 = 0.2$	$\rho_1 = 0.4$	$\rho_1 = 0.6$	$\rho_1 = 0.8$	$\rho_1 = 0.2$	$\rho_1 = 0.4$	$\rho_1 = 0.6$	$\rho_1 = 0.8$
$\rho_2 = 0.2$								
T= 50	0.020	0.027	0.050	0.143	0.024	0.031	0.053	0.146
	(0.019)	(0.025)	(0.046)	(0.132)	(0.025)	(0.030)	(0.049)	(0.136)
T= 100	0.007	0.010	0.021	0.077	0.008	0.011	0.022	0.078
	(0.007)	(0.010)	(0.020)	(0.070)	(0.008)	(0.011)	(0.021)	(0.071)
T= 200	0.002	0.004	0.009	0.035	0.003	0.004	0.009	0.036
	(0.002)	(0.004)	(0.008)	(0.033)	(0.003)	(0.004)	(0.008)	(0.033)
T= 500	0.001	0.001	0.002	0.011	0.001	0.001	0.002	0.011
	(0.001)	(0.001)	(0.002)	(0.010)	(0.001)	(0.001)	(0.002)	(0.010)
T=1000	0.000	0.000	0.001	0.004	0.000	0.000	0.001	0.004
	(0.000)	(0.000)	(0.001)	(0.004)	(0.000)	(0.000)	(0.001)	(0.004)
$\rho_2 = 0.4$								
T= 50	0.020	0.028	0.052	0.149	0.025	0.032	0.054	0.152
	(0.020)	(0.026)	(0.048)	(0.139)	(0.026)	(0.032)	(0.051)	(0.142)
T= 100	0.007	0.010	0.021	0.077	0.008	0.011	0.021	0.077
	(0.007)	(0.010)	(0.020)	(0.069)	(0.008)	(0.011)	(0.020)	(0.069)
T= 200	0.002	0.004	0.009	0.036	0.003	0.004	0.009	0.036
	(0.002)	(0.004)	(0.008)	(0.033)	(0.003)	(0.004)	(0.008)	(0.033)
T= 500	0.001	0.001	0.002	0.011	0.001	0.001	0.002	0.011
	(0.001)	(0.001)	(0.002)	(0.011)	(0.001)	(0.001)	(0.002)	(0.011)
T=1000	0.000	0.000	0.001	0.004	0.000	0.000	0.001	0.004
	(0.000)	(0.000)	(0.001)	(0.004)	(0.000)	(0.000)	(0.001)	(0.004)
$\rho_2 = 0.6$								
T= 50	0.021	0.027	0.051	0.155	0.026	0.032	0.055	0.159
	(0.021)	(0.026)	(0.047)	(0.139)	(0.029)	(0.034)	(0.052)	(0.144)
T= 100	0.007	0.010	0.022	0.079	0.008	0.011	0.022	0.080
	(0.007)	(0.010)	(0.020)	(0.071)	(0.008)	(0.011)	(0.021)	(0.071)
T= 200	0.003	0.004	0.009	0.037	0.003	0.004	0.009	0.037
	(0.002)	(0.004)	(0.008)	(0.033)	(0.003)	(0.004)	(0.008)	(0.033)
T= 500	0.001	0.001	0.002	0.011	0.001	0.001	0.002	0.011
	(0.001)	(0.001)	(0.002)	(0.010)	(0.001)	(0.001)	(0.002)	(0.010)
T=1000	0.000	0.000	0.001	0.004	0.000	0.000	0.001	0.004
	(0.000)	(0.000)	(0.001)	(0.004)	(0.000)	(0.000)	(0.001)	(0.004)
$\rho_2 = 0.8$								
T= 50	0.022	0.028	0.053	0.155	0.027	0.033	0.056	0.158
	(0.021)	(0.026)	(0.048)	(0.144)	(0.028)	(0.032)	(0.052)	(0.147)
T= 100	0.007	0.010	0.022	0.081	0.008	0.011	0.022	0.081
	(0.007)	(0.010)	(0.021)	(0.073)	(0.009)	(0.011)	(0.021)	(0.073)
T= 200	0.003	0.004	0.009	0.038	0.003	0.004	0.009	0.038
	(0.003)	(0.004)	(0.008)	(0.035)	(0.003)	(0.004)	(0.009)	(0.034)
T= 500	0.001	0.001	0.003	0.012	0.001	0.001	0.002	0.012
	(0.001)	(0.001)	(0.002)	(0.011)	(0.001)	(0.001)	(0.002)	(0.011)
T=1000	0.000	0.000	0.001	0.005	0.000	0.000	0.001	0.004
	(0.000)	(0.000)	(0.001)	(0.005)	(0.000)	(0.000)	(0.001)	(0.004)

Table 4: Empirical Rejection Probabilities of the LM Test when H_0 is True

	Based on OLS				Based on FM-OLS			
	$\rho_1 = 0.2$	$\rho_1 = 0.4$	$\rho_1 = 0.6$	$\rho_1 = 0.8$	$\rho_1 = 0.2$	$\rho_1 = 0.4$	$\rho_1 = 0.6$	$\rho_1 = 0.8$
$\rho_2 = 0.2$								
T = 50	0.962	0.960	0.966	0.971	0.058	0.037	0.025	0.023
100	0.972	0.971	0.965	0.961	0.034	0.026	0.020	0.028
200	0.981	0.981	0.979	0.975	0.032	0.030	0.029	0.045
500	0.992	0.996	0.993	0.992	0.034	0.036	0.047	0.083
1000	0.998	0.998	0.999	0.996	0.030	0.034	0.047	0.089
$\rho_2 = 0.4$								
T = 50	0.972	0.978	0.970	0.965	0.073	0.045	0.020	0.011
100	0.987	0.985	0.981	0.970	0.060	0.036	0.033	0.030
200	0.997	0.994	0.992	0.978	0.056	0.037	0.029	0.039
500	1.000	0.999	0.998	0.994	0.059	0.043	0.037	0.079
1000	1.000	0.999	1.000	0.999	0.066	0.041	0.043	0.084
$\rho_2 = 0.6$								
T = 50	0.994	0.993	0.980	0.967	0.113	0.065	0.031	0.022
100	0.999	0.995	0.987	0.975	0.104	0.061	0.033	0.033
200	0.999	0.998	0.992	0.988	0.111	0.071	0.036	0.047
500	1.000	1.000	0.998	0.997	0.115	0.064	0.036	0.087
1000	1.000	1.000	1.000	0.999	0.114	0.058	0.036	0.086
$\rho_2 = 0.8$								
T = 50	0.998	0.995	0.990	0.970	0.158	0.091	0.037	0.022
100	0.999	0.999	0.991	0.984	0.152	0.100	0.036	0.030
200	1.000	1.000	0.997	0.994	0.155	0.084	0.037	0.044
500	1.000	0.999	1.000	0.998	0.190	0.098	0.048	0.079
1000	1.000	1.000	0.999	1.000	0.195	0.092	0.036	0.086

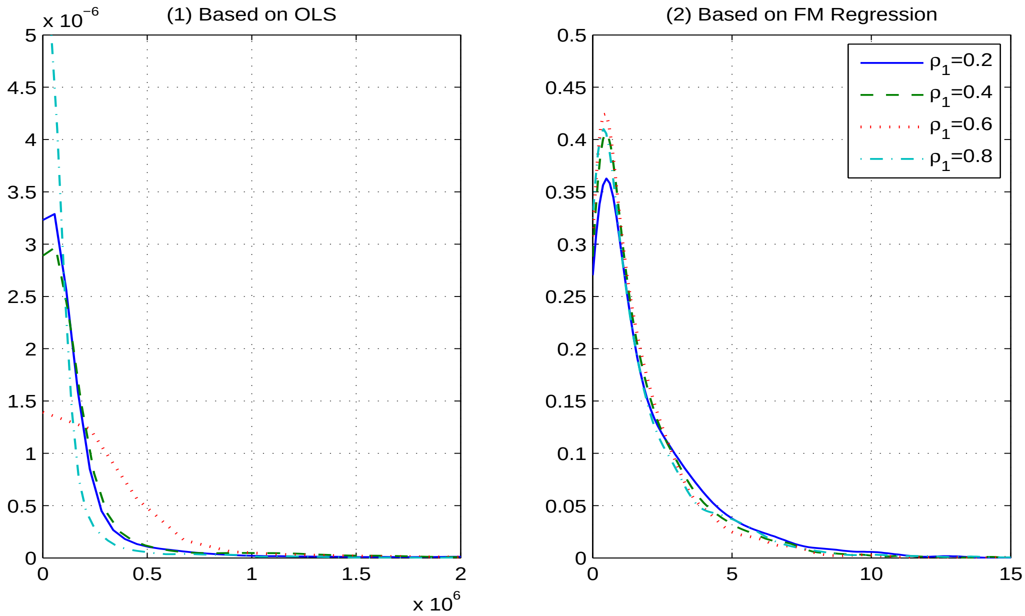
where we show kernel density estimates of the test statistics based on the 5,000 replications for the intermediate sample size $T = 500$. The kernel densities are based on using the Gaussian kernel with bandwidth chosen according to Silverman's rule of thumb. The left graphs display the OLS based statistics and the right graphs display the FM-OLS based statistics. Noting the scales for the left graphs makes clear why the OLS based statistics lead to rejections of the null hypotheses almost throughout. The FM-OLS based tests' densities in the right graphs remains remarkably unaffected by both serial correlation (upper panel) and endogeneity (lower panel).

For studying the power performance of the LM test, i.e. the empirical rejection probabilities for DGPs different from the estimated equation, we consider three alternative DGPs given by

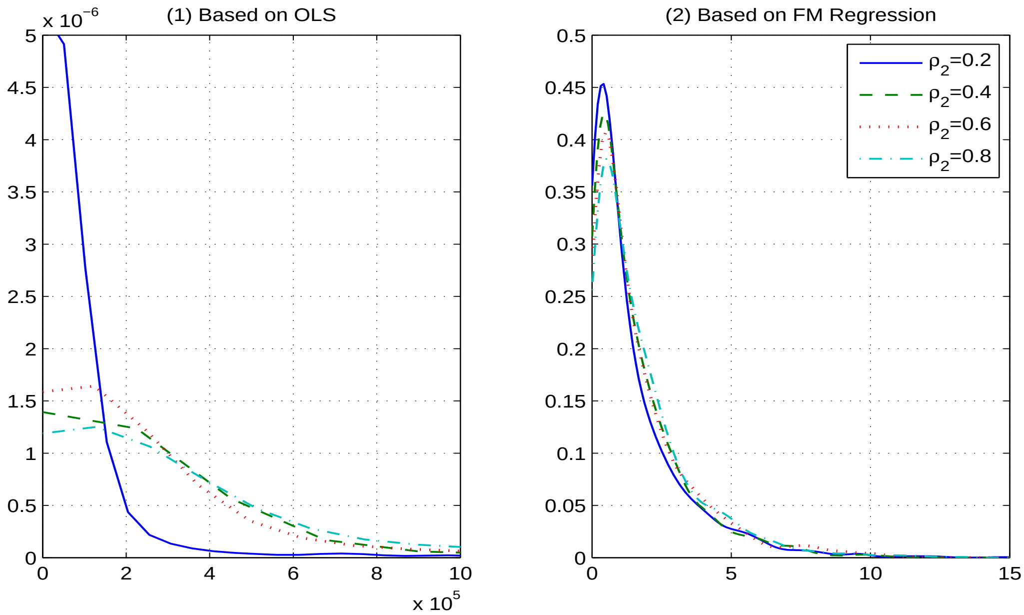
$$\begin{aligned}
(A) &\ : \ y_t = 1 + t - 15x_t + 5x_t^2 - 0.5x_t^3 + u_t \\
(B) &\ : \ y_t = 1 + t + 5x_t - 0.3x_t^2 + e_t\,, \ \ \text{where } e_t \text{ is an I(1) variable independent of } x_t \\
(C) &\ : \ y_t \text{ and } x_t \text{ are two independent I(1) variables}
\end{aligned}$$

To be precise in case (B) we use $e_t = \sum_{j=1}^{t} \varepsilon_t$ with $\varepsilon_t \sim \mathcal{N}(0,4)$ and in case (C) y_t is generated as $y_t = \sum_{j=1}^{t} \varepsilon_t$ with $\varepsilon_t \sim \mathcal{N}(0,1)$. The regressor x_t and for (A) u_t are throughout generated as described above. The three DGPs exemplify the main alternatives of interest. Alternative specification (A) covers the case of missing higher polynomial powers of the integrated regressor, alternative (B) corresponds to the case of a missing integrated regressor and alternative (C) corresponds to a spurious regression. The estimated equation is for all DGPs given by (18) and for alternative (A) again all combinations of ρ_1 and ρ_2 are considered.

The results in case of alternative (A) are displayed in Table 5 and for alternatives (B) and (C) the results are shown in Table 6. The results in Table 5 show that power is close to 1 or equal to 1 in case of missing higher polynomial powers for already the smallest considered sample size $T = 50$. This finding is robust with respect to the amount of serial correlation and endogeneity. However, power is much lower in case of alternatives (B) and (C). Alternatives (B) and (C) are not as well captured by the regressors F_t as in case of alternative (A), in which case the inclusion of F_t leads to a well-specified augmented regression. These results emphasize that the performance of the LM test depends, as in the stationary case, upon the relationship between the true but (in applications) unknown alternative and the auxiliary regressors collected in F_t. For empirical applications this means that one might consider to perform the LM test with several sets of auxiliary regressors, ignoring as is usual in empirical work all problems related to performing multiple inference.



(a) Effect of Serial Correlation: $T = 500$, $\rho_2 = 0.4$



(b) Effect of Endogeneity: $T = 500$, $\rho_1 = 0.4$

Figure 1: The Effects of Bias Correction on the Density of the LM Statistic

Table 5: Empirical Rejection Probabilities of the LM Test for Alternative (A)

	$\rho_1 = 0.2$	$\rho_1 = 0.4$	$\rho_1 = 0.6$	$\rho_2 = 0.8$
$\rho_2 = 0.2$				
T=50	0.999	0.999	0.998	0.957
100	1.000	1.000	1.000	1.000
200	1.000	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000
$\rho_2 = 0.4$				
T=50	1.000	1.000	0.999	0.954
100	1.000	1.000	1.000	1.000
200	1.000	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000
$\rho_2 = 0.6$				
T=50	1.000	1.000	0.998	0.956
100	1.000	1.000	1.000	1.000
200	1.000	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000
$\rho_2 = 0.8$				
T=50	1.000	1.000	0.999	0.950
100	1.000	1.000	1.000	1.000
200	1.000	1.000	1.000	1.000
500	1.000	1.000	1.000	1.000
1000	1.000	1.000	1.000	1.000

3.3 Performance of the KPSS Type Tests

We now briefly consider the performance of three sub-sample KPSS type tests, the modified Bonferroni procedures as proposed by Simes (1986) and Rom (1990) as well as the non-modified procedure used in Choi and Saikkonen (2005). Let us start with considering the behavior under the null hypothesis, reported in Table 7. A key observation is that the three versions of the sub-sample KPSS test perform very similarly under the null hypothesis. This implies that the modifications are not effective enough to completely offset the conservative behavior of the original Bonferroni procedure. The performance might be considered as not very good, but it has to be noted that the tests are performed on sub-samples that can be quite short, which of course leads to a deterioration of

Table 6: Empirical Rejection Probabilities of the LM Test for Alternatives (B) and (C)

	(B)	(C)
T=50	0.018	0.020
100	0.031	0.031
200	0.075	0.071
500	0.168	0.169
1000	0.286	0.285

finite sample performance. In particular this effect is observed also in the power experiments as illustrated in Tables 8 and 9. For all three considered alternatives the rejection probabilities are about 0.35 to 0.4 for $T = 500$ and around 0.75 for $T = 1000$. Note that the rejection probabilities are higher for the KPSS type tests than for the LM test for alternatives (B) and (C), which again reflects, from a comparative perspective, the dependence of the LM tests' performance upon the additional regressors F_t.

4 Environmental Kuznets Curves

As mentioned in the introduction, since the work of Grossmann and Krueger (1993, 1995) there has been a large body of both empirical and theoretical work studying the relationship between income and pollution or emissions measures. Brock and Taylor (2005) is an excellent recent discussion of EKCs which identifies three different mechanisms that link economic activity with pollution and emissions. These are the scale, composition and technique effects. For unchanging composition of output and unchanging technology emissions rise alongside with the scale of economic activity. For given scale and technique, emissions can rise or fall when the composition of output changes toward a more or less emissions intensive composition. Finally, emissions per unit of output, i.e. emissions intensity, can decrease by improvements in technology, e.g. via improved abatement technologies. Depending upon the relative importance of the three effects a monotonous, a U-shaped, an inverted U-shaped or in fact any pattern between per capita GDP and per capita emissions may emerge. Disentangling the relative importance of the three effects requires detailed structural modeling. However, the empirical EKC literature is typically less ambitious and focuses on reduced form modeling to address the issue whether the three mechanisms described *jointly operate* in a combination that leads to the emergence of an inverted U-shaped relationship.

Table 7: Empirical Rejection Probabilities of KPSS Type Tests when H_0 is True

	$\rho_1 = 0.2$			$\rho_1 = 0.4$			$\rho_1 = 0.6$			$\rho_1 = 0.8$		
Tests:	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
$\rho_2 = 0.2$												
T=50	0.061	0.057	0.071	0.032	0.030	0.039	0.012	0.011	0.018	0.017	0.016	0.033
100	0.030	0.030	0.029	0.019	0.019	0.021	0.040	0.040	0.042	0.138	0.135	0.149
200	0.010	0.010	0.010	0.013	0.013	0.012	0.041	0.039	0.045	0.191	0.183	0.197
500	0.007	0.007	0.008	0.015	0.015	0.019	0.058	0.057	0.058	0.304	0.295	0.297
1000	0.010	0.011	0.011	0.020	0.020	0.019	0.065	0.063	0.056	0.326	0.309	0.322
$\rho_2 = 0.4$												
T=50	0.057	0.055	0.064	0.030	0.028	0.034	0.012	0.011	0.017	0.015	0.014	0.030
100	0.024	0.024	0.025	0.017	0.017	0.019	0.042	0.042	0.042	0.148	0.144	0.155
200	0.008	0.008	0.011	0.015	0.015	0.017	0.044	0.044	0.049	0.201	0.196	0.213
500	0.006	0.006	0.008	0.016	0.016	0.018	0.064	0.062	0.065	0.292	0.283	0.294
1000	0.008	0.008	0.009	0.017	0.017	0.019	0.066	0.065	0.062	0.335	0.319	0.319
$\rho_2 = 0.6$												
T=50	0.046	0.044	0.057	0.027	0.025	0.030	0.011	0.010	0.016	0.012	0.010	0.028
100	0.019	0.018	0.022	0.016	0.016	0.016	0.038	0.036	0.041	0.143	0.140	0.148
200	0.005	0.005	0.008	0.011	0.010	0.014	0.042	0.041	0.042	0.199	0.191	0.209
500	0.004	0.004	0.005	0.014	0.014	0.014	0.058	0.057	0.058	0.298	0.287	0.294
1000	0.007	0.007	0.008	0.019	0.018	0.018	0.059	0.059	0.064	0.329	0.313	0.310
$\rho_2 = 0.8$												
T=50	0.038	0.035	0.043	0.023	0.021	0.028	0.011	0.010	0.019	0.015	0.013	0.032
100	0.016	0.016	0.016	0.014	0.013	0.015	0.039	0.039	0.042	0.153	0.148	0.152
200	0.003	0.003	0.005	0.011	0.011	0.011	0.038	0.037	0.041	0.198	0.191	0.210
500	0.005	0.005	0.004	0.014	0.014	0.015	0.052	0.052	0.057	0.289	0.275	0.268
1000	0.005	0.005	0.005	0.016	0.017	0.017	0.059	0.059	0.058	0.317	0.299	0.294

[Note] The three tests are: (1) Simes (1986), (2) Rom (1990) and (3) Choi and Saikkonen (2005)

Table 8: Empirical Rejection Probabilities of KPSS Type Tests for Alternative (A)

	$\rho_1 = 0.2$			$\rho_1 = 0.4$			$\rho_1 = 0.6$			$\rho_1 = 0.8$		
Tests:	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
$\rho_2 = 0.2$												
T=50	0.009	0.009	0.010	0.009	0.009	0.010	0.008	0.007	0.011	0.009	0.008	0.013
100	0.002	0.002	0.003	0.002	0.002	0.002	0.003	0.002	0.003	0.007	0.007	0.010
200	0.023	0.022	0.023	0.023	0.021	0.023	0.023	0.022	0.023	0.026	0.024	0.024
500	0.380	0.363	0.368	0.380	0.363	0.368	0.382	0.364	0.368	0.380	0.363	0.366
1000	0.762	0.739	0.733	0.762	0.738	0.734	0.763	0.738	0.734	0.764	0.740	0.734
$\rho_2 = 0.4$												
T=50	0.009	0.008	0.010	0.008	0.008	0.011	0.009	0.009	0.011	0.006	0.005	0.011
100	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.004	0.008	0.008	0.010
200	0.026	0.024	0.028	0.026	0.024	0.028	0.026	0.025	0.027	0.025	0.023	0.027
500	0.368	0.352	0.354	0.367	0.351	0.354	0.368	0.352	0.355	0.370	0.354	0.361
1000	0.764	0.741	0.735	0.764	0.741	0.735	0.763	0.742	0.735	0.764	0.742	0.735
$\rho_2 = 0.6$												
T=50	0.009	0.009	0.010	0.010	0.010	0.010	0.009	0.009	0.009	0.008	0.007	0.012
100	0.002	0.002	0.001	0.002	0.002	0.001	0.002	0.002	0.002	0.003	0.003	0.005
200	0.024	0.023	0.025	0.024	0.023	0.025	0.025	0.025	0.025	0.027	0.027	0.030
500	0.367	0.351	0.351	0.367	0.351	0.351	0.367	0.351	0.351	0.369	0.354	0.353
1000	0.754	0.733	0.729	0.754	0.733	0.729	0.754	0.733	0.729	0.754	0.732	0.730
$\rho_2 = 0.8$												
T=50	0.009	0.009	0.010	0.010	0.009	0.011	0.009	0.009	0.009	0.008	0.008	0.014
100	0.002	0.002	0.003	0.002	0.002	0.003	0.002	0.002	0.004	0.008	0.008	0.010
200	0.027	0.026	0.031	0.027	0.026	0.031	0.026	0.025	0.030	0.027	0.026	0.032
500	0.372	0.357	0.354	0.372	0.357	0.354	0.371	0.357	0.354	0.374	0.358	0.359
1000	0.758	0.734	0.730	0.758	0.734	0.730	0.757	0.734	0.729	0.756	0.734	0.729

[Note] The three tests are: (1) Simes (1986), (2) Rom (1990) and (3) Choi and Saikkonen (2005)

Table 9: Empirical Rejection Probabilities of KPSS Type Tests for Alternatives (B) and (C)

	(B)			(C)		
Tests:	(1)	(2)	(3)	(1)	(2)	(3)
T=50	0.009	0.009	0.011	0.006	0.005	0.011
100	0.003	0.003	0.004	0.008	0.008	0.010
200	0.026	0.025	0.027	0.025	0.023	0.027
500	0.368	0.352	0.355	0.370	0.354	0.361
1000	0.763	0.742	0.735	0.764	0.742	0.735

[Note] The three tests are: (1) Simes (1986), (2) Rom (1990) and (3) Choi and Saikkonen (2005)

Table 10: List of countries included in the empirical analysis. The sample range is 1870–2000 with the exception of New Zealand for which the sample ranges from 1878–2000 for CO_2 emissions.

Australia	Austria	Belgium	Canada
Denmark	Finland	France	Germany
Italy	Japan	Netherlands	New Zealand
Norway	Portugal	Spain	Sweden
Switzerland	UK	USA	

Typically, the following quadratic formulation in logarithms including a linear time trend is considered

$$e_t = c + \delta t + \beta_1 y_t + \beta_2 y_t^2 + u_t, \tag{19}$$

with e_t denoting the logarithm of per capita emissions and y_t the logarithm of per capita GDP.[3] A linear time trend is often included to allow e.g. for exogenous technical progress in abatement technologies. Studies like Grossmann and Krueger (1993, 1995) who investigate the relationship between pollutants and economic development with measurements taken at a disaggregated spatial level typically include further explanatory variables. For aggregate country-wide analysis, however, the above formulation (19) appears to be commonly used. Note that our theory allows for further deterministic, stationary and integrated regressors (including polynomial transformations of them) and hence can be used also for more detailed and less reduced form character modeling with additional explanatory variables as in the work of Grossman and Krueger.

In our application we consider annual data for GDP, carbon dioxide (CO_2) and sulfur dioxide (SO_2) emissions for 19 early industrialized countries listed in Table 10 over the period 1870–2000 and displayed in Appendix B in Figure 2. The GDP data are from the homepage of Angus Maddison (http://www.ggdc.net/maddison), the CO_2 emissions data have been downloaded from the homepage of the Carbon Dioxide Information Analysis Center (http://cdiac.ornl.gov) and the SO_2 emissions data are from Stern (2006). Note that for New Zealand the CO_2 emissions data only start in 1878.

Performing the battery of usual unit root tests on the log per capita GDP series leads to non-rejections of the unit root null hypothesis for all countries.[4] Note that the nonlinear cointegration

[3]The popular quadratic formulation appears to be due to Holtz-Eakin and Selden (1995), whereas Grossmann and Krueger (1995) use a cubic formulation, which we also investigate below.

[4]Detailed results for the usual unit root tests and specifications including only intercepts or both intercepts and linear trends are available upon request. The only borderline case is the US, where log per capita GDP over the

test results are qualitatively similar when we reduce the samples country specifically to exclude the break points for CO_2 emissions found for some countries in Lanne and Liski (2004) or in some countries during the two world wars.[5] This robustness of findings is probably driven by the fact that in those countries that are affected by e.g. the two world wars all variables are affected in a quite similar fashion, see the shaded areas in Figure 2 in Appendix B.

Before turning to formal statistical analysis of the relationship we consider the correlation between log per capita GDP on the one hand and log per capita emissions on the other in Figure 3, which is also relegated to Appendix B. The figure shows that for many countries there appears to be a relationship between income and both CO_2 and SO_2 emissions that is rising first and declining for higher income levels. However, in many cases this relationship appears to be asymmetric, especially for SO_2 emissions. For SO_2 emissions the quite rapid decline observed for many countries is at least partly due to changing legislation rather than only due to increasing income per se. Such an asymmetric behavior might better be captured by a third order polynomial than a second order polynomial as postulated in (19). For CO_2 emissions the relationship with income is monotonous for several countries for the sample period. Thus, even if there were an underlying quadratic relationship this will be hard to detect in these cases.

We present the result of the LM test and the three discussed versions of the KPSS type test for (19) in Table 11. In the application we use for the LM test, similarly to the simulations, $F_t = [y_t^3, y_t^4, q_t]$ with q_t generated and orthogonalized to the regressors as in Section 3. For our application testing by the LM test against F_t as just described is of particular relevance since as discussed for many countries the potential relationship appears to be asymmetric, which could be picked up by a third order polynomial of GDP. If we focus on the LM test results the null hypothesis is rejected for 11 (14) of the 19 countries at the 5% (10%) level for CO_2 emissions and for 14 (15) countries at the 5% (10%) level for SO_2 emissions. The three versions of the KPSS tests perform, unlike in the simulations, quite differently. The two modified versions à la Simes (1986) and Rom (1990) perform very similarly and lead to rejections of the null hypothesis for fewer countries than the Choi and Saikkonen (2005) version of the test. This might indicate that in the application the modifications correct to a certain extent the conservative behavior of the Choi and Saikkonen (2005) Bonferroni test, if the model under test is correct.

We focus on the results obtained by the LM test, which has shown good performance in case of

period 1870–2000 might also be considered as trend-stationary, depending upon test used.

[5]Note furthermore, in light of Remark 1, that we could incorporate broken deterministic trends in the analysis.

Table 11: Specification Test Results for the Quadratic Specification (19)

	CO_2				SO_2			
Tests:	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Australia	0.074	1(1)	1(1)	0.001	0.886	1(1)	1(1)	0.000
Austria	0.804	0(1)	0(1)	0.011	0.000	0(0)	0(0)	0.019
Belgium	0.238	1(1)	1(1)	0.000	0.001	1(1)	1(1)	0.004
Canada	0.000	0(0)	0(0)	0.032	0.891	0(1)	0(1)	0.012
Denmark	0.000	0(1)	0(1)	0.009	0.000	0(0)	0(0)	0.069
Finland	0.086	0(0)	0(0)	0.115	0.175	0(0)	0(0)	0.007
France	0.030	0(1)	0(1)	0.002	0.000	0(0)	0(0)	0.298
Germany	0.005	1(1)	1(1)	0.002	0.000	0(0)	0(0)	0.028
Italy	0.344	0(0)	0(0)	0.004	0.002	0(1)	0(1)	0.009
Japan	0.000	1(1)	0(1)	0.013	0.000	1(1)	1(1)	0.001
Netherlands	0.091	1(1)	0(1)	0.001	0.000	0(0)	0(0)	0.042
New Zealand	0.030	1(1)	1(1)	0.000	0.329	0(0)	0(0)	0.002
Norway	0.004	0(0)	0(0)	0.073	0.048	0(0)	0(0)	0.051
Portugal	0.004	1(1)	1(1)	0.002	0.004	1(1)	1(1)	0.002
Spain	0.411	0(1)	0(1)	0.006	0.098	1(1)	1(1)	0.023
Sweden	0.001	0(0)	0(0)	0.031	0.001	0(0)	0(0)	0.241
Switzerland	0.000	0(1)	0(1)	0.028	0.006	0(0)	0(0)	0.266
UK	0.199	1(1)	1(1)	0.000	0.000	1(1)	1(1)	0.001
USA	0.000	0(0)	0(0)	0.031	0.014	1(1)	1(1)	0.001

[Note] Test results for four specification tests of equation (19). The four tests are given by (1) LM test (p-value is reported), (2) Simes (1986) test where 1 indicates rejection and 0 non-rejection, (3) Rom (1990) test where 1 indicates rejection and 0 non-rejection and (4) Choi and Saikkonen (2005) test (p-value is reported). For tests (2) and (3) the numbers outside the brackets correspond to $\alpha = 5\%$ and the numbers in brackets to $\alpha = 10\%$.

alternatives well captured by the auxiliary regressors, whereas power is only increasingly slowly for the KPSS type tests. For CO_2 emissions the null hypothesis of a quadratic EKC is not rejected for Australia, Austria, Belgium, Finland, Italy, Netherlands, Spain and UK. In Figure 3 these are all countries in which the potential quadratic relationship is still primarily in the upward part with a tendency to flatten out at the highest income levels and with only few observations on the (potentially present) downward part. The exception here being the UK, where the scatterplot displays a rather wide inverted U-shape. Relatively similar observations can be made for SO_2 emissions, with non-rejections for Australia, Canada, Finland, New Zealand and Spain. For Canada and New Zealand the scatterplots indicate clearly an inverted U-shaped relationship.[6] For Australia, Finland and Spain the majority of observations is in the upward part of the potential quadratic relationship.

The estimation results for (19) are presented for all countries in Table 12 for CO_2 emissions and in Table 13 for SO_2 emissions. For completeness we also report the OLS estimates, whose reported standard errors have no formal statistical justification. For those countries for which the specification is not rejected the implied turning points, all corresponding to inverted U-shaped relationships, are with very few exceptions reasonable and within sample. The exceptions are for CO_2 emissions Italy with a turning point of almost 81,000 USD and for SO_2 emissions Australia with a turning point of the order 10^{32}. In both cases the scatterplots show that essentially all observations are in the upward part of the inverted U. Note also that not all coefficients corresponding to GDP squared are significantly different from 0. This happens e.g. for Australia in the case of SO_2 emissions, which is the only positive but not significant coefficient to squared GDP among the cases where the quadratic specification is not rejected. We observe some differences in the implied turning points between the OLS and FM-OLS estimates, where especially for CO_2 emissions typically FM-OLS estimation leads to smaller turning points. This is an important observation since part of the empirical EKC literature has observed unreasonably large turning points, in particular in case of CO_2 emissions.

Let us finally turn to the analysis of the *cubic* EKC, given by

$$e_t = c + \delta t + \beta_1 y_t + \beta_2 y_t^2 + \beta_3 y_t^3 + u_t, \tag{20}$$

where for the LM test, which we again focus on, we use $F_t = [y_t^4, q_t]$ as auxiliary regressors. For

[6] This graphical observation also applies to Norway, where the p-value of the LM test is 0.048, i.e. the null hypothesis is marginally rejected at the 5% level.

Table 12: Estimation Results for CO_2 Emissions for the Quadratic Specification (19)

	$\hat{\delta}$		$\hat{\beta}_1$		$\hat{\beta}_2$		Turning Point	
Country	OLS	FM-OLS	OLS	FM-OLS	OLS	FM-OLS	OLS	FM-OLS
Australia	0.035	0.037	5.836	10.884	-0.363	-0.645	3,113	4,609
	(7.338)	(8.046)	(1.644)	(2.483)	(-1.942)	(-2.756)		
Austria	-0.014	-0.014	7.124	10.652	-0.335	-0.535	40,924	21,178
	(-2.439)	(-3.034)	(2.079)	(3.424)	(-1.800)	(-3.069)		
Belgium	-0.007	-0.005	13.689	19.890	-0.726	-1.072	12,435	10,691
	(-3.871)	(-3.264)	(9.326)	(14.576)	(-9.270)	(-14.727)		
Canada	0.004	0.004	15.758	18.065	-0.854	-0.982	10,187	9,852
	(0.749)	(0.523)	(8.105)	(10.148)	(-7.730)	(-10.430)		
Denmark	-0.004	0.001	11.083	12.768	-0.554	-0.661	22,216	15,734
	(-0.801)	(0.241)	(9.502)	(10.408)	(-9.987)	(-11.202)		
Finland	-0.037	-0.029	18.184	16.964	-0.870	-0.816	34,472	32,495
	(-3.279)	(-3.165)	(8.975)	(7.907)	(-8.850)	(-7.363)		
France	-0.001	0.000	9.946	13.292	-0.532	-0.722	11,517	9,895
	(-0.899)	(0.092)	(12.400)	(15.287)	(-12.094)	(-15.336)		
Germany	-0.002	-0.001	9.956	14.071	-0.540	-0.774	10,164	8,880
	(-0.846)	(-0.255)	(8.561)	(9.275)	(-7.969)	(-9.251)		
Italy	-0.005	0.002	7.785	5.846	-0.355	-0.259	58,677	80,794
	(-0.416)	(0.269)	(2.274)	(2.317)	(-2.059)	(-1.890)		
Japan	0.021	0.013	12.735	14.721	-0.720	-0.823	6,922	7,652
	(2.144)	(0.888)	(4.365)	(4.085)	(-4.116)	(-4.228)		
Netherlands	0.003	0.007	9.749	13.077	-0.507	-0.702	14,921	11,113
	(2.017)	(4.235)	(9.610)	(10.709)	(-9.170)	(-10.570)		
New Zealand	0.003	0.023	2.058	-1.840	-0.080	0.065	391,430	1.3×10^6
	(0.419)	(4.588)	(0.573)	(-0.448)	(-0.403)	(0.290)		
Norway	0.043	0.071	-5.599	-10.302	0.273	0.479	28,804	46,640
	(2.723)	(5.918)	(-2.040)	(-3.557)	(2.180)	(3.348)		
Portugal	0.022	0.024	-3.989	-3.707	0.262	0.244	2,013	2,017
	(1.977)	(3.823)	(-1.732)	(-1.353)	(2.213)	(1.558)		
Spain	0.010	0.008	6.253	10.537	-0.320	-0.569	17,419	10,415
	(2.526)	(2.799)	(3.026)	(4.581)	(-2.825)	(-4.328)		
Sweden	-0.005	0.002	12.588	14.337	-0.654	-0.769	15,128	11,190
	(-0.664)	(0.164)	(5.181)	(5.368)	(-4.893)	(-5.581)		
Switzerland	-0.008	-0.014	5.849	11.745	-0.251	-0.559	116,343	36,243
	(-1.079)	(-1.748)	(2.802)	(5.501)	(-2.095)	(-5.101)		
UK	-0.005	-0.006	8.097	16.996	-0.427	-0.914	13,080	10,904
	(-3.064)	(-3.569)	(6.900)	(16.014)	(-7.170)	(-16.899)		
USA	0.000	-0.002	11.547	16.502	-0.603	-0.865	14,390	13,879
	(0.082)	(-0.373)	(8.587)	(12.148)	(-8.654)	(-12.679)		

[Note] Estimation results for (19). The sample period is $1870 - 2000$ with the exception of New Zealand where the sample starts in 1878. The t-statistics, in parentheses, are computed using the HAC estimator of Newey and West (1987) for the OLS estimator and for the FM-OLS estimator the t-statistics are computed as described in the text. The turning points are computed as $\exp\left(-\frac{\hat{\beta}_1}{2\hat{\beta}_2}\right)$. In bold we indicate the countries for which the null hypothesis of correct specification has not been rejected at the 5% level using the LM test as described in the text.

Table 13: Estimation Results for SO_2 Emissions for the Quadratic Specification (19)

	$\hat{\delta}$		$\hat{\beta}_1$		$\hat{\beta}_2$		Turning Point	
Country	OLS	FM-OLS	OLS	FM-OLS	OLS	FM-OLS	OLS	FM-OLS
Australia	0.018	0.020	-1.593	-0.487	0.070	0.003	92,885	3.7×10^{32}
	(3.126)	(3.875)	(-0.419)	(-0.100)	(0.356)	(0.012)		
Austria	-0.022	-0.025	23.024	25.844	-1.276	-1.427	8,279	8,582
	(-4.202)	(-5.735)	(5.703)	(8.511)	(-5.734)	(-8.389)		
Belgium	-0.008	-0.010	32.356	35.757	-1.781	-1.965	8,817	8,931
	(-2.667)	(-2.739)	(10.525)	(12.035)	(-10.832)	(-12.401)		
Canada	0.010	0.011	23.530	24.533	-1.347	-1.406	6,195	6,167
	(3.036)	(2.855)	(30.730)	(30.065)	(-32.072)	(-32.562)		
Denmark	-0.055	-0.062	29.075	30.716	-1.442	-1.516	23,888	25,109
	(-4.243)	(-4.755)	(6.219)	(9.259)	(-6.158)	(-9.505)		
Finland	0.005	-0.004	26.933	28.370	-1.501	-1.564	7,870	8,688
	(0.263)	(-0.326)	(6.050)	(9.121)	(-6.767)	(-9.728)		
France	-0.000	-0.002	18.923	19.737	-1.047	-1.088	8,421	8,665
	(-0.099)	(-0.464)	(7.546)	(9.064)	(-7.776)	(-9.227)		
Germany	-0.004	-0.007	16.737	20.010	-0.943	-1.120	7,140	7,585
	(-0.562)	(-1.171)	(3.580)	(6.450)	(-3.620)	(-6.547)		
Italy	-0.029	-0.029	14.101	12.855	-0.680	-0.608	31,918	38,755
	(-2.351)	(-3.511)	(3.138)	(4.039)	(-2.871)	(-3.524)		
Japan	-0.001	-0.006	17.880	18.498	-1.042	-1.070	5,309	5,674
	(-0.175)	(-0.690)	(10.654)	(8.540)	(-10.931)	(-9.146)		
Netherlands	-0.002	-0.005	30.880	34.213	-1.721	-1.897	7,873	8,240
	(-0.555)	(-1.395)	(8.038)	(12.442)	(-8.357)	(-12.689)		
New Zealand	0.010	0.010	23.437	26.999	-1.361	-1.560	5,477	5,719
	(1.559)	(2.075)	(6.625)	(6.155)	(-7.229)	(-6.534)		
Norway	-0.012	-0.006	23.289	21.713	-1.290	-1.214	8,303	7,636
	(-0.848)	(-0.643)	(9.552)	(9.259)	(-11.687)	(-10.480)		
Portugal	0.012	0.009	0.698	1.814	0.002	-0.060	0	3.8×10^{6}
	(2.514)	(2.296)	(0.449)	(0.996)	(0.024)	(-0.576)		
Spain	0.003	0.002	11.185	12.762	-0.624	-0.714	7,843	7,615
	(1.205)	(0.640)	(5.885)	(5.100)	(-5.728)	(-4.987)		
Sweden	-0.039	-0.042	38.474	39.419	-2.072	-2.117	10,778	11,030
	(-2.824)	(-3.914)	(10.555)	(13.148)	(-11.280)	(-13.690)		
Switzerland	-0.067	-0.081	28.611	31.887	-1.395	-1.539	28,336	31,564
	(-5.013)	(-7.771)	(9.403)	(11.627)	(-9.321)	(-10.926)		
UK	-0.009	-0.012	26.912	32.034	-1.474	-1.747	9,209	9,564
	(-2.086)	(-2.437)	(6.114)	(10.127)	(-6.347)	(-10.840)		
USA	-0.012	-0.013	17.820	20.049	-0.947	-1.063	12,221	12,437
	(-3.663)	(-2.308)	(14.809)	(12.861)	(-14.478)	(-13.579)		

[Note] Estimation results for (19). The sample period is $1870 - 2000$. The t-statistics, in parentheses, are computed using the HAC estimator of Newey and West (1987) for the OLS estimator and for the FM-OLS estimator the t-statistics are computed as described in the text. The turning points are computed as $\exp\left(-\frac{\hat{\beta}_1}{2\hat{\beta}_2}\right)$. In bold we indicate the countries for which the null hypothesis of correct specification has not been rejected at the 5% level using the LM test as described in the text.

CO_2 emissions the null hypothesis of correct specification is now not rejected – in addition to those countries for which the quadratic specification has not been rejected – for New Zealand, Switzerland and the USA. For the Netherlands the cubic specification is rejected, despite the non-rejection of the quadratic specification. For SO_2 emissions non-rejections occur in addition to those for the quadratic specification for Belgium, the Netherlands and the USA.

The Choi and Saikkonen (2005) test again leads to rejections for most countries for both pollutants. The two modified versions lead to identical test results for both pollutants for all countries and to a much smaller number of rejections than the Choi and Saikkonen (2005) test.

Table 14: Specification Test Results for the Cubic Specification (20)

	CO2				SO2			
Tests:	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Australia	0.133	0(1)	0(1)	0.000	0.726	1(1)	1(1)	0.000
Austria	0.718	0(0)	0(0)	0.084	0.004	0(0)	0(0)	0.023
Belgium	0.338	1(1)	1(1)	0.000	0.714	0(1)	0(1)	0.012
Canada	0.002	0(0)	0(0)	0.042	0.892	0(0)	0(0)	0.096
Denmark	0.000	0(1)	0(1)	0.001	0.000	0(0)	0(0)	0.037
Finland	0.881	0(0)	0(0)	0.093	0.216	0(0)	0(0)	0.014
France	0.027	1(1)	1(1)	0.000	0.002	0(0)	0(0)	0.052
Germany	0.002	1(1)	1(1)	0.002	0.000	0(1)	0(1)	0.015
Italy	0.193	0(0)	0(0)	0.002	0.042	1(1)	1(1)	0.011
Japan	0.000	0(0)	0(0)	0.129	0.001	0(0)	0(0)	0.007
Netherlands	0.038	1(1)	1(1)	0.008	0.749	0(0)	0(0)	0.020
New Zealand	0.099	1(1)	1(1)	0.000	0.815	0(0)	0(0)	0.046
Norway	0.014	0(0)	0(0)	0.027	0.016	0(0)	0(0)	0.078
Portugal	0.001	1(1)	1(1)	0.004	0.001	1(1)	1(1)	0.003
Spain	0.506	0(0)	0(0)	0.008	0.097	1(1)	1(1)	0.019
Sweden	0.001	0(0)	0(0)	0.249	0.000	0(0)	0(0)	0.277
Switzerland	0.379	1(1)	1(1)	0.001	0.006	0(0)	0(0)	0.468
UK	0.445	1(1)	1(1)	0.000	0.037	1(1)	1(1)	0.001
USA	0.759	1(1)	1(1)	0.000	0.631	1(1)	1(1)	0.075

[Note] Test results for four specification tests of equation (20). The four tests are given by (1) LM test (p-value is reported), (2) Simes (1986) test where 1 indicates rejection and 0 non-rejection, (3) Rom (1990) test where 1 indicates rejection and 0 non-rejection and (4) Choi and Saikkonen (2005) test (p-value is reported). For tests (2) and (3) the numbers outside the brackets correspond to $\alpha = 5\%$ and the numbers in brackets to $\alpha = 10\%$.

A comparison with the scatterplots displayed in Figure 3 shows that for the countries for which the quadratic specification is rejected but the cubic specification is not rejected by the LM test the scatterplots display asymmetrically shaped relatively smooth relationships with a substantial amount of observations in the upward and downward parts of the relationship. Thus, it appears that the estimation results and subsequently also the test results hinge upon the availability of enough observations spread out over the range of the potential relationship.

The estimation results for the cubic specification are displayed in Tables 15 and 16 in Appendix B. For SO_2 emissions the turning points for Belgium, the Netherlands and the USA are at reasonable in-sample values, for CO_2 emissions turning points are present only for New Zealand, whereas there are no turning points for Switzerland and the USA, which stems from the fact that for these countries complex conjugate roots occur.

5 Summary and Conclusions

The fast growing literature on environmental Kuznets curves has to date ignored the econometric implications of the presence of polynomial powers of integrated regressors in cointegrating regressions. Regressions involving nonlinear transformations of integrated regressors typically require different statistical analysis than standard linear cointegrating regressions. Based on the work of Park and Phillips (1999; 2001) we develop estimation and testing theory for regressions including stationary regressors, deterministic regressors and integrated regressors and their integer powers. This setup is clearly a rather special formulation of a nonlinear relationship but it offers some advantages. First, this setup leads to relationships that are linear in the parameters which implies that modified OLS estimation techniques will suffice, avoiding the need to resort to nonlinear estimation techniques that arises in more general formulations. Note that regressions involving also cross-products of the powers of the regressors can be studied by slightly modifying the results presented in this paper. We, however, believe that unless the application one has in mind leads one to consider such cross-products as being important the more parsimonious formulation without cross-products we focus on in this paper is a potentially good starting point for nonlinear cointegration analysis, not only for EKC analysis.

It turns out that the OLS estimator of the coefficients in regression equations considered in this paper behaves in many respects similar to the OLS estimator in linear cointegrating relationship as studied for instance in Phillips and Hansen (1990). The OLS estimator is consistent, but its limiting distribution is in general contaminated by second-order bias terms, which render valid inference infeasible. As in the linear case an FM-OLS estimator with a limiting distribution that is free of second order bias terms can be constructed. Consequently, similarly to the linear case the FM-OLS estimator is the basis for χ^2-inference on certain classes of hypotheses, including e.g. t-statistics.

We also consider specification and cointegration testing in detail by pursuing two avenues, one based on augmented respectively auxiliary regressions and the other one on studying KPSS type tests. Specification analysis is based on augmented regressions, using the Wald principle, or on auxiliary regressions, using the Lagrange Multiplier principle. By considering as additional regressors additional deterministic components, higher order powers of the integrated regressors and additional integrated regressors and their integer powers allows to stay, with appropriate extensions, in the estimation framework considered in the outset of the paper. As discussed in the paper the tests, leading to asymptotic χ^2-inference, are consistent against several forms of misspecification of the original equation (i.e. cointegration only in the augmented setup or no cointegration even in the augmented setup). The performance of the tests depends upon the relationship between the additional regressors and the unknown alternative. A direct test for cointegration is based on testing stationarity of the FM-OLS residuals of the original regression. This extends the work of Kwiatkowski, Phillips, Schmidt, and Shin (1992) to our nonlinear setup. The asymptotic distribution of this test statistic depends upon nuisance parameters related to the specification of the equation, i.e. the included deterministic components and the integrated variables and their powers. Following Choi and Saikkonen (2005) we present a sub-sample version of the test that has a nuisance parameter free limiting distribution. We also study in detail test procedures based on adjusted Bonferroni bounds, as considered by Hommel (1988), Simes (1986) and Rom (1990), to utilize the information from *all* the sub-sample statistics.

We investigate the performance of the proposed estimator and tests by means of a small simulation study. As expected it turns out that the point estimates obtained from OLS and FM-OLS estimation do not differ drastically, given that both estimators are consistent. However, typically the FM-OLS estimator leads to slightly smaller biases. Bias-correction by FM-OLS estimation is, as also expected, vital for inference, compare again Table 4 and Figure 1. The augmented respectively auxiliary regression based tests exhibit very good size performance, with their power performance depending upon the relationship between the additional regressors and the alternative considered. The sub-sample KPSS type test performance is suffering, as expected, from the fact that only short sub-samples are used. This affects also the power performance, which is not as satisfactory as for the Wald and LM tests but is on the other hand quite independent of the alternative considered. These findings suggest that the choice of test depends upon whether the researcher has a particular alternative in mind. In case one has a particular alternative in mind the corresponding variables should be used as additional regressors in the Wald or LM test. If one has no particular alternative in mind one might start by considering the KPSS type test, maybe in conjunction with the Wald and LM tests performed for several sets of additional regressors. In this respect it might be fruitful to study in some more detail the performance of nonlinear cointegration testing by adding superfluous deterministic trends as originally advocated for unit root and cointegration testing by Park and Choi (1988) and Park (1990).

Finally, we apply the developed methods to study the relationship between log per capita GDP and log per capita CO_2 and SO_2 emissions for 19 early industrialized countries over the period 1870–2000. We find evidence for the prevalence of a quadratic EKC in about half of the countries, where for a few countries cointegration is not rejected only in the cubic specification. We find evidence for an EKC in more countries for CO_2 emissions which might partly be due to the fact that the relationship between GDP and SO_2 emissions appears to be less smooth and symmetric then the relationship between GDP and CO_2 emissions. The implied turning points based on the FM-OLS estimates are with very few exceptions at reasonable in-sample values.

Acknowledgements

The helpful comments of Robert Kunst, of seminar participants in Hamburg and of conference participants at the 2nd International Workshop in Computational and Financial Econometrics in Neuchâtel are gratefully acknowledged. The usual disclaimer applies.

References

Bertinelli, L. and E. Strobl (2005). The Environmental Kuznets Curve Semi-Parametrically Revisited. *Economics Letters* **88**, 350–357.

Bradford, D., R. Fender, S.H. Shore, and M. Wagner (2005). The Environmental Kuznets Curve: Exploring a Fresh Specification. *Contributions to Economic Analysis and Policy* **4**, No. 1, Article 5. Berkeley Electronic Press.

Brock, W.A. and M.S. Taylor (2005). Economic Growth and the Environment: A Review of Theory and Empirics. In: Aghion, P. and S. Durlauf (Eds.), *Handbook of Economic Growth.* North-Holland, Amsterdam.

Chang, Y., J.Y. Park, and P.C.B. Phillips (2001). Nonlinear Econometric Models with Cointegrated and Deterministically Trending Regressors. *Econometrics Journal* **4**, 1–36.

Choi, I. and P. Saikkonen (2005). Tests for Nonlinear Cointegration. Mimeo.

de Benedictis, L.F. and D.E.A. Giles (1998). Diagnostic Testing in Econometrics: Variable Addition, RESET and Fourier Approximations. In Ullah, A. and D.E.A. Giles (Eds.) *Handbook of Applied Economic Statistics*, Marcel Dekker, New York, 383–417.

de Jong, R. (2002). Nonlinear Estimators with Integrated Regressors but without Exogeneity. Mimeo.

Grossmann, G.M. and A.B. Krueger (1993). Environmental Impacts of a North American Free Trade Agreement. In Garber, P. (Ed.) *The Mexico-US Free Trade Agreement*, 13–56, MIT Press, Cambridge.

Grossmann, G.M. and A.B. Krueger (1995). Economic Growth and the Environment. *Quarterly Journal of Economics* **110**, 353–377.

Holtz-Eakin, D. and T.M. Selden (1995). Stoking the Fires? CO_2 Emissions and Economic Growth. *Journal of Public Economics* **57**, 85–101.

Hommel, G.A. (1986). Multiple Test Procedures for Arbitrary Dependence Structures. *Metrika* **33**, 321–336.

Hommel, G.A. (1988). A Stagewise Rejective Multiplicative Test Procedure Based on a Modified Bonferroni Test. *Biometrika* **75**, 383–386.

Hong, S.H. and P.C.B. Phillips (2008). Testing Linearity in Cointegrating Relations with an Application to Purchasing Power Parity. Forthcoming in *Journal of Business and Economic Statistics.*

Kuznets, S. (1955). Economic Growth and Income Inequality. *American Economic Review* **45**, 1–28.

Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, and Y. Shin (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**, 159–178.

Lanne, M. and M. Liski (2004). Trends and Breaks in Per-Capita Carbon Dioxide Emissions, 1870–2028. *Energy Journal* **25**, 41–65.

Lee, T.H., H. White, and C.W.J. Granger (1993). Testing for Neglected Nonlinearity in Time Series Models: A Comparison of Neural Network Methods and Alternative Tests. *Journal of Econometrics* **56**, 269–290.

Millimet, D.L., J.A. List, and T. Stengos (2003). The Environmental Kuznets Curve: Real Progress or Misspecified Models? *Review of Economics and Statistics* **85**, 1038–1047.

Müller-Fürstenberger G. and M. Wagner (2007). Exploring the Environmental Kuznets Hypothesis: Theoretical and Econometric Problems. *Ecological Economics* **62**, 648–660.

Newey, W. and K. West (1987). A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator. *Econometrica* **50**, 703–708.

Park, J. Y. (1990). Testing for Unit Roots and Cointegration by Variable Addition. In Fomby, T. B. and G. F. Rhodes (Eds.) *Advances in Econometrics, Vol. 8: Co-Integration, Spurious Regression, and Unit Roots*. JAI Press, Greenwich, 107–133.

Park, J.Y. (1992). Canonical Cointegrating Regressions. *Econometrica* **60**, 119–143.

Park, J. Y. and B. Choi (1988). A New Approach to Testing for a Unit Root. CAE Working Paper 88-23, Cornell University.

Park, J.Y. and P.C.B. Phillips (1999). Asymptotics for Nonlinear Transformations of Integrated Time Series. *Econometric Theory* **15**, 269–298.

Park, J.Y. and P.C.B. Phillips (2001). Nonlinear Regressions with Integrated Time Series. *Econometrica* **69**, 117–161.

Phillips, P.C.B. (1983). Best Uniform and Modified Padé Approximants to Probability Densities in Econometrics. In Hildenbrand, W. (Ed.) *Advances in Econometrics*, Cambridge University Press, Cambridge, 123–167.

Phillips, P.C.B. and B.E. Hansen (1990). Statistical Inference in Instrumental Variables Regression with I(1) Processes. *Review of Economic Studies* **57**, 99–125.

Phillips, P.C.B and V. Solo (1992). Asymptotics for Linear Processes. *The Annals of Statistics* **20**, 971–1001.

Ramsey, J.B. (1969). Tests for Specification Errors in Classical Linear Least-Squares Regression Analysis. *Journal of the Royal Statistical Society B* **31**, 350–371.

Rom, D.M. (1990). A Sequentially Rejective Test Procedure Based on a Modified Bonferroni Inequality. *Biometrika* **77**, 663–665.

Romano, J.P. and M. Wolf (2001). Subsampling Intervals in Autoregressive Models with Linear Time Trend. *Econometrica* **69**, 1283–1314.

Saikkonen, P. (1991). Asymptotically Efficient Estimation of Cointegrating Regressions. *Econometric Theory* **7**, 1–21.

Schmalensee, R., T.M. Stoker, and R.A. Judson (1998). World Carbon Dioxide Emissions: 1950–2050. *Review of Economics and Statistics* **80**, 15–27.

Simes, R.J. (1986). An Improved Bonferroni Procedure for Multiple Tests of Significance. *Biometrika* **73**, 751–754.

Stern, D.I. (2004). The Rise and Fall of the Environmental Kuznets Curve. *World Development* **32**, 1419–1439.

Stern, D.I. (2006). Reversal of the Trend in Global Anthropogenic Sulfur Emissions. *Global Environmental Change* **16**, 207–220.

Wagner, M. (2008). The Carbon Kuznets Curve: A Cloudy Picture Emitted by Bad Econometrics? *Energy and Resource Economics* **30**, 388–408.

Yandle, B., M. Bjattarai, and M. Vijayaraghavan (2004). Environmental Kuznets Curves: A Review of Findings, Methods, and Policy Implications. Research Study 02.1 update, PERC.

Appendix A: Proofs

Proof of Proposition 1

The OLS estimator $\hat{\theta}$ in (2) can be written as

$$\begin{aligned} G^{-1}(\hat{\theta}-\theta) &= (GZ'ZG)^{-1}GZ'u \\ &= \begin{bmatrix} G_w w' w G_w & G_w w' D G_D & G_w w' X G_X \\ G_D D' w G_w & G_D D' D G_D & G_D D' X G_X \\ G_X X' w G_w & G_X X' D G_D & G_X X' X G_X \end{bmatrix}^{-1} \begin{bmatrix} G_w w' u \\ G_D D' u \\ G_X X' u \end{bmatrix} \\ &\Rightarrow \begin{bmatrix} \Sigma_{ww} & 0 & 0 \\ 0 & \int DD' & \int D\mathbf{B}_v' \\ 0 & \int \mathbf{B}_v D' & \int \mathbf{B}_v \mathbf{B}_v' \end{bmatrix}^{-1} \begin{bmatrix} N_{wu} \\ \int D dB_u \\ \int \mathbf{B}_v dB_u + M \end{bmatrix}. \end{aligned}$$

The stated convergence results for the entries in the respective matrices are straightforward to establish. For all components in $GZ'ZG$ that involve integrated processes we can use the results of Chang, Park, and Phillips (2001, Lemma 5). Due to the assumptions made for $G_w W' W G_w$ a law of large numbers applies. Thus it follows that

$$\begin{bmatrix} G_w w' w G_w & G_w w' D G_D & G_w w' X G_X \\ G_D D' w G_w & G_D D' D G_D & G_D D' X G_X \\ G_X X' w G_w & G_X X' D G_D & G_X X' X G_X \end{bmatrix} \Rightarrow \begin{bmatrix} \Sigma_{ww} & 0 & 0 \\ 0 & \int DD' & \int D\mathbf{B}_v' \\ 0 & \int \mathbf{B}_v D' & \int \mathbf{B}_v \mathbf{B}_v' \end{bmatrix}$$

It is the asymptotic orthogonality between the stationary processes and both the deterministic and integrated processes that leads to the structure of the limiting matrix above. Using the definitions of $\tilde{D}$ and $\tilde{\mathbf{B}}_v$ the inverse can be written as

$$\begin{aligned} &\begin{bmatrix} \Sigma_{ww} & 0 & 0 \\ 0 & \int DD' & \int D\mathbf{B}_v' \\ 0 & \int \mathbf{B}_v D' & \int \mathbf{B}_v \mathbf{B}_v' \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \Sigma_{ww}^{-1} & 0 & 0 \\ 0 & \left(\int \tilde{D}\tilde{D}'\right)^{-1} & -\left(\int \tilde{D}\tilde{D}'\right)^{-1} \int D\mathbf{B}_v' \left(\int \mathbf{B}_v \mathbf{B}_v'\right)^{-1} \\ 0 & -\left(\int \tilde{\mathbf{B}}_v \tilde{\mathbf{B}}_v'\right)^{-1} \int \mathbf{B}_v D' \left(\int DD'\right)^{-1} & \left(\int \tilde{\mathbf{B}}_v \tilde{\mathbf{B}}_v'\right)^{-1} \end{bmatrix}. \qquad (21) \end{aligned}$$

Let us now turn to the three blocks in $GZ'u$. The convergence result for the first block has been established e.g. in Park (1992, Lemma A.1(a)). For the quantity $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} w_t u_t$ a central limit theorem can be established using similar arguments as in Phillips and Solo (1992, Theorem 3.8 and Lemma 5.9). The variance of the limiting normal distribution depends upon the coefficients $c_{w,j}$, $d_{u,j}$, $\Sigma_{\eta\eta}$ and σ_ζ^2 and can be derived explicitly by cumbersome computations which are available upon request. For notational brevity we denote this random variable by N_{wu}. A typical entry of the third block is given by $T^{-\frac{k+1}{2}}\sum_{t=1}^{T} x_{jt}^k u_t$. For this quantity de Jong (2002, Lemma 1) has established convergence to $\int B_{v_j}^k dB_u + k\Lambda_{v_j u} \int B_{v_j}^{k-1}$. The result then follows by stacking the coordinates and using the definition of M.

The result as given in the proposition for $G_w^{-1}(\hat{\theta}_w - \theta_w)$ now follows from straightforward multiplications. The results for $G_D^{-1}(\hat{\theta}_D - \theta_D)$ and $G_X^{-1}(\hat{\theta}_X - \theta_X)$ follow from using again the definitions of $\tilde{D}$, $\tilde{\mathbf{B}}_{\mathbf{v}}$, $B_{u.v} = B_u - \Omega_{uv}\Omega_{vv}^{-1}B_v$ and the relationship

$$\int \tilde{\mathbf{B}}_v dB_u \quad = \quad \int \tilde{\mathbf{B}}_v dB_{u.v} - \int \tilde{\mathbf{B}}_v dB_v' \Omega_{vv}^{-1}\Omega_{vu},$$

which completes the proof of the proposition.

Proof of Proposition 2

From the definition of $\hat{\theta}^+$ in the main text we obtain

$$G^{-1}\left(\hat{\theta}^+ - \theta\right) = \left(GZ'ZG\right)^{-1}\left(GZ'u^+ - GA^*\right),$$

with $u_t^+ := u_t - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_t$ and $u^+ = [u_1^+, \ldots, u_T^+]'$. The limit of $(GZ'ZG)^{-1}$ has already been analyzed in the proof of Proposition 1. Therefore we only need to investigate the second matrix above, the cross-products and the correction terms, with blocks given by

$$\left(\begin{array}{c} G_w w' u^+ \\ G_D D' u^+ \\ G_X X' u^+ - G_X M^* \end{array} \right).$$

With consistent long-run variance estimators $\hat{\Omega}_{uv}$ and $\hat{\Omega}_{vv}$ we obtain for $0 \leq r \leq 1$ that

$$\begin{aligned} \frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor rT \rfloor} u_t^+ \quad &= \quad \frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor rT \rfloor} u_t - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor rT \rfloor} v_t \\ &\Rightarrow \quad B_u(r) - \Omega_{uv}\Omega_{vv}^{-1}B_v(r) = B_{u.v}(r) \end{aligned}$$

and $G_D D' u^+ \Rightarrow \int D dB_{u.v}$. Convergence of $G_w w' u^+ \Rightarrow N_{wu.v}$, with $N_{wu.v}$ normally distributed, follows from similar arguments as the convergence of $\frac{1}{\sqrt{T}}\sum_{t=1}^T w_t u_t$ established in Proposition 1. Let us now consider a typical entry of the third block

$$\begin{aligned} T^{-\frac{k+1}{2}}\sum_{t=1}^T x_{jt}^k u_t^+ \quad &= \quad T^{-\frac{k+1}{2}}\sum_{t=1}^T x_{jt}^k u_t - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1} T^{-\frac{k+1}{2}}\sum_{t=1}^T x_{jt}^k v_t \qquad (22) \\ &\Rightarrow \quad \int B_{v_j}^k dB_u + k\Lambda_{v_j u}\int B_{v_j}^{k-1} - \Omega_{uv}\Omega_{vv}^{-1}\left(\int B_{v_j}^k dB_v + k\Lambda_{vv_j}\int B_{v_j}^{k-1}\right) \\ &\Rightarrow \quad \int B_{v_j}^k dB_{u.v} + k\left(\Lambda_{v_j u} - \Omega_{uv}\Omega_{vv}^{-1}\Lambda_{vv_j}\right)\int B_{v_j}^{k-1}, \end{aligned}$$

where the result concerning $T^{-\frac{k+1}{2}}\sum_{t=1}^T x_{jt}^k u_t$ has already been used in Proposition 1 and the result for $T^{-\frac{k+1}{2}}\sum_{t=1}^T x_{jt}^k v_t$ is contained in the proof of Lemma 4 of Hong and Phillips (2008).

Given the definition of the bias correction term M^*, it follows that the third block converges to $\int \mathbf{B}_v dB_{u.v}$. The proposition then follows from similar calculations as in Proposition 1 by using again the definitions of $\tilde{D}$ and $\tilde{\mathbf{B}}_v$.

Proof of Proposition 3

Denote with $G_R \in \mathbb{R}^{s \times s}$ a weighting matrix capturing the convergence rates of the coefficients involved in the respective hypotheses formulated by the rows of R. With either of the two formulated conditions on the set of testable hypotheses it holds that

$$G_R^{-1} R = R G_N^{-1},$$

with $G_N \in \mathbb{R}^{(q+1+p)\times(q+1+p)}$ the weighting matrix defined in Remark 2. This implies $G_R^{-1} R \left(\hat{\theta}_N^{+} - \theta_N\right) = R G_N^{-1} \left(\hat{\theta}_N^{+} - \theta_N\right)$ with $G_N^{-1} \left(\hat{\theta}_N^{+} - \theta_N\right) \Rightarrow (JJ')^{-1} \int J dB_{u.v}$. Therefore, under the null hypothesis we have

$$\begin{aligned} T &= \left[G_R^{-1} R \left(\hat{\theta}_N^{+} - \theta_N\right)\right]' \left[\hat{\omega}_{u.v} G_R^{-1} R \left(Z_N' Z_N\right)^{-1} R' G_R^{-1}\right]^{-1} \left[G_R^{-1} R \left(\hat{\theta}_N^{+} - \theta_N\right)\right] \\ &= \left[R G_N^{-1} \left(\hat{\theta}_N^{+} - \theta_N\right)\right]' \left[\hat{\omega}_{u.v} R G_N^{-1} \begin{pmatrix} D'D & D'X \\ X'D & X'X \end{pmatrix}^{-1} G_N^{-1} R'\right]^{-1} \left[R G_N^{-1} \left(\hat{\theta}_N^{+} - \theta_N\right)\right] \\ &\Rightarrow \left[R \left(\int JJ'\right)^{-1} \int J dB_{u.v}\right]' \left[\omega_{u.v} R \left(\int JJ'\right)^{-1} R'\right]^{-1} \left[R \left(\int JJ'\right)^{-1} \int J dB_{u.v}\right], \end{aligned}$$

which concludes the proof since the distribution in the above line is given by a quadratic form of a mean zero normal mixture with variance given by the expression inverted in the middle.

Proof of Proposition 4

Clearly the result in this proposition is a special case of a hypothesis covered by Proposition 3 which leads due to the form of the restrictions to a particularly simple form of the test statistic. In the augmented regression (10) the restriction $\theta_F = 0$ corresponds to

$$\left[\begin{array}{cc} 0 & I_b \end{array}\right] \left[\begin{array}{c} \theta \\ \theta_F \end{array}\right] = 0.$$

This immediately implies that $\left(R \left[\begin{array}{cc} Z_N' Z_N & Z_N' F \\ F' Z_N & F'F \end{array}\right]^{-1} R'\right)^{-1} = \tilde{F}_N' \tilde{F}_N$, with $R = \left[\begin{array}{cc} 0 & I_b \end{array}\right]$ and $\tilde{F}_N$ as defined in the main text.

Proof of Proposition 5

The proof is in many respects similar to the proofs of Propositions 1 and 2 in showing that the correction terms given in the proposition (asymptotically) correct the second order bias terms of the OLS estimator. Let us start by defining the weighting matrices corresponding to the additional regressors F, i.e. let $G_F(T) := \text{diag}\left[G_{\bar{X}_1}(T), \dots, G_{\bar{X}_m}(T), G_{Q_1}(T), \dots, G_{Q_k}(T)\right]$, with

$$G_{\bar{X}_j}(T) := \begin{bmatrix} T^{-\frac{p_j+2}{2}} & & \\ & \ddots & \\ & & T^{-\frac{p_j+r_j+1}{2}} \end{bmatrix}, \quad G_{Q_i}(T) := \begin{bmatrix} T^{-1} & & \\ & \ddots & \\ & & T^{-\frac{s_i+1}{2}} \end{bmatrix}.$$

The OLS estimate of $\theta_{\tilde{F}}$ of (10) is given by

$$\begin{aligned} \hat{\theta}_{\tilde{F}} &:= (\tilde{F}'\tilde{F})^{-1}\tilde{F}'\hat{u}^+ \\ &= G_F(G_F\tilde{F}'\tilde{F}G_F)^{-1}G_F\tilde{F}'\hat{u}^+. \end{aligned}$$

Next define the stacked Brownian motion vectors corresponding to the higher order polynomial powers of x_{jt} and to the polynomial powers of q_{it}. For t such that $\lim_{T\to\infty} t/T = r$ we consider

$$\begin{aligned} \lim_{T\to\infty} \sqrt{T}G_{\bar{X}_j}(T)\bar{X}_{jt} &= \lim_{T\to\infty} \begin{pmatrix} T^{-\frac{p_j+1}{2}} & & \\ & \ddots & \\ & & T^{-\frac{p_j+r_j}{2}} \end{pmatrix} \begin{pmatrix} x_{jt}^{p_j} \\ \vdots \\ x_{jt}^{p_j+r_j} \end{pmatrix} = \begin{pmatrix} B_{v_j}^{p_j} \\ \vdots \\ B_{v_j}^{p_j+r_j} \end{pmatrix} =: \mathbf{B}_{v_j}^F(r), \\ \lim_{T\to\infty} \sqrt{T}G_{Q_i}(T)Q_{it} &= \lim_{T\to\infty} \begin{pmatrix} T^{-\frac{1}{2}} & & \\ & \ddots & \\ & & T^{-\frac{s_i}{2}} \end{pmatrix} \begin{pmatrix} q_{it} \\ \vdots \\ q_{it}^{s_i} \end{pmatrix} = \begin{pmatrix} B_{v_i^*} \\ \vdots \\ B_{v_i^*}^{s_i} \end{pmatrix} =: \mathbf{B}_{v_i^*}^F(r) \end{aligned}$$

stacked together as $\mathbf{B}^F(r) := [\mathbf{B}_{v_1}^F(r)', \dots, \mathbf{B}_{v_m}^F(r)', \mathbf{B}_{v_1^*}^F(r)', \dots, \mathbf{B}_{v_k^*}^F(r)']'$. To establish the result three things have to be considered, namely the asymptotic behavior of $(G_F\tilde{F}'\tilde{F}G_F)^{-1}$, the asymptotic behavior of $G_F\tilde{F}'\hat{u}^+$ and that the proposed correction factors annihilate the bias terms arising in the limit of $G_F\tilde{F}'\hat{u}^+$. Since $\tilde{F}$ corresponds to the regression residuals of F on Z, it is not surprising that in the limiting quantities correspondingly adjusted Brownian motions will appear. These can be written in various forms with the most convenient one for this proposition given by

$$\tilde{\mathbf{B}}^F := \mathbf{B}^F - \int \mathbf{B}^F\tilde{D}' \left(\int \tilde{D}\tilde{D}'\right)^{-1} D - \int \mathbf{B}^F\tilde{\mathbf{B}}_v' \left(\int \tilde{\mathbf{B}}_v\tilde{\mathbf{B}}_v'\right)^{-1} \mathbf{B}_v. \tag{23}$$

With the just defined quantity it follows by appropriate rearrangements of terms that $(G_F\tilde{F}'\tilde{F}G_F)^{-1} \Rightarrow \left(\int \tilde{\mathbf{B}}^F\tilde{\mathbf{B}}^{F'}\right)^{-1}$. Let us next consider

$$\begin{aligned} G_F\tilde{F}'\hat{u}^+ &= G_F\tilde{F}'\left(u^+ - Z(\hat{\theta}^+ - \theta)\right) \\ &= G_F\tilde{F}'u^+ \\ &= G_FF'u^+ - G_FF'ZG(GZ'ZG)^{-1}GZ'u^+, \end{aligned}$$

with the first equality following from $\hat{u}^+ = u^+ - Z(\hat{\theta}^+ - \theta)$, the second from $\tilde{F}'Z = 0$ and the third from the definition of $\tilde{F}$. Next we use from Proposition 2 that $G^{-1}(\hat{\theta}^+ - \theta) + (GZ'ZG)^{-1}GA^* = (GZ'ZG)^{-1}GZ'u^+$ to obtain

$$G_F\tilde{F}'\hat{u}^+ \quad = \quad G_F F'u^+ - G_F F'ZG\left(G^{-1}(\hat{\theta}^+ - \theta) - (GZ'ZG)^{-1}GA^*\right). \tag{24}$$

Several of the above terms have been already analyzed in Proposition 2. Using similar arguments as in the derivation of equation (22) in Proposition 2 it follows that $G_F F'u^+ \Rightarrow \int \mathbf{B}^F dB_{u.v} + O^F + M^F$, with $M^F := [M_1^{F'}, \dots, M_m^{F'}, M_{m+1}^{F'}, \dots, M_{m+k}^{F'}]'$ and

$$M_j^F := \Lambda_{v_j u}^+ \begin{bmatrix} p_j \int B_{v_j}^{p_j - 1} \\ \vdots \\ (p_j + r_j) \int B_{v_j}^{p_j + r_j - 1} \end{bmatrix}, \; M_{m+i}^F := \Lambda_{v_i^* u}^+ \begin{bmatrix} 1 \\ 2\int B_{v_i^*} \\ \vdots \\ s_i \int B_{v_i^*}^{s_i - 1} \end{bmatrix}.$$

and with

$$O^F = \Omega_{u\tilde{v}}\Omega_{\tilde{v}\tilde{v}}^{-1}\int \mathbf{B}^F d\mathbf{B}_{\tilde{v}} - \Omega_{uv}\Omega_{vv}^{-1}\int \mathbf{B}^F d\mathbf{B}_v.$$

This modified results stems from the fact that $B_{u.v}$ is not orthogonal to $\mathbf{B}_{v^*}$, which requires some additional modifications to arrive at an orthogonalized process. Thus, define $B_{u.\tilde{v}} := B_u - \Omega_{u\tilde{v}}\Omega_{\tilde{v}\tilde{v}}^{-1}B_{\tilde{v}}$, which is by construction orthogonal to both v and v^*. Using this quantity the relevant cross-moments, i.e. the quantities corresponding to (22), can be written as

$$\begin{aligned} T^{-\frac{k+1}{2}}\sum_{t=1}^T x_{jt}^k u_t^+ \quad &= \quad T^{-\frac{k+1}{2}}\sum_{t=1}^T x_{jt}^k u_t - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}T^{-\frac{k+1}{2}}\sum_{t=1}^T x_{jt}^k v_t \\ &\Rightarrow \int B_{v_j}^k dB_u + k\Lambda_{uv_j}\int B_{v_j}^{k-1} - \Omega_{uv}\Omega_{vv}^{-1}\left(\int B_{v_j}^k dB_v + k\Lambda_{vv_j}\int B_{v_j}^{k-1}\right) \\ &\Rightarrow \int B_{v_j}^k dB_{u.\tilde{v}} + k\left(\Lambda_{v_j u} - \Omega_{uv}\Omega_{vv}^{-1}\Lambda_{vv_j}\right)\int B_{v_j}^{k-1}, \\ &\quad +\Omega_{u\tilde{v}}\Omega_{\tilde{v}\tilde{v}}^{-1}\int B_{v_j}^k dB_{\tilde{v}} - \Omega_{uv}\Omega_{vv}^{-1}\int B_{v_j}^k dB_v \end{aligned} \tag{25}$$

and

$$\begin{aligned} T^{-\frac{k+1}{2}}\sum_{t=1}^T q_{it}^k u_t^+ \quad &= \quad T^{-\frac{k+1}{2}}\sum_{t=1}^T q_{it}^k u_t - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}T^{-\frac{k+1}{2}}\sum_{t=1}^T q_{it}^k v_t \\ &\Rightarrow \int B_{v_i^*}^k dB_u + k\Lambda_{uv_i^*}\int B_{v_i^*}^{k-1} - \Omega_{uv}\Omega_{vv}^{-1}\left(\int B_{v_i^*}^k dB_v + k\Lambda_{vv_i^*}\int B_{v_i^*}^{k-1}\right) \\ &\Rightarrow \int B_{v_i^*}^k dB_{u.\tilde{v}} + k\left(\Lambda_{uv_i^*} - \Omega_{uv}\Omega_{vv}^{-1}\Lambda_{vv_i^*}\right)\int B_{v_i^*}^{k-1} \\ &\quad +\Omega_{u\tilde{v}}\Omega_{\tilde{v}\tilde{v}}^{-1}\int B_{v_i^*}^k dB_{\tilde{v}} - \Omega_{uv}\Omega_{vv}^{-1}\int B_{v_i^*}^k dB_v, \end{aligned} \tag{26}$$

which by stacking leads to the additional bias correction term O^F.

Straightforward computations show that $G_F F' Z G \Rightarrow \left[\int \mathbf{B}^F D' \; 0 \; \int \mathbf{B}^F \mathbf{B}_v'\right]$. The terms in the brackets on the right hand side of (24) have all been considered already in Proposition 2. Tedious and lengthy computations to arrive at a useful formulation of the limiting quantity lead to

$$G_F \tilde{F}' \hat{u}^+ \;\; \Rightarrow \;\; \int \tilde{\mathbf{B}}^F dB_{u.v} + O^F + M^F - \left(\mathbf{B}^F \tilde{\mathbf{B}}_v'\right)\left(\tilde{\mathbf{B}}_v \tilde{\mathbf{B}}_v'\right)^{-1} M.$$

We are thus left to show that the correction terms stated in the formulation of the proposition converge to the same limits. Therefore consider the components of

$$G_F \hat{\theta}_{\tilde{F}} = (G_F \tilde{F}' \tilde{F} G_F)^{-1}(G_F \tilde{F}' \hat{u}^+ - G_F O^{F*} - G_F M^{F*} + G_F k^{F*} M^*)$$

in some detail. With consistent variance estimators it holds that $G_F O^{F*} \Rightarrow O^F$ and $G_F M^{F*} \Rightarrow M^F$. Let us next consider the last term

$$G_F k^{F*} M^* \;\; = \;\; G_F F' \tilde{X} G_X (G_X \tilde{X}' \tilde{X} G_X)^{-1} G_X M^*.$$

From this representation we immediately obtain $G_F F' \tilde{X} G_X \Rightarrow \int \mathbf{B}^F \tilde{\mathbf{B}}_v'$, $(G_X \tilde{X}' \tilde{X} G_X)^{-1} \Rightarrow \left(\int \tilde{\mathbf{B}}_v \tilde{\mathbf{B}}_v'\right)^{-1}$ and $G_X M^* \Rightarrow M$. This implies that the asymptotic distribution of the estimator $\hat{\theta}_{\tilde{F}}^+$ is under the null hypothesis given by

$$G_F^{-1} \hat{\theta}_{\tilde{F}}^+ \Rightarrow \left(\int \tilde{\mathbf{B}}^F \tilde{\mathbf{B}}^{F\prime}\right)^{-1} \int \tilde{\mathbf{B}}^F dB_{u.\tilde{v}}. \tag{27}$$

The asymptotic χ^2 distribution of the test statistic now follows from the same arguments as used in Propositions 3 and 4.

Proof of Proposition 6

By definition we have $\hat{u}_t^+ = u_t^+ - Z_t'\left(\hat{\theta}^+ - \theta\right)$. From the proof of Proposition 2 we already know that $\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor rT \rfloor} u_t^+ \Rightarrow B_{u.v}(r)$ and thus we only need to investigate the second term

$$\begin{aligned}
\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor rT \rfloor} Z_t' G G^{-1}\left(\hat{\theta}^+ - \theta\right) \;\; &= \;\; \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor rT \rfloor} \left[w_t' G_w \;\; D_t' G_D \;\; X_t' G_X\right]\right) \begin{bmatrix} G_w^{-1}(\hat{\theta}_w^+ - \theta_w) \\ G_D^{-1}(\hat{\theta}_D^+ - \theta_D) \\ G_X^{-1}(\hat{\theta}_X^+ - \theta_X) \end{bmatrix} \\
&\Rightarrow \;\; \left[\; r\mathbb{E}(w_t)' \quad \int_0^r D' \quad \int_0^r \mathbf{B}_v' \;\right] \begin{bmatrix} \Sigma_{ww}^{-1} N_{wu.v} \\ \left[\int \tilde{D}\tilde{D}'\right]^{-1} \int \tilde{D} dB_{u.v} \\ \left[\int \tilde{\mathbf{B}}_v \tilde{\mathbf{B}}_v'\right]^{-1} \int \tilde{\mathbf{B}}_v dB_{u.v} \end{bmatrix}.
\end{aligned}$$

Since by assumption $\mathbb{E}(w_t) = 0$ it follows that

$$\begin{aligned}
\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor rT \rfloor} \hat{u}_t^+ \;\; &\Rightarrow \;\; B_{u.v}(r) - \int_0^r D' \left[\int \tilde{D}\tilde{D}'\right]^{-1} \int \tilde{D} dB_{u.v} - \int_0^r \mathbf{B}_v' \left[\int \tilde{\mathbf{B}}_v \tilde{\mathbf{B}}_v'\right]^{-1} \int \tilde{\mathbf{B}}_v dB_{u.v} \\
&= \;\; B_{u.v}^*(r).
\end{aligned}$$

This implies due to the assumption of consistency of $\hat{\omega}_{u.v}$ that

$$CT \Rightarrow \frac{1}{\omega_{u.v}} \int (B^*_{u.v})^2.$$

Proof of Proposition 7

Let $0 \leq r \leq 1$ and $i \leq t = \lfloor br \rfloor + i - 1 \leq i + b - 1$. Similar to the proof of Proposition 6 a functional central limit theorem applies for the sub-sample of residuals and we obtain

$$\frac{1}{\sqrt{b}} \sum_{j=i}^{t} \hat{u}_j^+ = \frac{1}{\sqrt{b}} \sum_{j=i}^{t} u_j^+ + \left(\frac{1}{\sqrt{b}} \sum_{j=i}^{t} Z_j' G(b) \right) \left(G(b)^{-1} G(T) \right) \left(G(T)^{-1} (\hat{\theta}^+ - \theta) \right) \tag{28}$$

Also similar to the proof of Proposition 6 one can show, since $b \to \infty$ and $\sqrt{\frac{b}{T}} \to 0$, that $\lim_{T\to\infty} \frac{1}{\sqrt{b}} \sum_{j=i}^{t} u_j^+ = B_{u.v}(r)$. The first and the third bracketed terms composing the product on the right hand side above, i.e. $\left(\frac{1}{\sqrt{b}} \sum_{j=i}^{t} Z_j' G(b) \right)$ and $\left(G(T)^{-1} (\hat{\theta}^+ - \theta) \right)$, converge in distribution. The term in the middle is of order $O\left(\sqrt{\frac{b}{T}}\right)$, which implies that the right hand side product term in (28) is $O_p\left(\sqrt{\frac{b}{T}}\right)$. Therefore, since by assumption $\sqrt{\frac{b}{T}} \to 0$, we have established that $\frac{1}{\sqrt{b}} \sum_{j=i}^{t} \hat{u}_j^+ \Rightarrow B_{u.v}(r)$. The result then follows from the assumption of consistency of $\hat{\omega}_{u.v} \to \omega_{u.v}$.

Appendix B: Additional Material for the Empirical Study

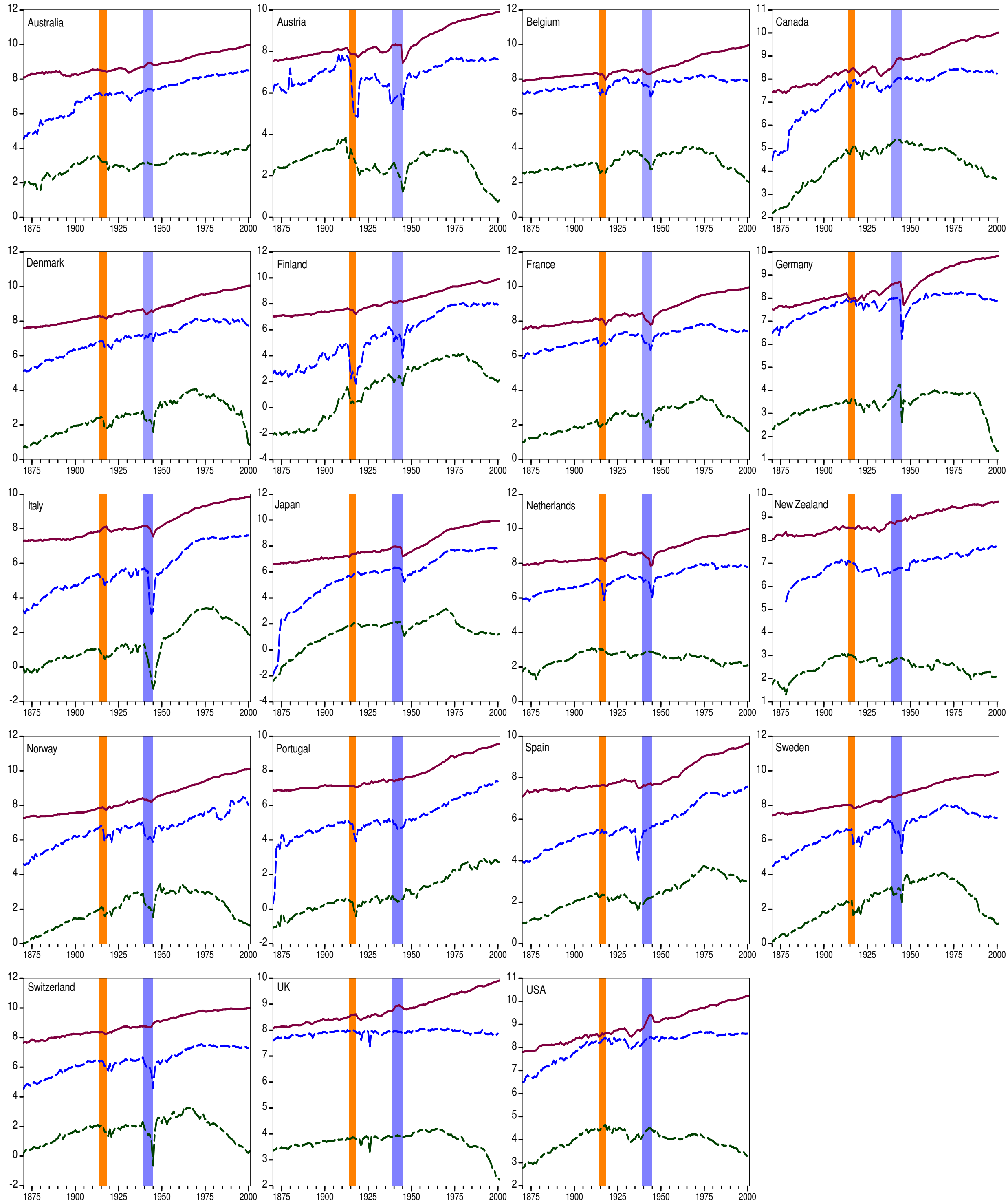


Figure 2: Time series plots of the three variables (log per capita) GDP, CO_2 and SO_2 emissions. The solid lines display GDP, the upper dashed lines CO_2 emissions and the lower dashed lines SO_2 emissions. The two shaded areas indicate world war I (1914–1918) and world war II (1939–1945).

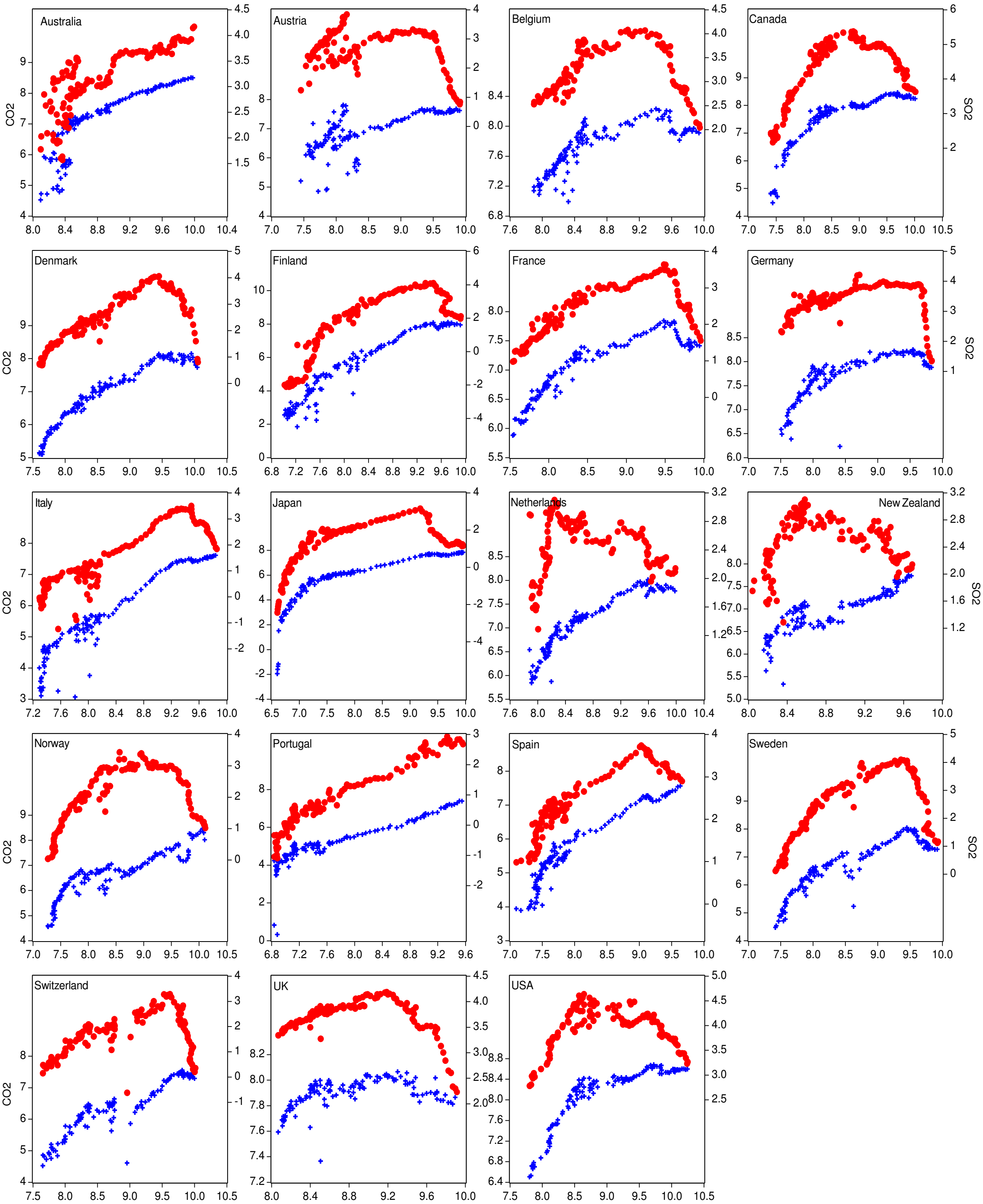


Figure 3: Scatterplots of log per capita GDP versus log per capita CO_2 and SO_2 emissions. The lines with the + symbols correspond to CO_2 emissions (left scale) and the line with the full circle symbols correspond to SO_2 emissions (right scale).

Table 15: Estimation Results for CO_2 Emissions for the Cubic Specification (20)

Country	$\hat{\delta}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	Turning Points	
Australia	0.035	388.196	-42.527	1.547	5,463	16,615
	(7.690)	(4.084)	(-4.031)	(3.969)		
Austria	-0.015	152.685	-16.998	0.634	–	–
	(-3.305)	(2.753)	(-2.646)	(2.563)		
Belgium	-0.012	268.446	-28.950	1.041	–	–
	(-7.510)	(11.661)	(-11.224)	(10.810)		
Canada	0.000	200.525	-22.102	0.812	7,467	10,229
	(0.064)	(10.472)	(-9.987)	(9.548)		
Denmark	-0.006	86.127	-8.966	0.314	–	–
	(-1.229)	(5.414)	(-4.988)	(4.622)		
Finland	-0.052	97.963	-10.250	0.368	–	–
	(-5.161)	(3.793)	(-3.395)	(3.115)		
France	-0.004	127.939	-13.873	0.502	–	–
	(-2.391)	(7.833)	(-7.413)	(7.029)		
Germany	-0.006	207.289	-23.112	0.859	6,994	8,848
	(-2.461)	(7.778)	(-7.507)	(7.257)		
Italy	0.001	16.378	-1.494	0.048	–	–
	(0.132)	(0.411)	(-0.321)	(0.265)		
Japan	-0.025	172.620	-19.784	0.758	–	–
	(-1.888)	(6.008)	(-5.762)	(5.533)		
Netherlands	0.006	134.432	-14.392	0.514	–	–
	(3.509)	(5.173)	(-4.912)	(4.672)		
New Zealand	0.020	104.900	-11.906	0.448	3,538	14,169
	(4.894)	(1.012)	(-1.021)	(1.025)		
Norway	0.037	80.229	-9.556	0.374	1,845	13,373
	(2.063)	(1.928)	(-2.060)	(2.153)		
Portugal	0.012	51.350	-6.314	0.261	–	–
	(1.272)	(1.185)	(-1.219)	(1.262)		
Spain	0.000	172.153	-19.742	0.757	–	–
	(0.093)	(4.952)	(-4.795)	(4.659)		
Sweden	-0.034	211.025	-23.160	0.853	–	–
	(-2.587)	(3.778)	(-3.641)	(3.519)		
Switzerland	-0.034	264.167	-29.101	1.076	–	–
	(-5.413)	(8.386)	(-8.182)	(8.032)		
UK	-0.017	371.170	-40.188	1.452	–	–
	(-10.420)	(21.586)	(-21.101)	(20.624)		
USA	-0.009	210.194	-22.404	0.797	–	–
	(-2.636)	(13.685)	(-13.134)	(12.631)		

[Note] Estimation results for the specification $e_t = c+\delta t+\beta_1 y_t+\beta_2 y_t^2+\beta_3 y_t^3+u_t$. The sample period is 1870–2000 with the exception of New Zealand where the sample starts in 1878. The t-statistics, in parentheses, are computed as described in the text. The turning points, if present, are given by the exponentially transformed real roots of the polynomial $\hat{\beta}_1 + 2\hat{\beta}_2 z + 3\hat{\beta}_3 z^2 = 0$. In bold we indicate the countries for which the null hypothesis of correct specification has not been rejected at the 5% level using the LM test as described in the text.

Table 16: Estimation Results for SO_2 Emissions for the Cubic Specification (20)

Country	$\hat{\delta}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	Turning Points	
Australia	0.019	89.048	-9.938	0.367	4,009	16,983
	(3.733)	(0.811)	(-0.816)	(0.816)		
Austria	-0.024	-75.088	10.245	-0.448	465	8,842
	(-6.087)	(-1.570)	(1.849)	(-2.102)		
Belgium	-0.005	-96.788	12.873	-0.553	603	9,015
	(-1.654)	(-2.274)	(2.699)	(-3.107)		
Canada	0.010	51.568	-4.534	0.120	5,950	1.3×10^7
	(2.696)	(3.471)	(-2.641)	(1.823)		
Denmark	-0.044	-138.328	17.605	-0.722	767	15,019
	(-4.826)	(-4.731)	(5.330)	(-5.788)		
Finland	0.018	-50.001	7.553	-0.355	194	7,380
	(1.089)	(-1.194)	(1.543)	(-1.855)		
France	0.004	-103.791	13.066	-0.540	1,167	8,749
	(1.141)	(-3.252)	(3.573)	(-3.869)		
Germany	-0.003	-91.247	11.733	-0.494	954	7,910
	(-0.524)	(-1.662)	(1.850)	(-2.026)		
Italy	-0.017	-123.204	15.339	-0.622	1,078	12,731
	(-2.121)	(-2.833)	(3.012)	(-3.133)		
Japan	-0.022	83.263	-8.847	0.311	31,015	5,621
	(-2.432)	(4.265)	(-3.791)	(3.340)		
Netherlands	-0.001	-174.476	21.594	-0.880	1,441	8,842
	(-0.484)	(-4.373)	(4.801)	(-5.214)		
New Zealand	0.005	400.225	-43.612	1.579	5,784	17,257
	(0.952)	(3.259)	(-3.152)	(3.039)		
Norway	0.013	-33.340	4.914	-0.229	255	6,302
	(0.880)	(-0.968)	(1.280)	(-1.593)		
Portugal	0.004	26.723	-3.028	0.118	–	–
	(0.662)	(0.905)	(-0.858)	(0.838)		
Spain	0.003	1.810	0.585	-0.051	0	7,519
	(0.748)	(0.048)	(0.132)	(-0.293)		
Sweden	-0.024	-52.225	8.303	-0.397	120	9,569
	(-1.621)	(-0.809)	(1.130)	(-1.417)		
Switzerland	-0.074	-54.756	8.261	-0.370	145	20,497
	(-6.914)	(-1.018)	(1.360)	(-1.615)		
UK	-0.006	-60.385	8.415	-0.373	372	9,084
	(-2.036)	(-1.887)	(2.374)	(-2.848)		
USA	-0.018	148.216	-15.316	0.527	19,029	13,455
	(-3.958)	(7.092)	(-6.599)	(6.143)		

[Note] Estimation results for the specification $e_t = c + \delta t + \beta_1 y_t + \beta_2 y_t^2 + \beta_3 y_t^3 + u_t$. The sample period is $1870 - 2000$. The t-statistics, in parentheses, are computed as described in the text. The turning points, if present, are given by the exponentially transformed real roots of the polynomial $\hat{\beta}_1 + 2\hat{\beta}_2 z + 3\hat{\beta}_3 z^2 = 0$. In bold we indicate the countries for which the null hypothesis of correct specification has not been rejected at the 5% level using the LM test as described in the text.

Authors: Seung Hyun Hong, Martin Wagner

Title: Nonlinear Cointegration Analysis and the Environmental Kuznets Curve

Reihe Ökonomie / Economics Series 224

Editor: Robert M. Kunst (Econometrics)
Associate Editors: Walter Fisher (Macroeconomics), Klaus Ritzberger (Microeconomics)

ISSN: 1605-7996

Stumpergasse 56, A-1060 Vienna • ☎ +43 1 59991-0 • Fax +43 1 59991-555 • http://www.ihs.ac.at

ISSN: 1605-7996