SFB 823

# The environmental Kuznets curve for carbon dioxide emissions: A seemingly unrelated cointegrating polynomial regressions approach

Martin Wagner, Peter Grabarczyk,

Discussion Paper

Nr. 75/2016

SFB 823

# The Environmental Kuznets Curve for Carbon Dioxide Emissions: A Seemingly Unrelated Cointegrating Polynomial Regressions Approach

Martin Wagner
Faculty of Statistics
Technical University Dortmund
Dortmund, Germany
&
Institute for Advanced Studies
Vienna, Austria
&
Bank of Slovenia
Ljubljana, Slovenia

Peter Grabarczyk
Faculty of Statistics
Technical University Dortmund
Dortmund, Germany

**Abstract**

We present estimation and inference techniques for systems of seemingly unrelated cointegrating polynomial regressions. In particular, we present two fully modified-type estimators and Wald-type hypothesis tests based upon them. We develop tests for poolability of subsets of coefficients over subsets of equations. For the case that these restrictions are not rejected, we provide the correspondingly pooled estimators. This *group-wise pooling* turns out to be very useful in our application where we analyze the environmental Kuznets curve for $CO_2$ emissions for seven early industrialized countries. Group-wise pooled estimation leads to almost the same results as unrestricted estimation whilst reducing the number of estimated parameters by about one third. Fully pooled, panel-data type estimation performs poorly in comparison.

**JEL Classification**: C12, C13, C32, Q20

**Keywords**: Cointegrating Polynomial Regression, Environmental Kuznets Curve, Fully Modified Estimation, Poolability, Seemingly Unrelated Regression

# 1 Introduction

The environmental Kuznets curve (EKC) hypothesis postulates an inverted U-shaped relationship between measures of economic development, typically GDP per capita, and measures of per capita pollution or emissions. 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.

Starting with the pioneering work of Grossman and Krueger (1991, 1993, 1995) and Shafik and Bandyopadhyay (1992) a large and still growing body of research, both theoretical and empirical, has been devoted to the EKC hypothesis. Theoretical contributions include Andreoni and Levinson (2001), Arrow *et al.* (1995), Brock and Taylor (2005, 2010), Cropper and Griffiths (1994), Dinda (2005), Jones and Manuelli (2001), Selden and Song (1995) or Stokey (1998).[1] Müller-Fürstenberger and Wagner (2007) discuss problems that arise at the intersection of theoretical and empirical EKC analysis. Additional early empirical contributions on top of the mentioned seminal papers include Agras and Chapman (1999), Antweiler *et al.* (2001), Hilton and Levinson (1998), Holtz-Eakin and Selden (1995),[2] Kahn (1998), List and Gallet (1999) or Torras and Boyce (1998).

Criticism of the EKC is as old as the EKC itself, both on theoretical as well as on econometric grounds. In this paper we focus on discussing two problems related to (i) using unit root and cointegration methods for (ii) multi-country (or multi-regional) data in a parametric approach to the EKC. The problems addressed also impact – if unit root nonstationary behavior of explanatory variables is indeed present – the validity of other estimation approaches to the EKC, including non-parametric approaches (see, e.g., Millimet *et al.*, 2003), semi-parametric approaches (see, e.g., Bertinelli and Strobl, 2005) or specifications based on spline interpolations (see, e.g., Schmalensee *et al.*, 1998).

Given that a significant part of the empirical literature uses unit root and cointegration techniques, understanding the implications of (i) and (ii) is important for empirical practice. Papers that use time series unit root and cointegration methods include Esteve and Tamarit (2012), Fosten *et al.*

[1] A relatively recent survey of economic models for analyzing the EKC is given by Kijima *et al.* (2010). Uchiyama (2016, Chapter 2) contains a detailed discussion of the model of Stokey (1998) as well as an overview discussion of empirical work on the EKC. Already early survey papers like Stern (2004) or Yandle *et al.* (2004) find more than 100 refereed publications; and many more written since then.

[2] The quadratic formulation, i.e., the functional form that can literally lead to an inverted U-shape has first been used in this paper, whereas Grossman and Krueger used a third order polynomial.

(2012), Friedl and Getzner (2003), He and Richard (2010), Jalil and Mahmud (2009) and Lindmark (2002). Panel data studies using unit root and cointegration techniques include Apergis (2016), Auffhammer and Carson (2008), Baek (2015), Bernard *et al.* (2015), Dijkgraaf and Vollebergh (2005), Dinda and Coondoo (2006), Galeotti *et al.* (2006), Perman and Stern (2003) or Romero-Avila (2008). As pointed out by Wagner (2015), based on Wagner and Hong (2016), these papers ignore the fact that powers of integrated processes are not themselves integrated processes (see also Wagner, 2012). Therefore, a regression of (the logarithm of) emissions per capita on (the logarithm of) GDP per capita and its powers is not a *standard* cointegrating regression, but in the terminology of Wagner and Hong (2016, eq. (1)) a *cointegrating polynomial regression* (CPR); if this specific form of nonlinear cointegration prevails and the regression is not spurious.[3]

In the presence of powers of integrated regressors in cointegrating regressions, estimators like the fully modified OLS (FM-OLS) estimator (introduced for the linear cointegration case in Phillips and Hansen, 1990) can be adapted by using *appropriately constructed* additive correction terms. The precise form of these correction terms depends upon the specification of the relationship. They differ from the correction terms in the linear case, see Wagner and Hong (2016).[4] The implications of this difference for EKC analysis for time series data are illustrated in Wagner (2015). The asymptotic behavior of treating unit root process and and their powers all as unit root processes and using the standard FM-OLS estimator this way in the CPR setting is discussed in Stypka *et al.* (2016).

The part of the empirical EKC literature that uses panel unit root and cointegration techniques relies entirely upon methods for *linear* cointegration developed for cross-sectionally independent panels. Thus, a fortiori the above-mentioned problems continue to be present. Importantly, additionally the assumption of cross-sectional independence that is employed in these studies, utilizing standard panel cointegration techniques like Kao and Chiang (2000), Phillips and Moon (1999) or Pedroni (2000), is clearly often unrealistic.[5] Also, the tacit assumption of these studies that

[3]Clearly, tests for nonlinear cointegration in EKC-type relationships need to be discussed, see, e.g., Choi and Saikkonen (2010), Wagner (2013) or Wagner and Hong (2016).

[4]Important earlier work in this respect has been undertaken by Park and Phillips (1999, 2001), Chang *et al.* (2001) or Ibragimov and Phillips (2008). The difference between the work of Wagner and Hong (2016) and, e.g., Chang *et al.* (2001) is that the latter assume that the regressors are pre-determined and the errors serially uncorrelated. Wagner and Hong (2016) remove these two assumptions and consider the "standard" setting in cointegration analysis with endogenous regressors and serially correlated errors.

[5]Apergis (2016) and Romero-Avila (2008) acknowledge the potential of cross-sectional dependencies in time series panels by considering some form of cross-sectional dependence testing. That alone, however, does not solve the associated problems.

all coefficients (except for, usually, the intercepts) are indeed identical, i.e., can be *pooled*, for all cross-section members may be too restrictive in many applications. In case that the cross-sectional dimension is small (compared to the time series dimension) a *seemingly unrelated regressions* (SUR) approach allows to relax both the cross-sectional independence as well as the poolability assumption. Based on Hong and Wagner (2014) we present in Section 2 fully modified OLS SUR estimators for systems of *seemingly unrelated cointegrating polynomial regressions* (SUCPR) formulated here for the quadratic EKC specification as used in the application.[6] In the SUCPR setting we allow for cross-sectional dependence of both the regressors and the errors and do not impose any poolability assumptions on the coefficients. Instead of having to impose poolability of the coefficients, we can test for any form of pooling and then estimate the parameters pooled correspondingly. Some basic forms of pooling related to panel analysis are reviewed and stated in Appendix A.1: (P) all coefficients but the intercepts are pooled, (S) only the coefficients corresponding to log GDP per capita and its powers are pooled, and (T) only the coefficient corresponding to the linear time trend is pooled. More generally, however, it may be the case that only some coefficients can be pooled over (potentially) different subsets of cross-section members. This turns out to be the case in the application in Section 3. Therefore we discuss estimation in partially pooled settings of a form relevant for our application in detail in Section 2.1.

The application of our methodology to study the EKC for $CO_2$ emissions for seven early industrialized countries over the period 1870–2009 highlights the usefulness of the SUCPR approach. Group-wise pooled estimation of the EKC leads to almost equal to almost the same results (estimated parameters, turning points, and fitted values) as those obtained with unrestricted individual or SUCPR estimation. This happens despite the reduction of the number of parameters to be estimated by about one third. Fully pooled estimation, rejected by poolability testing, on the other hand, performs drastically worse. This shows that the situation specific approach to pooling that our methodology provides is a helpful addition to the EKC analysis toolkit. The flexibility of the approach will allow for fruitful applicability also when modeling other relationships with data sets with a small cross-sectional dimension compared to a large time series dimension.

The paper is organized as follows: In Section 2 we present the econometric methodology, i.e., two fully modified least squares estimators for systems of seemingly unrelated cointegrating polynomial

[6]In terms of econometric methodology the paper discusses an extension of SUR cointegration analysis from the linear cointegration SUR case (see, e.g., Park and Ogaki, 1991; Mark *et al.*, 2005; Moon, 1999; Moon and Perron, 2005) to the SUCPR case. This is similar in scope – now for the SUR case – to the extension of FM-OLS from the linear cointegration to the CPR case presented in Wagner and Hong (2016).

regressions including a discussion of group-wise pooling – both with respect to testing for poolability as well estimation imposing the corresponding pooling – of a form relevant for our application. Section 3 presents and discusses the empirical findings and Section 4 briefly summarizes and concludes. Two appendices follow the main text: Appendix A contains some additional material and results concerning different variations of pooled estimation relevant for EKC analysis. Appendix A also contains the derivation of the limiting distributions of the group-wise pooled estimators. Appendix B contains additional empirical results.

We use the following notation: $\lfloor x \rfloor$ denotes the integer part of $x \in \mathbb{R}$ and $\mathrm{diag}(\cdot)$ denotes a diagonal matrix with entries specified throughout. For a vector $x = (x_i)_{i=1,\ldots,n}$ we denote by $\|x\|^2 = \sum_{i=1}^n x_i^2$, and for a matrix $M$ we denote by $\|M\| = \sup_x \frac{\|Mx\|}{\|x\|}$. For a square matrix $A$ we denote its determinant by $|A|$. We denote the $m$-dimensional identity matrix by $I_m$, with $0_{m\times n}$ a $(m \times n)$-matrix with all entries equal to zero, with $\mathbf{1}_s = [1,\ldots,1]' \in \mathbb{R}^s$ and with $e_{i,N}$ the $i$-th unit vector in $\mathbb{R}^n$. For (block-)matrices $M$ we denote the (i,j)-(block-)element with $M^{i,j}$, the $i$-th (block-)row with $M^{i,\cdot}$ and the $j$-th (block-)column with $M^{\cdot,j}$. With $\mathbb{1}_{\{\cdot\}}$ we denote the indicator function. Furthermore, $\otimes$ denotes the Kronecker product, $\mathbb{E}(\cdot)$ denotes the expected value and $L$ denotes the backward-shift operator, i.e., $L\{z_t\}_{t\in\mathbb{Z}} = \{z_{t-1}\}_{t\in\mathbb{Z}}$. Definitional equality is signified by $:=$ and $\Rightarrow$ denotes weak convergence. Brownian motions are denoted $B(r)$ or short-hand by $B$, with covariance matrices specified in the context. For integrals of the form $\int_0^1 B(s)ds$ or $\int_0^1 B(s)dB(s)$, we often use the short-hand notation $\int B$ or $\int BdB$ and drop function arguments for notational simplicity.

## 2 Seemingly Unrelated Cointegrating Polynomial Regressions

For the discussion in this paper it suffices to consider the following special case of a system of seemingly unrelated quadratic polynomial regressions. In the application in the following section $y_{i,t}$ denotes log $CO_2$ emissions per capita and $x_{i,t}$ log GDP per capita in year $t$ in country $i$:

$$\begin{aligned} y_{i,t} &= c_i + \delta_i t + \beta_{1,i} x_{i,t} + \beta_{2,i} x_{i,t}^2 + u_{i,t}, \quad i = 1,\ldots,N, \ t = 1,\ldots,T, \qquad (1) \\ &= Z_{i,t}' \theta_i + u_{i,t} \\ x_{i,t} &= x_{i,t-1} + v_{i,t}, \end{aligned}$$

with $Z_{i,t} := [1, t, x_{i,t}, x_{i,t}^2]'$ and $\theta_i := [c_i, \delta_i, \beta_{1,i}, \beta_{2,i}]'$. Denoting with $x_t := [x_{1,t}, \ldots, x_{N,t}]'$, with $u_t := [u_{1,t}, \ldots, u_{N,t}]'$ and with $v_t := [v_{1,t}, \ldots, v_{N,t}]'$, we assume for $\xi_t := [u_t', v_t']'$ that

$$\begin{aligned} u_t &:= \Psi(L)\zeta_t = \sum_{j=0}^{\infty} \Psi_j \zeta_{t-j}, \\ \Delta x_t = v_t &:= \Phi(L)\epsilon_t = \sum_{j=0}^{\infty} \Phi_j \epsilon_{t-j}, \end{aligned} \tag{2}$$

with $\sum_{j=0}^{\infty} j\|\Phi_j\| < \infty$, $\sum_{j=0}^{\infty} j\|\Psi_j\| < \infty$. Furthermore, we assume $|\Phi(1)| \neq 0$, which excludes stationarity of $\{x_t\}$, and $|\Psi(1)| \neq 0$, since we need regularity of this matrix for the construction of the *modified SUR* estimator, a term coined by Park and Ogaki (1991) in the linear SUR cointegration setting. The stacked process $\{\xi_t^0\}_{t\in\mathbb{Z}} := \{[\epsilon_t', \zeta_t']'\}_{t\in\mathbb{Z}}$ is assumed to be a strictly stationary and ergodic martingale difference sequence with respect to the natural filtration $\mathcal{F}_t$ with positive definite conditional variance matrix $\Sigma := \mathbb{E}\left(\xi_t^0(\xi_t^0)'|\mathcal{F}_{t-1}\right)$ and $\sup_{t\geq 1} \mathbb{E}(\|\xi_t^0\|^r|\mathcal{F}_{t-1}) < \infty$ a.s. for some $r > 4$.

**Remark 1** *The above setting in* (1) *can be generalized in several ways: First, several integrated regressors and their powers can be included, with the specifications allowed to be equation specific. In the above example this means that different powers can be included in the different equations. Second, more general (equation-specific) deterministic components can be included. Third, predetermined (or even more easily strictly exogenous) stationary regressors can be included as well. Fourth, common non-cointegrated nonstationary regressors can also be included in the equation system, which may be of particular relevance in, e.g., regional applications where country-wide variables may be important determinants for all regions. For more details in these respects see Hong and Wagner (2014).*

The above assumptions are sufficient for

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor rT \rfloor} \xi_t = \frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor rT \rfloor} \begin{bmatrix} u_t \\ v_t \end{bmatrix} \Rightarrow B(r) = \begin{bmatrix} B_u(r) \\ B_v(r) \end{bmatrix} = \Omega^{1/2} W(r), \tag{3}$$

with $W(r)$ a $2N$-dimensional standard Wiener process and $\Omega := \sum_{h=-\infty}^{\infty} \mathbb{E}(\xi_0 \xi_h')$ the so-called long run variance of $\{\xi_t\}_{t\in\mathbb{Z}}$. For later usage we define also the one-sided long run variance given by $\Delta := \sum_{h=0}^{\infty} \mathbb{E}(\xi_0 \xi_h')$ and both matrices are partitioned according to the partitioning of $\xi_t$:

$$\Omega := \begin{bmatrix} \Omega_{uu} & \Omega_{uv} \\ \Omega_{vu} & \Omega_{vv} \end{bmatrix}, \ \Delta := \begin{bmatrix} \Delta_{uu} & \Delta_{uv} \\ \Delta_{vu} & \Delta_{vv} \end{bmatrix} \tag{4}$$

The above set of $N$ equations (1) can be jointly written as

$$y_t = Z_t'\theta + u_t, \tag{5}$$

with

$$y_t := \begin{bmatrix} y_{1,t} \\ \vdots \\ y_{N,t} \end{bmatrix} \in \mathbb{R}^N, \; Z_t := \begin{bmatrix} Z_{1,t} & & \\ & \ddots & \\ & & Z_{N,t} \end{bmatrix} \in \mathbb{R}^{4N \times N}, \; u_t := \begin{bmatrix} u_{1,t} \\ \vdots \\ u_{N,t} \end{bmatrix} \in \mathbb{R}^N,$$

and with $\theta := [\theta_1', \ldots, \theta_N']'$. Stacking all $T$ observations for the above equation (5) we arrive at

$$y = Z\theta + u, \tag{6}$$

with

$$y := \begin{bmatrix} y_1 \\ \vdots \\ y_T \end{bmatrix} \in \mathbb{R}^{NT}, \quad Z := \begin{bmatrix} Z_1' \\ \vdots \\ Z_T' \end{bmatrix} \in \mathbb{R}^{NT \times 4N}.$$

A few basic observations concerning parameter estimation in (6) can already be made: First, it is straightforward to show that the OLS estimator of $\theta$ in (6) is consistent with a limiting distribution contaminated by second order bias terms, just as in the linear seemingly unrelated cointegrating case studied in Park and Ogaki (1991) or Moon (1999). Alternatively, the results for the OLS estimator given in Wagner and Hong (2016) for the single equation case, of course, generalize to the SUCPR case. Second, in the classical SUR approach of Zellner (1962) the errors are typically assumed to be serially uncorrelated (and the regressors nonstochastic). Correspondingly, the weighting matrix used in SUR estimation, i.e., in GLS estimation, is an estimate of the contemporaneous error covariance matrix. In the cointegration setting we allow for error serial correlation (and in addition for endogenous regressors). To take this into account, Park and Ogaki (1991) define a *modified SUR* (MSUR) estimator using an estimate of the long run variance matrix of the errors as weighting matrix. The asymptotic behavior of these two estimators is derived in Hong and Wagner (2014, Proposition 1) for the SUCPR case. As in the time series case, the limiting distributions of the OLS and additionally the MSUR estimator are the starting points to perform the two-part FM-type corrections.[7] One of the corrections is as in the linear case, i.e., the dependent variable

[7]For completeness, the OLS estimator is (as always) given by $\hat{\theta}_{\text{OLS}} := (Z'Z)^{-1} Z'y$ and the MSUR estimator is defined as $\hat{\theta}_{\text{MSUR}} := \left(Z'\left(I_T \otimes \hat{\Omega}_{uu}^{-1}\right) Z\right)^{-1} \left(Z'\left(I_T \otimes \hat{\Omega}_{uu}^{-1}\right) y\right)$.

$y_t$ is replaced by $y_t^+ := y_t - \hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_t$, with consistent estimators of the long run variances.[8] The second transformation consists of subtracting an appropriately constructed correction term. In the SUR setting we need two sets of correction terms, depending upon estimator considered as starting point (OLS or MSUR). For our specification (1) these are given by $A^* := [A_1^{*\prime}, \ldots, A_N^{*\prime}]'$ and $\tilde{A}^* := [\tilde{A}_1^{*\prime}, \ldots, \tilde{A}_N^{*\prime}]'$, with

$$A_i^* := (\hat{\Delta}_{vu}^+)^{i,i} \begin{bmatrix} 0_{2\times 1} \\ T \\ 2\sum_{t=1}^{T} x_{i,t} \end{bmatrix}, \; \tilde{A}_i^* := (\hat{\Delta}_{vu}^+)^{i,\cdot}(\hat{\Omega}_{u.v}^{-1})^{\cdot,i} \begin{bmatrix} 0_{2\times 1} \\ T \\ 2\sum_{t=1}^{T} x_{i,t} \end{bmatrix}, \tag{7}$$

where $(\hat{\Delta}_{vu}^+)^{i,i}$ is a consistent estimator of $(\Delta_{vu}^+)^{i,i} := \Delta_{vu}^{i,i} - (\Delta_{vv})^{i,\cdot}\Omega_{vv}^{-1}\Omega_{vu}^{\cdot,i}$ and $\hat{\Omega}_{u.v}$ is a consistent estimator of $\Omega_{u.v} := \Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}$.

In order to finally define the two fully modified estimators and to state their asymptotic distributions we still need to define some additional quantities. We define, again for our special case, the weighting matrix $G = G(T) := I_N \otimes G_\bullet(T)$, with $G_\bullet(T) := \text{diag}(T^{-1/2}, T^{-3/2}, T^{-1}, T^{-3/2})$ and a stochastic process $J(r) := \text{diag}\,(J_1(r), \ldots, J_N(r))$ with $J_i(r) := [1, r, B_{v_i}(r), B_{v_i}^2(r)]'$, with $B_{v_i}(r)$ denoting the $i$-th coordinate of $B_v(r)$.

**Proposition 1 (Hong and Wagner 2014, Proposition 2)** *Let $y_t$ be generated by (1) with the assumptions given in place. Assume furthermore that, based on the OLS residuals, all required long run variances are estimated consistently. Using the correction factors defined in (7) the fully modified systems OLS (FM-SOLS) and the fully modified SUR (FM-SUR) estimators are given by:*

$$\hat{\theta}_{\text{FM-SOLS}} := \left(Z'Z\right)^{-1}\left(Z'y^+ - A^*\right), \tag{8}$$

$$\hat{\theta}_{\text{FM-SUR}} := \left(Z'\left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right)Z\right)^{-1}\left(Z'\left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right)y^+ - \tilde{A}^*\right), \tag{9}$$

*with $y^+ := [y_1^{+\prime}, \ldots, y_T^{+\prime}]'$. As $T \to \infty$ it holds that*

$$G^{-1}\left(\hat{\theta}_{\text{FM-SOLS}} - \theta\right) \Rightarrow \left(\int JJ'\right)^{-1}\int JdB_{u.v}, \tag{10}$$

$$G^{-1}\left(\hat{\theta}_{\text{FM-SUR}} - \theta\right) \Rightarrow \left(\int J\Omega_{u.v}^{-1}J'\right)^{-1}\int J\Omega_{u.v}^{-1}dB_{u.v}, \tag{11}$$

*where $B_{u.v}(r) := B_u(r) - \Omega_{uv}\Omega_{vv}^{-1}B_v(r)$ is a Brownian motion with variance matrix $\Omega_{u.v}$. By construction $B_{u.v}(r)$ is independent of $B_v(r)$.*

[8] The results of, e.g., Jansson (2002) apply in our setting and provide conditions on kernels and bandwidths that allow for consistent long run variance estimation.
Note for completeness that a more detailed discussion concerning possibilities to construct FM-type estimators in the SUR case is given in Hong and Wagner (2014) and Moon (1999).

The above zero mean Gaussian mixture limiting distributions given in (10) and (11) form the basis for asymptotic chi-squared inference using, e.g., Wald-type tests. Because the vectors $\hat{\theta}_{\text{FM-SOLS}}$ and $\hat{\theta}_{\text{FM-SUR}}$ contain elements that converge at different rates, obtaining formal results for the Wald statistics requires a condition on the restriction matrix $R$ that is unnecessary when all estimated coefficients converge at the same rate (see, e.g., Park and Phillips, 1988, 1989). We posit in the following proposition a sufficient (asymptotic) rank condition that ensures that the Wald-type test statistics have asymptotic chi-squared null distributions. Note that if none of the hypotheses mixes coefficients with different convergence rates no additional complications compared to a standard situation arise.

**Proposition 2 (Hong and Wagner 2014, Proposition 3)** *Let $y_t$ be generated by* (1) *and the given assumptions in place. Consider $s$ linearly independent restrictions collected in $H_0 : R\theta = r$ with $R \in \mathbb{R}^{s\times 4N}$ with full row rank $s$, $r \in \mathbb{R}^s$ and suppose that there exists a (matrix sequence) $G_R = G_R(T)$ such that $\lim_{T\to\infty} G_R R G = R^*$ with $R^* \in \mathbb{R}^{s\times 4N}$ of full rank $s$.*

*Then it holds that under $H_0$ the Wald-type statistics:*

$$W_{FM\text{-}SOLS} := \left(R\hat{\theta}_{FM\text{-}SOLS} - r\right)' \left[R\left(Z'Z\right)^{-1} Z' \left(I_T \otimes \hat{\Omega}_{u.v}\right) Z \left(Z'Z\right)^{-1} R'\right]^{-1} \left(R\hat{\theta}_{FM\text{-}SOLS} - r\right), \tag{12}$$

$$W_{FM\text{-}SUR} := \left(R\hat{\theta}_{FM\text{-}SUR} - r\right)' \left[R\left(Z'\left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right) Z\right)^{-1} R'\right]^{-1} \left(R\hat{\theta}_{FM\text{-}SUR} - r\right) \tag{13}$$

*are asymptotically chi-squared distributed with $s$ degrees of freedom.*

## 2.1 Testing for Poolability and Pooled Estimation

As outlined in the introduction a key advantage of a SUR setting is that it allows to test for *in principle arbitrary forms of* poolability rather than assuming poolability from the outset as in panel analysis. Clearly, the results from Propositions 1 and 2 allow to test for poolability of the coefficients. In Appendix A.1 we briefly present the test statistics and the correspondingly pooled estimators for three "standard" pooling tests involving all cross-section members. These are labelled as: (P), where all coefficients except for the intercepts are pooled; (S), where only the coefficients to $x_{i,t}$ and $x_{i,t}^2$ are pooled and (T), where only the linear trend coefficient is pooled.

The first variant of pooling corresponds closely to a fixed-effects panel model, with individual specific fixed effects. Note, however, that the literature does not provide the theory for panel estimation methods (with $N \to \infty$) for cross-sectionally dependent panels of cointegrating polynomial regressions. de Jong and Wagner (2016) provide the theory for the cross-sectionally independent case for the cubic formulation with one- and two-way fixed effects.[9]

If the considered null hypothesis is not rejected, then pooled estimation of a smaller number of parameters allows to lift some efficiency gains in estimation. For our data, however, it turns out that all three "global" hypotheses (P), (S) and (T) are rejected.[10]

A more careful analysis reveals that the coefficient corresponding to the linear time trend can be pooled in two subgroups and the coefficients to GDP and its square, the stochastic regressors, can be pooled in four subgroups, with two of these subgroups comprising only one country. To exploit the possibilities of group-wise pooling thus necessitates writing down the corresponding Wald-type statistics as well as the correspondingly group-wise pooled estimators. This is discussed in the following subsection for the setting relevant in our application. Along similar lines any form of group-wise pooling can be considered in more general SUCPR settings.

## 2.2 Group-Wise Pooling

In this subsection we consider the situation where we test the null hypothesis that the coefficients for the linear time trend are group-wise pooled over a partition of $k$ subsets $I_{n_j},\ j = 1, \ldots, n_k$ with $I := \{1, \ldots, N\} = \bigcup_{j=1}^{k} I_{n_j}$. Similarly, we consider a partition over $l$ subsets $I_{n_j},\ j = 1, \ldots, n_l$ for the regressors $x_{i,t}$ and $x_{i,t}^2$, i.e., $I = \bigcup_{j=1}^{l} I_{m_j}$. Without loss of generality order the subsets according to decreasing cardinality, i.e., $|I_{n_1}| \geq \ldots \geq |I_{n_k}|$ and $|I_{m_1}| \geq \ldots \geq |I_{m_l}|$, denoting here with $|S|$ the number of elements of a set $S$.

The null hypothesis corresponding to group-wise poolability of the coefficients corresponding to the above partitioning is given by:

$$\begin{aligned} H_0^{\mathrm{GW}}: \qquad \delta_i = \delta_j \qquad & \forall\, i,j \in I_{n_d}\ \forall\, d \in \{\{1, \ldots, k\} : |I_{n_d}| > 1\} \\ \begin{pmatrix} \beta_{1,i} \\ \beta_{2,i} \end{pmatrix} = \begin{pmatrix} \beta_{1,j} \\ \beta_{2,j} \end{pmatrix} \quad & \forall\, i,j \in I_{m_p}\ \forall\, p \in \{\{1, \ldots, l\} : |I_{m_p}| > 1\}. \end{aligned} \tag{14}$$

To construct the Wald-type test statistics discussed in Proposition 2 for this specific situation it is convenient to define a few more quantities. First, denote with $N_j = |I_{n_j}|,\ j = 1, \ldots, k$ and

[9] Note again that the part of the empirical EKC literature that uses panel cointegration methods, estimates a system of equations like (1) with methods for linear cointegration developed for panels of cross-sectionally independent units. The SUCPR approach overcomes these two limitations, allowing for cross-sectional dependence and taking into account the specific form of nonlinear cointegration.

[10] As will be seen in Section 3, for the 19 countries considered, (non-)cointegration tests lead to evidence for a CPR relationship in seven countries. The CPR and SUCPR analysis is consequently performed with the data for these seven countries.

$M_j = |I_{m_j}|,\ j = 1, \ldots, l$. Furthermore, the elements of the index set $I_{n_j}$, $a_{j,n_j}$ say, are considered sorted, i.e., $I_{n_j} = (a_{1,n_j}, a_{2,n_j}, \ldots, a_{N_j,n_j})$ with $1 \leq a_{1,n_j} < a_{2,n_j} < \cdots < a_{N_j,n_j} \leq N$ for $j = 1, \ldots, k$ and similarly for the sets $I_{m_j},\ j = 1, \ldots, l$. Using this notation and setting the restriction matrix to test for (the considered form of) group-wise poolability can be written as

$$R^{\mathrm{GW}} := [R'_{n_1}, \ldots, R'_{n_k}, R'_{m_1}, \ldots, R'_{m_l}]' \in \mathbb{R}^{s \times 4N} \tag{15}$$

with

$$R_{n_j} := \left( \left( \mathbf{1}_{(N_j-1)} \otimes e'_{a_{1,n_j},N} \right) - \begin{pmatrix} e'_{a_{2,n_j},N} \\ \vdots \\ e'_{a_{N_j,n_j},N} \end{pmatrix} \right) \otimes e'_{2,4} \in \mathbb{R}^{(N_j-1) \times 4N} \tag{16}$$

for $j$ such that $N_j > 1$ and $R_{n_j} = \emptyset$ otherwise; and

$$R_{m_j} := \left( \left( \mathbf{1}_{(M_j-1)} \otimes e'_{a_{1,m_j},N} \right) - \begin{pmatrix} e'_{a_{2,m_j},N} \\ \vdots \\ e'_{a_{M_j,m_j},N} \end{pmatrix} \right) \otimes \begin{pmatrix} e'_{3,4} \\ e'_{4,4} \end{pmatrix} \in \mathbb{R}^{2(M_j-1) \times 4N} \tag{17}$$

for $j$ such that $M_j > 1$ and $R_{m_j} = \emptyset$ otherwise. The total number of restrictions is

$$s = \sum_{j=1}^{k} (N_j - 1) + 2 \sum_{j=1}^{l} (M_j - 1) \tag{18}$$

and, clearly, $r = 0$ (in $R\theta = r$) here. Using either the FM-SOLS estimates or the FM-SUR estimates, the two test statistics (12) and (13) can be calculated to test the considered null hypothesis $H_0^{\mathrm{GW}}$

**Remark 2** *In the above definition of the blocks of the restriction matrix, setting, e.g., $R_{n_j} = \emptyset$ for $N_j = 1$, merely states that for groups of size one, of course, no poolability hypothesis testing is performed. Including in $R^{GW}$ only the subsets of size larger than one would require to define another index, say $n_k^*$, until which the groups – ordered according to non-increasing size – comprise more than one element.*

In case that the null hypothesis discussed above is not rejected, the corresponding group-wise pooled estimators can be (defined and) employed. To this end consider

$$\ddot{D}_t := [D_{n_1,t}, \ldots, D_{n_k,t}]' \in \mathbb{R}^{k \times N}, \tag{19}$$

where

$$D_{n_j,t} := \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{n_j}\}} \cdot t \cdot e'_{i,N}. \tag{20}$$

For the stochastic regressors we similarly have

$$\ddot{X}_t := [X'_{m_1,t}, \ldots, X'_{m_l,t}]' \in \mathbb{R}^{2l\times N}, \tag{21}$$

with

$$X_{m_j,t} := \sum_{i=1}^{N} \mathbb{1}_{\{i\in I_{m_j}\}} \cdot \left(e'_{i,N} \otimes X_{i,t}\right) \tag{22}$$

and $X_{i,t} := [x_{i,t}, x_{i,t}^2]'$. With these quantities the group-wise pooled model can be compactly written as

$$y_t = \ddot{Z}_t'\theta^{\mathrm{GW}} + u_t, \tag{23}$$

with $y_t := [y_{1,t}, \ldots, y_{N,t}]'$, $u_t := [u_{1,t}, \ldots, u_{N,t}]'$, $\ddot{Z}_t := [I_N, \ddot{D}_t', \ddot{X}_t']' \in \mathbb{R}^{(N+k+2l)\times N}$ and the parameter vector $\theta^{\mathrm{GW}} := [c_1, \ldots, c_N, \delta_{n_1}, \ldots, \delta_{n_k}, \beta'_{m_1}, \ldots, \beta'_{m_l}]' \in \mathbb{R}^{N+k+2l}$. Finally, stacking the quantities over time gives

$$y = \ddot{Z}\theta^{\mathrm{GW}} + u, \tag{24}$$

with $y = [y_1, \ldots, y_T]'$, $u = [u_1, \ldots, u_T]'$ and $\ddot{Z} = [\ddot{Z}_1, \ldots, \ddot{Z}_T]'$.

The correction terms for the group-wise pooled FM-SOLS and FM-SUR estimators are defined as $A^{\mathrm{GW}} := [0_{1\times(N+n_k)}, A_{m_1}^{\mathrm{GW}\prime}, \ldots, A_{m_l}^{\mathrm{GW}\prime}]'$, $\ddot{A}^{\mathrm{GW}} := [0_{1\times(N+n_k)}, \ddot{A}_{m_1}^{\mathrm{GW}\prime}, \ldots, \ddot{A}_{m_l}^{\mathrm{GW}\prime}]'$, with

$$A_{m_j}^{\mathrm{GW}} := \sum_{i=1}^{N} \mathbb{1}_{\{i\in I_{m_j}\}} \cdot \left(\hat{\Delta}_{vu}^{+}\right)^{i,i} \cdot \begin{pmatrix} T \\ 2\sum_{t=1}^{T} x_{i,t} \end{pmatrix}, \tag{25}$$

$$\ddot{A}_{m_j}^{\mathrm{GW}} := \sum_{i=1}^{N} \mathbb{1}_{\{i\in I_{m_j}\}} \cdot \left(\hat{\Delta}_{vu}^{+}\right)^{i,\cdot} \left(\hat{\Omega}_{u\cdot v}^{-1}\right)^{\cdot,i} \cdot \begin{pmatrix} T \\ 2\sum_{t=1}^{T} x_{i,t} \end{pmatrix}. \tag{26}$$

For group-wise pooled estimation the weighting matrix is given by $\ddot{G}(T) := \mathrm{diag}(T^{-1/2} \cdot I_N, T^{-3/2} \cdot I_k, I_l \otimes \mathrm{diag}(T^{-1}, T^{-3/2}))$ and the limit stochastic process is given by $\ddot{J}(r) := [I_N, \ddot{J}_D', \ddot{J}_X']'$. Here $\ddot{J}_D(r) := [J_{D_{n_1}}(r)', \ldots, J_{D_{n_k}}(r)']'$ is composed of $J_{D_{n_j}}(r) := \sum_{i=1}^{N} \mathbb{1}_{\{i\in I_{n_j}\}} \cdot r \cdot e'_{i,N}$ and $\ddot{J}_X(r) := [J_{X_{m_1}}(r)', \ldots, J_{X_{m_l}}(r)']'$ is composed of $J_{X_{m_j}}(r) := \sum_{i=1}^{N} \mathbb{1}_{\{i\in I_{m_j}\}} \cdot \left(e'_{i,N} \otimes \begin{pmatrix} B_{v_i}(r) \\ B_{v_i}^2(r) \end{pmatrix}\right)$.

**Proposition 3** *Let $y_t$ be generated by* (24)*, the discussed restricted version of* (1) *with group-wise pooled parameters, with the assumptions given in place. Assume furthermore that, based on the OLS residuals, all required long run variances are estimated consistently. Using the correction factors* *defined in* (25) *and* (26)*, the* group-wise *FM-SOLS and FM-SUR estimators are given by:*

$$\ddot{\theta}^{GW}_{FM\text{-}SOLS} := \left(\ddot{Z}'\ddot{Z}\right)^{-1}\left(\ddot{Z}'y^{+} - A^{GW}\right), \tag{27}$$

$$\ddot{\theta}^{GW}_{FM\text{-}SUR} := \left(\ddot{Z}'\left(I_T \otimes \hat{\Omega}^{-1}_{u.v}\right)\ddot{Z}\right)^{-1}\left(\ddot{Z}'\left(I_T \otimes \hat{\Omega}^{-1}_{u.v}\right)y^{+} - \ddot{A}^{GW}\right). \tag{28}$$

*As* $T \to \infty$ *it holds that*

$$\ddot{G}^{-1}\left(\ddot{\theta}^{GW}_{FM\text{-}SOLS} - \theta^{GW}\right) \Rightarrow \left(\int \ddot{J}\ddot{J}'\right)^{-1}\int \ddot{J}dB_{u\cdot v}, \tag{29}$$

$$\ddot{G}^{-1}\left(\ddot{\theta}^{GW}_{FM\text{-}SUR} - \theta^{GW}\right) \Rightarrow \left(\int \ddot{J}\Omega^{-1}_{u\cdot v}\ddot{J}'\right)^{-1}\int \ddot{J}\Omega^{-1}_{u\cdot v}dB_{u\cdot v}. \tag{30}$$

In the following empirical analysis we discuss and compare unrestricted, pooled and group-wise pooled estimation results.

# 3 Empirical Analysis

The empirical analysis builds upon Wagner (2015), who performs individual country FM-CPR analysis for 19 early industrialized countries. The first step, prior to the SUR analysis performed here, is to reassess the findings of the earlier paper, since we now have data ranging from 1870–2009 rather than only until 2000. The $CO_2$ emissions data are from the Carbon Dioxide Information Analysis Center of the US Department of Energy and comprise total $CO_2$ emissions from fossil fuel usage.[11] The GDP data, measured in 1990 Geary-Khamis Dollars, are from the Maddison project at the University of Groningen.[12] The data are used in logarithms of per capita quantities. Throughout, for all estimators and all tests we use the Bartlett kernel and the bandwidth chosen according to Newey and West (1994).

For all 19 early industrialized countries investigated, the unit root null hypothesis is not rejected for log GDP per capita using the unit root tests of Phillips and Perron (1988) as well as the fixed-$b$ versions of this test developed by Vogelsang and Wagner (2013).[13] Using the tests for cointegration in CPRs of Wagner (2013) and Wagner and Hong (2016) leads to evidence for a quadratic cointegrating EKC for $CO_2$ emissions for the following seven countries: Austria (AT), Belgium (BE), Denmark (DK), Finland (FI), the Netherlands (NL), Switzerland (CH) and the

[11] See Boden *et al.* (2016) and `http://cdiac.ornl.gov`.

[12] See Bolt and van Zanden (2014) and `http://www.ggdc.net/maddison/maddison-project/home.htm`.

[13] The results are given in Table 4 in Appendix B.

| | $\hat{\delta}$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | TP | $\hat{\delta}$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | TP |
|---|---|---|---|---|---|---|---|---|
| | **Austria** | | | | **Belgium** | | | |
| FM-CPR | -0.017 | 6.439 | -0.288 | 71509.279 | -0.004 | 10.989 | -0.579 | 13201.586 |
| ($t$-values) | -3.672 | 2.446 | -1.979 | | -2.385 | 8.876 | -8.878 | |
| FM-SOLS | -0.018 | 8.727 | -0.413 | 39004.716 | -0.005 | 11.974 | -0.631 | 13254.808 |
| ($t$-values) | -4.620 | 3.899 | -3.341 | | -3.274 | 11.927 | -11.939 | |
| FM-SUR | -0.014 | 7.503 | -0.356 | 37456.789 | -0.004 | 10.333 | -0.543 | 13667.603 |
| ($t$-values) | -4.559 | 4.610 | -3.813 | | -3.773 | 13.414 | -13.112 | |
| | **Denmark** | | | | **Finland** | | | |
| FM-CPR | -0.013 | 12.055 | -0.585 | 29852.349 | -0.030 | 15.775 | -0.745 | 39396.139 |
| ($t$-values) | -2.441 | 10.323 | -10.554 | | -3.169 | 8.543 | -7.999 | |
| FM-SOLS | -0.011 | 12.025 | -0.587 | 28069.339 | -0.039 | 16.154 | -0.745 | 51377.431 |
| ($t$-values) | -2.363 | 11.028 | -11.369 | | -4.708 | 9.509 | -8.670 | |
| FM-SUR | -0.009 | 11.656 | -0.573 | 25925.716 | -0.030 | 15.607 | -0.732 | 42424.906 |
| ($t$-values) | -2.693 | 14.047 | -13.646 | | -6.021 | 13.014 | -11.151 | |
| | **Netherlands** | | | | **Switzerland** | | | |
| FM-CPR | 0.001 | 9.370 | -0.477 | 18323.280 | -0.023 | 6.981 | -0.232 | $3.4\times10^6$ |
| ($t$-values) | 0.643 | 7.734 | -7.365 | | -6.539 | 5.686 | -3.431 | |
| FM-SOLS | 0.001 | 9.761 | -0.498 | 17882.497 | -0.023 | 6.138 | -0.188 | $1.2\times10^7$ |
| ($t$-values) | 0.608 | 8.564 | -8.187 | | -7.876 | 6.285 | -3.483 | |
| FM-SUR | 0.002 | 9.173 | -0.469 | 17848.695 | -0.022 | 6.664 | -0.223 | $3.1\times10^6$ |
| ($t$-values) | 1.466 | 11.802 | -10.728 | | -7.893 | 6.718 | -4.053 | |
| | **UK** | | | | **Pooled** | | | |
| FM-CPR | -0.007 | 7.697 | -0.397 | 16160.237 | | | | |
| ($t$-values) | -2.734 | 5.243 | -5.442 | | | | | |
| FM-SOLS | -0.006 | 8.908 | -0.465 | 14388.146 | -0.015 | 13.329 | -0.653 | 27173.094 |
| ($t$-values) | -2.746 | 6.946 | -7.305 | | -8.749 | 21.481 | -19.390 | |
| FM-SUR | -0.005 | 6.754 | -0.352 | 14720.830 | -0.013 | 13.207 | -0.652 | 24864.948 |
| ($t$-values) | -3.041 | 6.969 | -7.226 | | -19.738 | 46.701 | -41.859 | |

Table 1: FM-CPR, FM-SOLS, FM-SUR and pooled FM-SOLS and FM-SUR estimation results for Equation (1). The estimated turning points TP are computed as $\exp\left(-\frac{\hat{\beta}_1}{2\hat{\beta}_2}\right)$.

UK.[14]

Table 1 shows the results of estimating the quadratic EKC (1) using both individual country FM-CPR (as used in Wagner, 2015) and the two SUR estimators discussed in Section 2, FM-SUR and FM-SOLS, for the seven countries listed above. In addition, the lower right block of the table contains the results when estimating the EKC "fully" pooled, allowing only for country specific intercepts (this form of pooling is referred to as (P) in Section 2.1).[15] The following messages

[14] This is slightly different from Wagner (2015) who finds evidence for a quadratic EKC for $CO_2$ emissions for only four out of the seven countries above: Austria, Belgium, Finland and the UK. These differences may stem from the different sample range and/or the fact that the $CO_2$ emissions data have been updated.

[15] In formal terms, estimation of (1) is performed under the restrictions $\delta_i = \delta$, $\beta_{1,i} = \beta_1$ and $\beta_{2,1} = \beta_2$ for $i = 1, \ldots, 7$.

emerge from the table: First, the estimated coefficients (all significant with "correct" signs) and a fortiori the estimated turning points do usually not differ strongly across the three methods for each country. The exception here is Austria where the FM-CPR turning point is about twice as large as the FM-SUR and FM-SOLS turning points. For Switzerland, the turning point is estimated far outside the sample range. This is related to the fact that, see Figure 2, per capita $CO_2$ emissions are essentially constant since about 1980 in Switzerland. Second, with respect to the two SUR estimators the differences are mostly very minor, with the one exception being Finland. For this reason we focus from now on on the FM-SUR estimator in the discussion.[16] Third, the estimated coefficients and consequently the estimated turning points differ substantially across countries and this heterogeneity can – by construction – not be captured by the pooled, i.e., almost panel-type, estimation results in the lower right block. This finding highlights that commonly used panel methods need to be considered very carefully, or maybe not used at all.[17]

The results from Table 1 are displayed graphically in Figures 1 and 2. The first figure displays the estimated EKCs, given by using 140 equidistant values for the explanatory variable from the range of log GDP per capita associated with values of the time trend ranging from 1,...,140 and inserting these values in Equation (1) using the coefficient estimates obtained from both FM-CPR (solid with x-marks) and FM-SUR (solid). Additionally the graphs include the scatter plots between log GDP per capita and log $CO_2$ emissions per capita. The similar coefficient estimates translate, as expected or in fact necessary, into very similar estimated EKCs. Figure 2 displays the actual values of log per capita $CO_2$ emissions with the fitted values obtained from both FM-CPR and FM-SUR estimation. Clearly, the two fitted value lines corresponding to FM-CPR and FM-SUR are very close to each other for all countries, with the still small but relatively largest differences for Austria (for which also the estimated turning point differs most between the two methods). In general the fit is very good, especially for the period since the second world war.

Performing the poolability tests (P), (S) and (T) described in Section 2.1 and in more detail in Appendix A.1 for the considered seven countries leads throughout to rejections of the respective null hypotheses for both tests, i.e., the tests based on the FM-SOLS estimator (12) and the FM-

[16]The similarity of the findings with both the FM-SOLS and the FM-SUR estimators is made clearly visible in Figures 7 and 8 in Appendix B.

[17]Building upon the seminal work of Phillips and Moon (1999), de Jong and Wagner (2016) consider a panel version of FM-type estimators for panels of cointegrating polynomial regressions under the assumption of cross-sectional independence. Under appropriate assumptions it may be the case that the pooled estimates converge to "average coefficients", see Phillips and Moon (1999) for details. These issues remain to be studied for the cointegrating polynomial regression case.

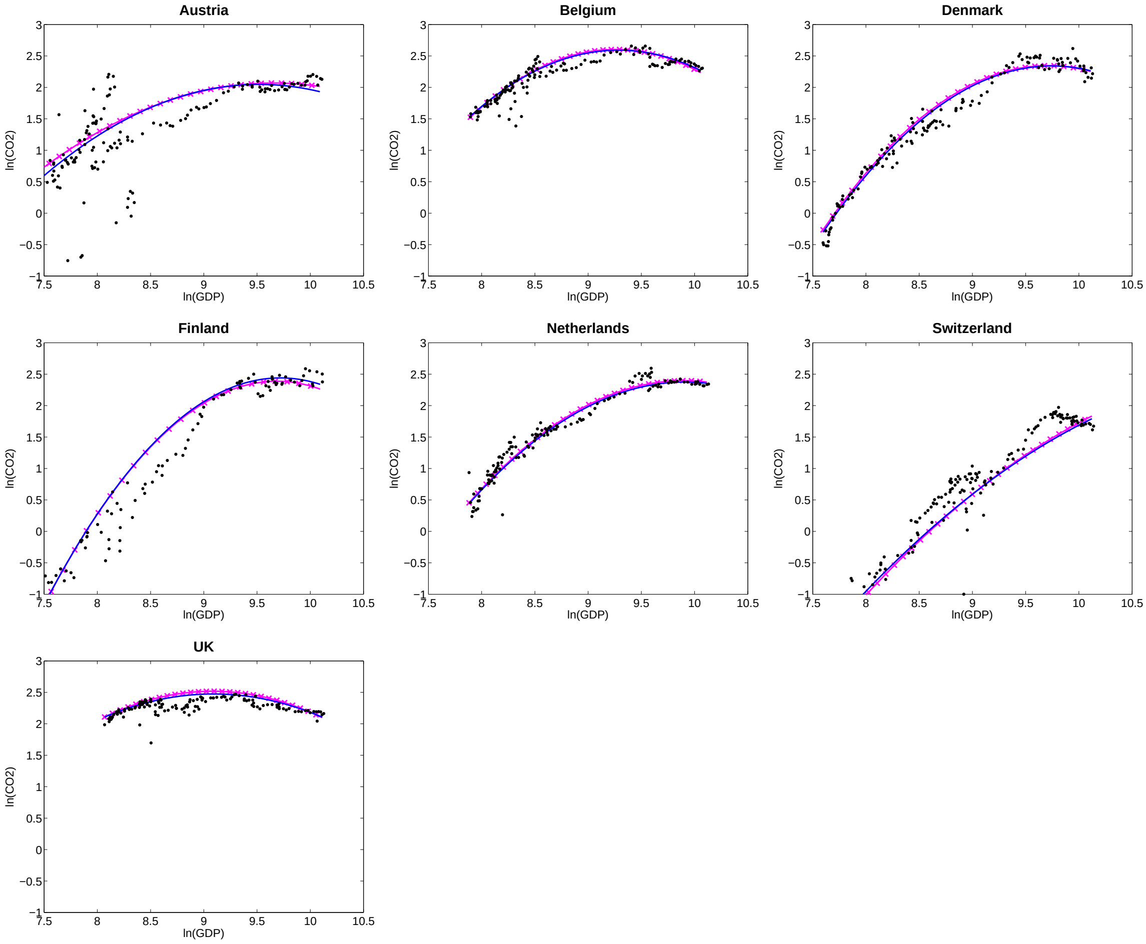


Figure 1: EKC estimation results for Equation (1): scatter plot and EKC. The dots show the pairs of observations of $\ln(GDP)$ per capita and $\ln(CO_2)$ emissions per capita. The lines show results based on inserting 140 equidistant points from the sample range of $\ln(GDP)$ per capita, with corresponding values of the linear trend given by $t = 1, \ldots, 140$ in the estimated relationship (1). The solid lines with x-marks correspond to the FM-CPR estimates and the solid lines to the FM-SUR estimates.

SUR estimator (13). For the hypothesis (P) this is already expected, given the cross-country heterogeneity of the unrestricted estimates, compare again the results in Table 1. The prize to be paid when applying pooled estimation, allowing only for country specific intercepts, despite this restriction being rejected, is clearly visible when looking at Figures 3 and 4, which are similar in structure to Figures 1 and 2. The fitted fully pooled EKC only fits well - and is consequently very similar to the FM-SUR EKC – for Denmark. For the other six countries the differences are partly enormous, both with respect to slope and shape. These differences translate directly into partly drastic reductions of fit, when considering the fitted value graphs in Figure 4. Thus, testing for group-wise poolability and potentially group-wise pooled estimation, as outlined in Section 2.2, are

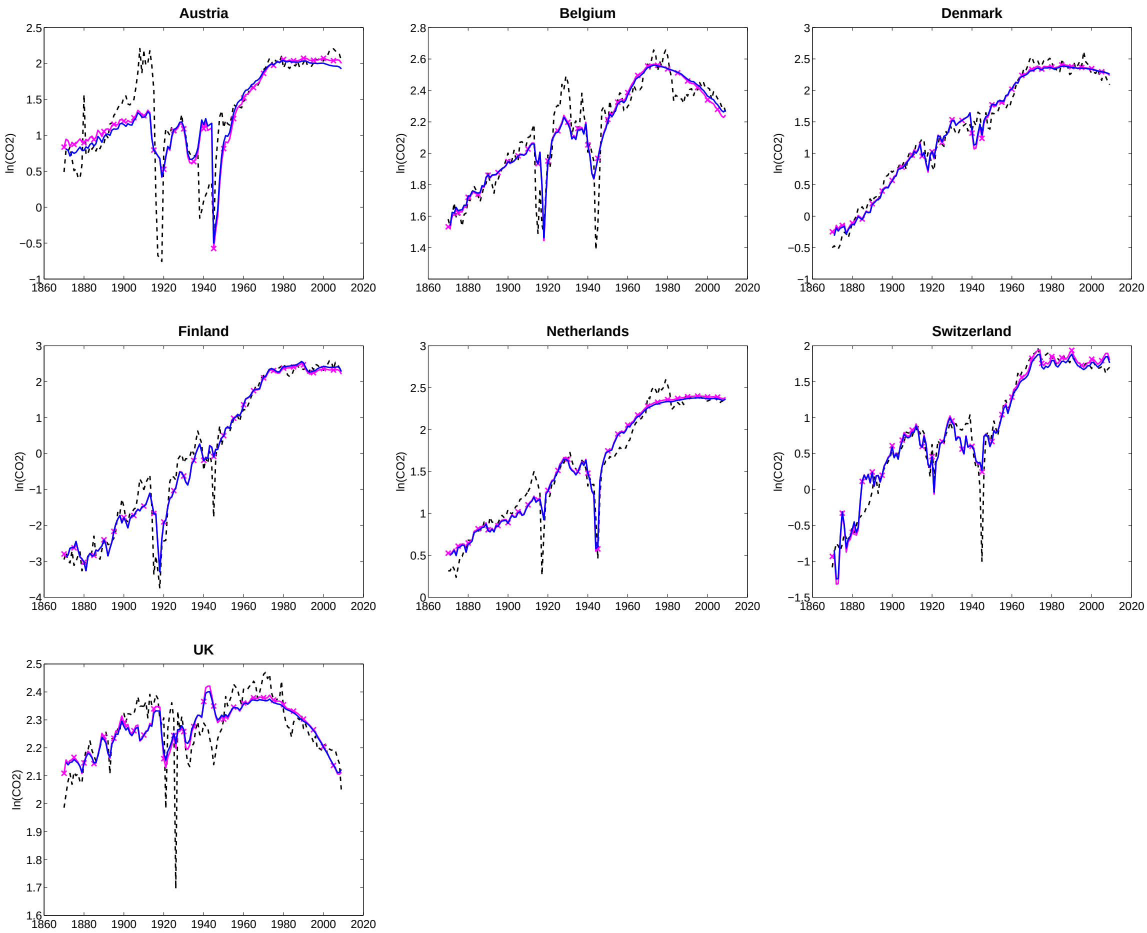


Figure 2: EKC estimation results for Equation (1): actual and fitted values. The dashed lines show the actual values of $\ln(CO_2)$ per capita emissions, the solid lines with x-marks the FM-CPR fitted values and the solid lines the FM-SUR fitted values.

the logical next steps.

In many applications the researcher may have some prior knowledge concerning candidates for group-wise pooling. To a certain extent this is also the case here, as one expects that very similar countries, e.g., Belgium and the Netherlands, may have very similar EKCs. Here, however, we pursue a more exploratory approach and test for the discussed three forms of pooling – (P), (S) and (T) – in all possible sub-groups. This means that we test for these forms of poolability in all possible 21 country-pairs, 35 country-triples and so on.[18] The results are given in Table 2 and

[18]Note that we test for the three forms of poolability using only data for the subset of countries under investigation. We do not perform all possible tests of group-wise poolability in all possible partitions into multiple subgroups using the data for all seven countries. Doing that would entail a rather large number of tests to be performed. Let us stress also that the approach is to be understood exploratory, since neither of the complications resulting from multiple testing is even addressed, let alone solved.

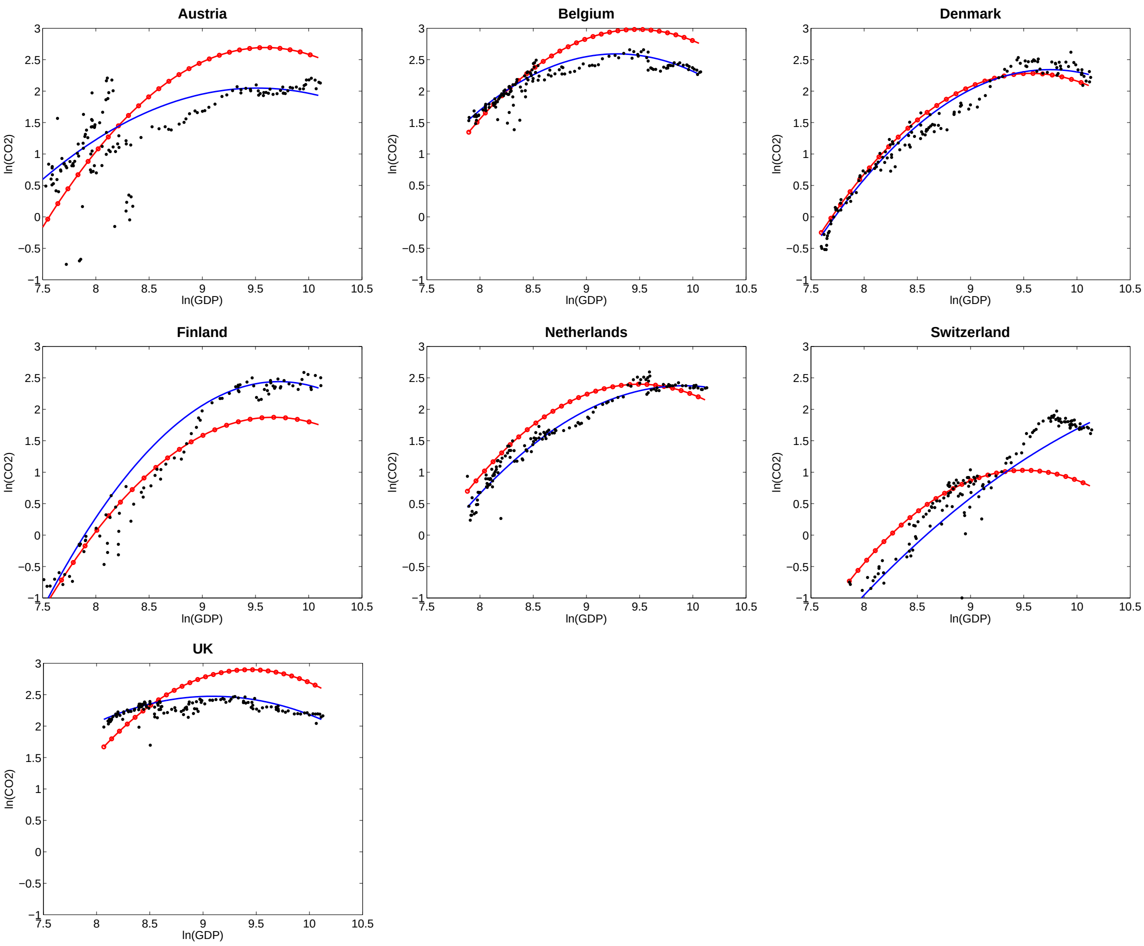


Figure 3: EKC estimation results for Equation (1): scatter plot and EKC. The solid lines correspond to the FM-SUR estimates and the solid lines with o-marks to the pooled FM-SUR estimates. For further explanations see notes to Figure 1.

Table 6 in Appendix B. Table 2 contains the numbers of groups of the respective sizes for which the corresponding poolability hypothesis cannot be rejected, with the group members displayed in Table 6. As for the coefficients, also for the tests the differences are minor between the FM-SOLS and FM-SUR results and thus we again focus again on the results obtained with FM-SUR. The full pooling hypothesis (P) is not rejected only for the pair Denmark and Finland. The slope parameters $\beta_1$ and $\beta_2$ can be pooled for (the pooling hypothesis (S) is not rejected for) six country-pairs and three country-triples. With respect to the trend parameters there are four country groups of size four, for which the trend slope can be pooled. In all three of these groups Austria and Denmark are included.

We take the above results as starting point to estimate the EKC for the seven considered countries

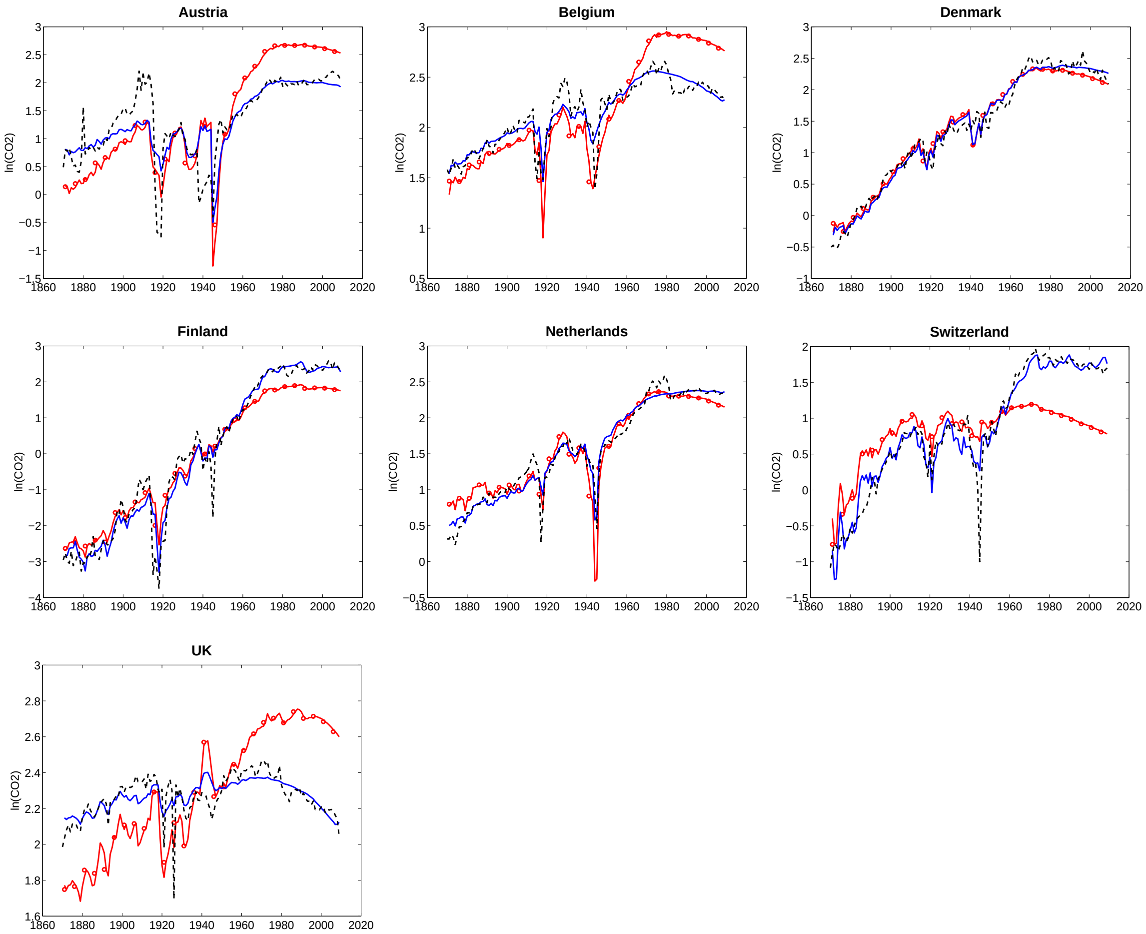


Figure 4: EKC estimation results for Equation (1): actual and fitted values. The dashed lines show the actual values, the solid lines the FM-SUR fitted values and the solid lines with o-marks the pooled FM-SUR fitted values.

in a group-wise pooled fashion. In particular we consider: the trend slope pooled in three groups, comprising Austria, Denmark, Finland and Switzerland; Belgium and the UK; and the Netherlands respectively. The slope parameters are pooled in four groups, given by Belgium, the Netherlands and the UK; Denmark and Finland; Austria; and Switzerland.[19] Table 3 displays the estimation results. As observed up to now, again the estimates are very similar for the now group-wise pooled FM-SOLS and FM-SUR estimates. Looking at the coefficients in the individual groups it can be clearly seen that the group-wise pooled estimates are – almost by construction when using group-wise pooled least squares estimation – close to the averages of the country specific estimates given

[19]We take this group of four countries, since for this group the poolability hypothesis is not rejected also for all subgroups of two or three countries of these four countries. The choice is made using similar arguments also for the slope parameters. Poolability of the slope parameters is not rejected for all pairs of countries of the triple comprising Belgium, the Netherlands and the UK.

| $k$ | 2 | 3 | 4 | 5 | 6 | 7 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Total nr. of groups of size $k$ | 21 | 35 | 35 | 21 | 7 | 1 | 21 | 35 | 35 | 21 | 7 | 1 |
| | **FM-SOLS** | | | | | | **FM-SUR** | | | | | |
| Linear Trend & Stochastic Regressors (P) | 1 | | | | | | 1 | | | | | |
| Stochastic Regressors (S) | 5 | 2 | | | | | 6 | 4 | | | | |
| Linear Trend (T) | 11 | 8 | 2 | | | | 12 | 9 | 3 | | | |

Table 2: Testing for group-wise poolability of subsets of coefficients. The numbers indicate the number of groups of size $k$ for which the indicated null hypothesis of group-wise poolability is not rejected. The members of the groups are given in Table 6 in Appendix B. Empty entries correspond to zeros. The left column displays the results for the FM-SOLS test statistic (12) and the right column displays the results for the FM-SUR test statistic (13). Individual test decisions are performed at the 1% significance level.

in Table 1. Of course, group-wise pooled estimation is not simply *mean-group* estimation, and thus the group-wise pooled coefficients estimates do not simply coincide with the averages. The same observations as for the coefficients hold, of course again by implication, for the estimated turning points.

The benefit of group-wise pooling becomes clearly visible when considering the results graphically in Figures 5 and 6. These two figures, again similar in structure to Figures 1 and 2, show clearly that imposing non-rejected group-wise poolability restrictions in group-wise pooled FM-SUR estimation (solid lines with square symbols) leads to very similar estimates of the EKCs compared to non-pooled FM-SUR estimation (solid lines). Importantly, also the (unavoidable) reduction in fit is negligible (see Figure 6), with the exception of the UK to some extent. Recall for comparison the drastic reduction in fit when pooling all slope and trend coefficients over all countries displayed in Figures 3 and 4.[20] Group-wise pooling of a form adapted to the situation leads to a sizeable reduction of the number of parameters to be estimated, in our case from 28 to 18, without any clearly visible losses in terms of approximation quality. Unthoughtful global pooling, i.e., panel-type estimation, leads to drastically worse results. These findings illustrate that a seemingly unrelated CPR approach is indeed very useful for the analysis of the EKC and similar relationships in situations with multi-country or multi-regional data where the cross-sectional dimension is small.

[20]Figures 9 and Figure 10 in Appendix B compare the group-wise pooled and pooled FM-SUR results. These two figures clearly make the same point as the figures in the main text, but contrasting group-wise pooled and pooled estimation results in the same figure highlights the benefits of group-wise pooling compared to pooling nicely.

| | $\hat{\delta}_{n_1}$ | $\hat{\delta}_{n_2}$ | $\hat{\delta}_{n_3}$ |
|---|---|---|---|
| Countries | **AT-DK-FI-CH** | **BE-UK** | **NL** |
| FM-SOLS | -0.023 | -0.008 | 0.001 |
| ($t$-values) | -9.215 | -7.290 | -1.032 |
| FM-SUR | -0.020 | -0.008 | 0.002 |
| ($t$-values) | -17.744 | -11.372 | 2.473 |

| | $\hat{\beta}_{1,m_1}$ | $\hat{\beta}_{2,m_1}$ | $\hat{\beta}_{1,m_2}$ | $\hat{\beta}_{2,m_2}$ | $\hat{\beta}_{1,m_3}$ | $\hat{\beta}_{2,m_3}$ | $\hat{\beta}_{1,m_4}$ | $\hat{\beta}_{2,m_4}$ |
|---|---|---|---|---|---|---|---|---|
| Countries | **BE-NL-UK** | | **DK-FI** | | **AT** | | **CH** | |
| FM-SOLS | 11.270 | -0.584 | 14.491 | -0.678 | 10.714 | -0.511 | 6.129 | -0.188 |
| ($t$-values) | 15.079 | -14.935 | 20.358 | -16.948 | 5.397 | -4.600 | 6.323 | -3.497 |
| TP | 15540.291 | | 37832.799 | | 35459.781 | | $1.2\times10^7$ | |
| FM-SUR | 10.464 | -0.542 | 14.887 | -0.720 | 10.191 | -0.495 | 7.351 | -0.267 |
| ($t$-values) | 22.322 | -21.899 | 35.055 | -30.992 | 7.166 | -6.131 | 8.638 | -5.750 |
| TP | 15676.923 | | 30887.742 | | 29412.644 | | $9.7\times10^5$ | |

Table 3: Group-wise pooled estimation results for Equation (1) using FM-SOLS and FM-SUR. The trend parameter $\delta$ is pooled in three groups (of sizes four, two and one) and the slope parameters $\beta_1$, $\beta_2$ are pooled in four groups (of sizes three, two and twice one). The estimated turning points TP are computed as $\exp\left(-\frac{\hat{\beta}_1}{2\hat{\beta}_2}\right)$.

## 4 Summary and Conclusions

We provide tools for multi-country (or multi-regional) cointegration analysis of the environmental Kuznets curve (EKC) by pursuing a seemingly unrelated cointegrating polynomial regressions (SUCPR) approach advocated by Hong and Wagner (2014). The approach can be applied also in other contexts in which inverted U-shaped relationships are studied, such as the intensity of use (IOU) relationship between GDP and energy or material intensity (see, e.g., Guzmán *et al*, 2005; Labson and Crompton, 1993).

The SUCPR approach advocated in this paper addresses three of the main challenges of the existing literature: First, it takes into account that powers of integrated processes are themselves not integrated processes and that consequently cointegration analysis of the EKC needs to resort to methods designed for this specific form of nonlinear relationship, labelled cointegrating polynomial regression by Wagner and Hong (2016). The implications of this fact for single country EKC analysis have been pointed out earlier in Wagner (2015); the present paper translates and extends this discussion to the multi-country data case. Second, it is not necessarily the case that, e.g., emissions and GDP data for different countries are independent of each other, an assumption typically made in the panel EKC literature. Third, furthermore the EKC relationship, if present,

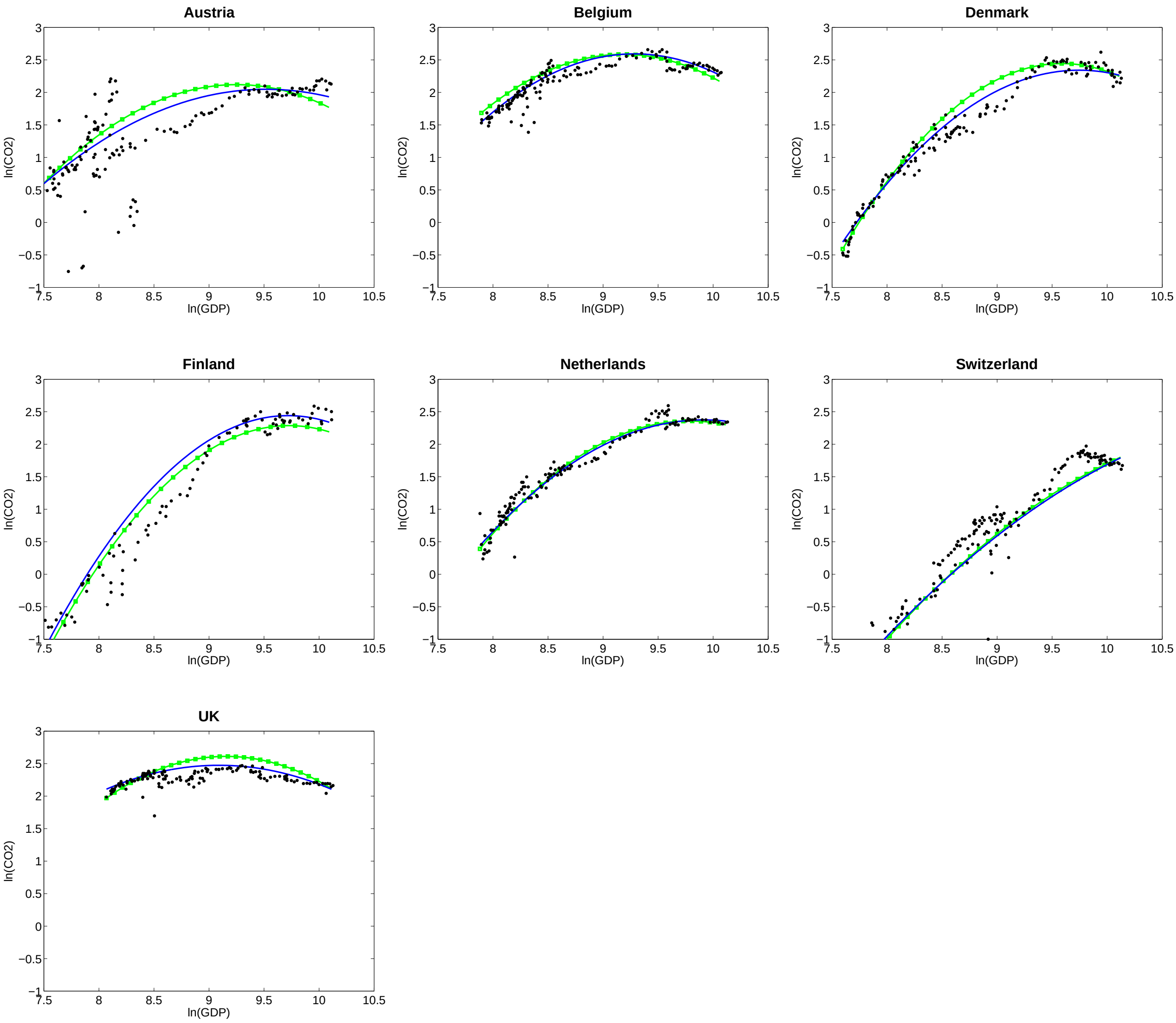


Figure 5: EKC estimation results for Equation (1): scatter plot and EKC. The solid lines correspond to the FM-SUR estimates and the solid lines with the square symbols to the group-wise pooled FM-SUR estimates. For further explanations see notes to Figure 1.

need not be identical (potentially up to country specific individual effects) across countries, which is the the key assumption underlying pooling which panel data analysis rests upon. A SUCPR approach (based on Hong and Wagner, 2014) addresses these three issues and provides new tools for group-wise poolability testing and, in case the restrictions are not rejected, correspondingly pooled estimation.

Developing poolability tests and correspondingly pooled estimators for general sets of restrictions is shown to be extremely useful in our application to $CO_2$ emissions data for seven early industrialized countries over the period 1870–2009. It turns out that the trend respectively slope parameters can

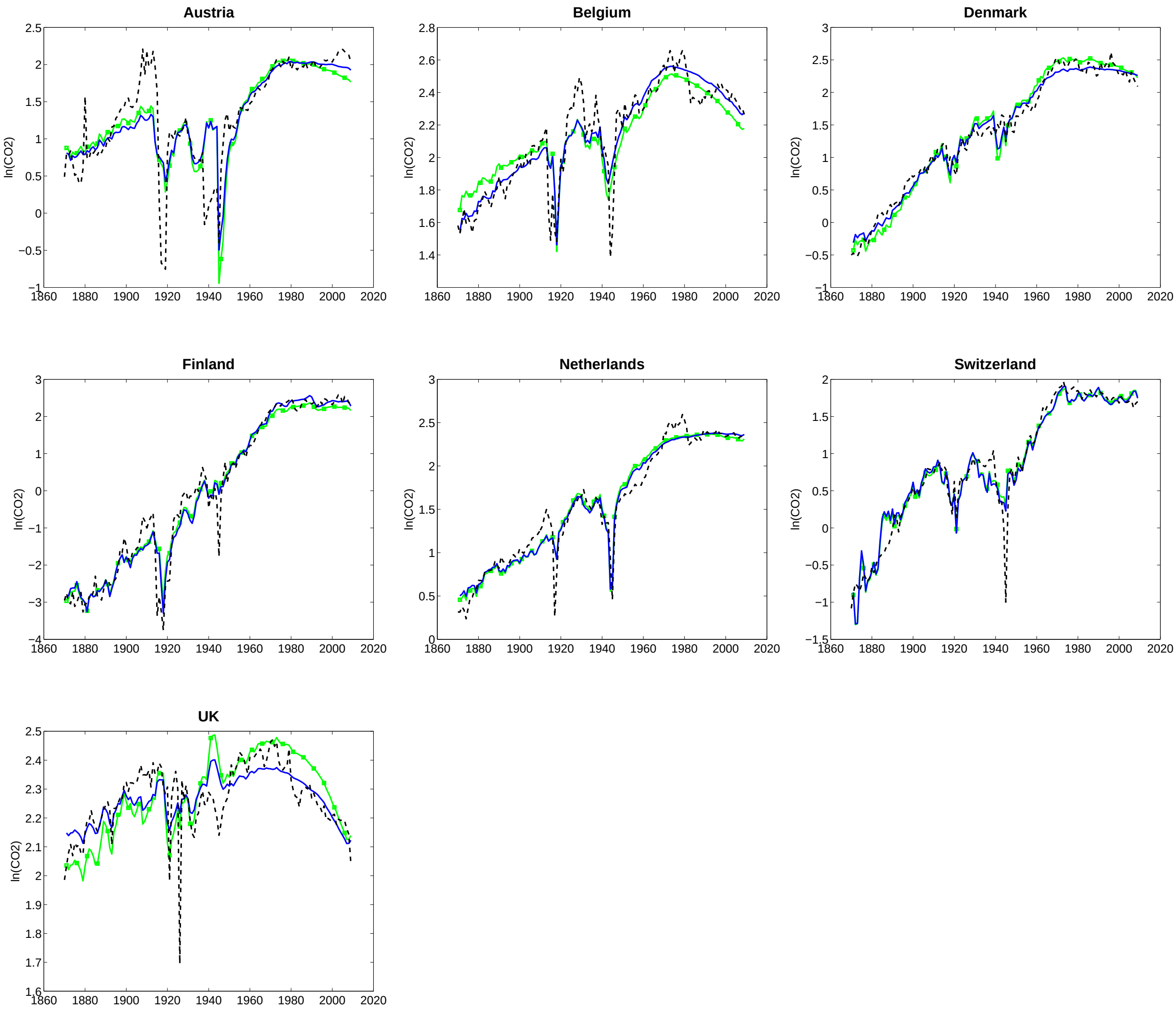


Figure 6: EKC estimation results for Equation (1): actual and fitted values. The dashed lines show the actual values, the solid lines the FM-SUR fitted values and the solid lines with square symbols the group-wise pooled FM-SUR fitted values.

be pooled over different country sub-groups, a situation that we label group-wise pooling. The results show that group-wise pooled estimation provides fits that are close to the fits from either individual country or unrestricted SUCPR estimation, whilst the number of parameters to be estimated is substantially reduced. Altogether, the simple reduced form SUCPR EKC analysis leads to very good fit, especially since the second world war, and meaningful estimates of the turning points. Performing SUCPR estimation in a fully pooled fashion with only country specific intercepts, by comparison leads to substantial losses in terms of fit. A major limitation of any SUR approach is the limitation to situations with a relatively small cross-sectional dimension. For data sets with large cross-sectional dimension panel data approaches will need to be pursued, with all advantages and disadvantages. For a first step in this direction see de Jong and Wagner (2016), who in turn build upon the seminal work of Phillips and Moon (1999) for the linear cointegration case.

The empirical results of this paper illustrate the usefulness of SUCPR analysis of the EKC, but the reduced form character of the analysis presented here dictates the necessary next steps of the research agenda: First, for certain applications it may be necessary to extend the methodology to allow for the inclusion of stationary regressors.[21] This is a pertinent issue in, e.g., IOU analysis. In case of substitution possibilities between different metals (see, e.g., Stuermer, 2016) or energetic resources, the inclusion of *relative prices* is of key importance to capture substitution elasticities. Note in this respect that the SUR approach also can be used to study EKC or IOU relationships for a set of different emissions variables or resource intensities for a given country or a small number of countries. This allows to study the interrelationships in a system of cointegrating polynomial regressions. Second, in particular for regional data it may be important to allow for the inclusion of common *aggregate* variables, i.e., technically speaking for the inclusion of common (nonstationary) regressors.[22] Third, it is always important to strive for extending the discussed methods to allow for a more structural analysis of EKC- or IOU-type relationships by considering more general specifications. Extensions along all three dimensions are or will be investigated in ongoing and planned research.

## Acknowledgments

Both authors acknowledge financial support from the Collaborative Research Center 823: *Statistical Modelling of Nonlinear Dynamic Processes*. The first author furthermore acknowledges research support from the Jubiläumsfonds of the Oesterreichische Nationalbank via several grants. The fruitful discussions at the Climate Econometrics Workshop in Aarhus, Denmark, have been very beneficial for this paper. The usual disclaimer applies.

[21] Pre-determined stationary regressors can be accommodated more easily than endogenous stationary regressors. Endogenous stationary regressors will require to construct an instrumental variables-type extension of the estimators discussed here. Even if an IV-type estimator is developed, the availability of valid and relevant instruments will, as always, be a challenge in actual applications.

[22] This may on a bigger scheme also be relevant for multi-country data, e.g., EU data with common EU-wide variables to be included. These could be related to common policies or regulations.

## References

Agras, J. and D. Chapman (1999). A Dynamic Approach to the Environmental Kuznets Curve Hypothesis. *Ecological Economics* **28**, 267–277.

Andreoni, J. and A. Levinson (2001). The Simple Analytics of the Environmental Kuznets Curve. *Journal of Public Economics* **80**, 269–286.

Antweiler, W., B.R. Copeland, and M.S. Taylor (2001). Is Free Trade Good for the Environment? *American Economic Review* **91**, 877–908.

Apergis, N. (2016). Environmental Kuznets Curves: New Evidence on both Panel and Country-Level $CO_2$ Emissions. *Energy Economics* **54**, 263–271.

Arrow, K.J., B. Polin, R. Costanza, P. Dasgupta, C. Folke, C.S. Holling, B.O. Jansson, S. Levin, K.G. Maler, C. Perrings, and D. Pimentel (1995). Economic Growth, Carrying Capacity, and the Environment. *Science* **268**, 520–521.

Auffhammer, M. and R. T. Carson (2008). Forecasting the Path of China's $CO_2$ Emissions using Province-level Information. *Journal of Environmental Economics and Management* **55**, 229–247.

Baek, J. (2015). Environmental Kuznets Curve for $CO_2$ Emissions: The Case of Arctic Countries. *Energy Economics* **50**, 13–17.

Bernard, J. T., M. Gavin, L. Khalaf, and M. Voia (2015). Environmental Kuznets Curve: Tipping Points, Uncertainty and Weak Identification. *Environmental and Resource Economics* **60**, 285–315.

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

Boden, T.A., G. Marland and R.J. Andres (2016). Global, Regional, and National Fossil-Fuel $CO_2$ Emissions. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

Bolt, J. and J. L. van Zanden (2014). The Maddison Project: Collaborative Research on Historical National Accounts. *The Economic History Review* **67**, 627–651.

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.

Brock, W.A. and M.S. Taylor (2010). The Green Solow Model. *Journal of Economic Growth* **15**, 127–153.

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 (2010). Tests for Nonlinear Cointegration. *Econometric Theory* **26**, 682–709.

Cropper, M. and C. Griffiths (1994). The Interaction of Population Growth and Environmental Quality. *American Economic Review* **84**, 250–254.

de Jong, R. M. and M. Wagner (2016). Panel Nonlinear Cointegration Analysis of the Environmental Kuznets Curve. *Mimeo.*

Dijkgraaf, E., and H. R. Vollebergh (2005). A Test for Parameter Homogeneity in $CO_2$ Panel EKC Estimations. *Environmental and Resource Economics* **32**, 229–239.

Dinda, S. (2005). A Theoretical Basis for the Environmental Kuznets Curve. *Ecological Economics* **53**, 403–413.

Dinda, S. and D. Coondoo (2006). Income and Emission: A Panel Data-Based Cointegration Analysis. *Ecological Economics* **57**, 167-181.

Esteve, V. and C. Tamarit (2012). Threshold Cointegration and Nonlinear Adjustment between CO2 and Income: The Environmental Kuznets Curve in Spain, 1857-2007. *Energy Economics* **34**, 2148–2156.

Fosten, J., B. Morley, and T. Taylor (2012). Dynamic Misspecification in the Environmental Kuznets Curve: Evidence from $CO_2$ and $SO_2$ Emissions in the United Kingdom. *Ecological Economics* **76**, 25–33.

Friedl, B. and M. Getzner (2003). Determinants of $CO_2$ Emissions in a Small Open Economy. *Ecological Economics* **45**, 133–148.

Galeotti, M., A. Lanza, and F. Pauli (2006). Reassessing the Environmental Kuznets Curve for $CO_2$ Emissions: A Robustness Exercise. *Ecological Economics* **57**, 152–163.

Grossman, G.M. and A.B. Krueger (1991). Environmental Impacts of a North American Free Trade Agreement. NBER Working paper No. 3914.

Grossman, 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.

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

Guzmán, J.I., T. Nishiyama and J.E. Tilton (2005). Trends in the Intensity of Copper Use in Japan since 1960. *Resources Policy* **30**, 21–27.

He, J. and P. Richard (2010). Environmental Kuznets Curve for $CO_2$ in Canada. *Ecological Economics* **69**, 1083–1093.

Hilton, F.G.H. and A. Levinson (1998). Factoring the Environmental Kuznets Curve: Evidence from Automobile Lead Emissions. *Journal of Environmental Economics and Management* **35**, 126–141.

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

Hong, S.H. and M. Wagner (2014). Seemingly Unrelated Cointegrating Polynomial Regressions: Fully Modified OLS Estimation and Inference. Mimeo.

Ibragimov, R. and P.C.B. Phillips (2008). Regression Asymptotics Using Martingale Convergence Methods. *Econometric Theory* **24**, 888–947.

Jalil, A. and S.F. Mahmud (2009). Environment Kuznets Curve for $CO_2$ Emissions: A Cointegration Analysis for China. *Energy Policy* **37**, 5167–5172.

Jansson, M. (2002). Consistent Covariance Matrix Estimation for Linear Processes. *Econometric Theory* **18**, 1449–1459.

Jones, L.E. and R.E. Manuelli (2001). Endogenous Policy Choice: The Case of Pollution and Growth. *Review of Economic Dynamics* **4**, 369–405.

Kahn, M. E. (1998). A Household Level Environmental Kuznets Curve. *Economics Letters* **59**, 269–273.

Kao, C. and M.-H. Chiang (2000). On the Estimation and Inference of a Cointegrated Regression in Panel Data. In Baltagi, B.H. (Ed.) *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, Elsevier, Amsterdam.

Kijima, M., K. Nishide, and A. Oyama (2010). Economic Models for the Environmental Kuznets Curve: A Survey. *Journal of Economic Dynamics and Control* **34**, 1187–1201.

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

Labson, B.S. and P.L. Crompton (1993). Common Trends in Economic Activity and Metals Demand: Cointegration and the Intensity of Use Debate. *Journal of Environmental Economics and Management* **25**, 147–161.

Lindmark, M. (2002). An EKC-pattern in Historical Perspective: Carbon Dioxide Emissions, Technology, Fuel Prices and Growth in Sweden 1870–1997. *Ecological Economics* **42**, 333–347.

List, J.A. and C.A. Gallet (1999). The Environmental Kuznets Curve: Does One Size Fit All? *Ecological Economics* **31**, 409–423.

Mark, N.C., M. Ogaki, and D. Sul (2005). Dynamic Seemingly Unrelated Cointegrating Regressions. *Review of Economic Studies* **72**, 797–820.

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.

Moon, H.R. (1999). A Note on Fully-Modified Estimation of Seemingly Unrelated Regressions Models with Integrated Regressors. *Economics Letters* **65**, 25–31.

Moon, H.R. and B. Perron (2005). Efficient Estimation of the SUR Cointegration Regression Model and Testing for Purchasing Power Parity. *Econometric Reviews* **23**, 293–323.

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 (1994). Automatic Lag Selection in Covariance Matrix Estimation. *Review of Economic Studies* **61**, 631–654.

Park, J.Y. and M. Ogaki (1991). Seemingly Unrelated Canonical Cointegrating Regressions. Mimeo.

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

Park, J.Y. and P.C.B. Phillips (1989). Statistical Inference in Regressions with Integrated Processes: Part II. *Econometric Theory* **5**, 95–131.

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.

Pedroni, P. (2000). Fully Modified OLS for Heterogeneous Cointegrated Panels. In Baltagi, B.H. (Ed.) *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, Elsevier, Amsterdam.

Perman, R. and D.I. Stern (2003). Evidence from Panel Unit Root and Cointegration Tests that the Environmental Kuznets Curve does not exist. *The Australian Journal of Agricultural and Resource Economics* **47**, 325–347.

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 H.R. Moon (1999). Linear Regression Limit Theory for Nonstationary Panel Data. *Econometrica* **67**, 1057–1111.

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

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

Romero-Avila, D. (2008). Questioning the Empirical Basis of the Environmental Kuznets Curve for $CO_2$: New Evidence from a Panel Stationarity Test Robust to Multiple Breaks and Cross-Dependence. *Ecological Economics* **64**, 559–574.

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

Selden, D.M. and D. Song (1995). Neoclassical Growth, the J-Curve for Abatement, and the Inverted U Curve for Pollution. *Journal of Environmental Economics and Management* **29**, 162–168.

Shafik, N. and S. Bandyopadhyay (1992). Economic Growth and Environmental Quality: Time-series and Cross-country Evidence. World Bank Policy Research Working Paper, WPS 904.

Shin, Y. (1994). A Residual Based Test for the Null of Cointegration Against the Alternative of No Cointegration. *Econometric Theory* **10**, 91–115.

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

Stokey, N.L. (1998). Are there Limits to Growth? *International Economic Review* **39**, 1–31.

Stuermer, M. (2016). 150 Years of Boom and Bust: What Drives Mineral Commodity Prices? *Macroeconomic Dynamics*. Forthcoming.

Stypka, O., Grabarczyk, P., Kawka, R. and M. Wagner (2016). "Linear" Fully Modified OLS Estimation of Cointegrating Polynomial Regressions. Mimeo.

Torras, M. and J. K. Boyce (1998). Income, Inequality, and Pollution: A Reassessment of the Environmental Kuznets Curve. *Ecological Economics* **25**, 147–160.

Uchiyama, K. (2016). Environmental Kuznets Curve Hypothesis and Carbon Dioxide Emissions. Springer Japan.

Vogelsang, T.J., and M. Wagner (2013). A Fixed-$b$ Perspective on the Phillips-Perron Tests. *Econometric Theory*, **29**, 609–628.

Wagner, M. (2012). The Phillips Unit Root Tests for Polynomials of Integrated Processes. *Economics Letters* **114**, 299–303.

Wagner, M. (2013). Residual Based Cointegration and Non-Cointegration Tests for Cointegrating Polynomial Regressions. Mimeo.

Wagner, M. (2015). The Environmental Kuznets Curve, Cointegration and Nonlinearity. *Journal of Applied Econometrics* **30**, 948–967.

Wagner, M. and S.H. Hong (2016). Cointegrating Polynomial Regressions: Fully Modified OLS Estimation and Inference. *Econometric Theory* **32**, 1289–1315.

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.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association* **57**, 348–368.

# Appendix A: More Details on Pooling

## Appendix A.1: Details for Pooling Cases (P), (S) and (T)

Here we consider three cases of pooling:

$$H_0^P : \begin{bmatrix} \delta_1 \\ \beta_{1,1} \\ \beta_{2,1} \end{bmatrix} = \begin{bmatrix} \delta_2 \\ \beta_{1,2} \\ \beta_{2,2} \end{bmatrix} = \cdots = \begin{bmatrix} \delta_N \\ \beta_{1,N} \\ \beta_{2,N} \end{bmatrix} \tag{31}$$

$$H_0^S : \begin{bmatrix} \beta_{1,1} \\ \beta_{2,1} \end{bmatrix} = \begin{bmatrix} \beta_{1,2} \\ \beta_{2,2} \end{bmatrix} = \cdots = \begin{bmatrix} \beta_{1,N} \\ \beta_{2,N} \end{bmatrix} \tag{32}$$

$$H_0^T : \delta_1 = \delta_2 = \cdots = \delta_N. \tag{33}$$

The corresponding restriction matrices for the Wald-type test are given by:

$$R^P = \begin{bmatrix} (0_{3\times1}, I_3) & (0_{3\times1}, -I_3) & 0_{3\times4} & \dots & 0_{3\times4} \\ \vdots & 0_{3\times4} & (0_{3\times1}, -I_3) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0_{3\times4} \\ (0_{3\times1}, I_3) & 0_{3\times4} & \dots & 0_{3\times4} & (0_{3\times1}, -I_3) \end{bmatrix} \in \mathbb{R}^{3(N-1)\times4N}, \; r = 0_{3(N-1)\times1}$$

$$R^S = \begin{bmatrix} (0_{2\times2}, I_2) & (0_{2\times2}, -I_2) & 0_{2\times4} & \dots & 0_{2\times4} \\ \vdots & 0_{2\times4} & (0_{2\times2}, -I_2) & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0_{2\times4} \\ (0_{2\times2}, I_2) & 0_{2\times4} & \dots & 0_{2\times4} & (0_{2\times2}, -I_2) \end{bmatrix} \in \mathbb{R}^{2(N-1)\times4N}, \; r = 0_{2(N-1)\times1}$$

$$R^T = \begin{bmatrix} (0, 1, 0_{1\times2}) & (0, -1, 0_{1\times2}) & 0_{1\times4} & \dots & 0_{1\times4} \\ \vdots & 0_{1\times4} & (0, -1, 0_{1\times2}) & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0_{1\times4} \\ (0, 1, 0_{1\times2}) & 0_{1\times4} & \dots & 0_{1\times4} & (0, -1, 0_{1\times2}) \end{bmatrix} \in \mathbb{R}^{(N-1)\times4N}, \; r = 0_{(N-1)\times1}$$

In case the respective null hypotheses are not rejected, correspondingly pooled estimation is the logical next step to reap the possible efficiency gains from reducing the number of parameters to be estimated. This basically entails a corresponding redefinition of the regressor matrices, the parameter vectors; and for the asymptotic analysis the weighting matrices and the limit processes. We discuss the three given cases in turn and start by defining the necessary adapted quantities:

(P) :

$$Z^* := \begin{bmatrix} Z_1^{*\prime} \\ Z_2^{*\prime} \\ \vdots \\ Z_T^{*\prime} \end{bmatrix}, \quad Z_t^* := \begin{bmatrix} I_N \\ X_t^* \end{bmatrix} \quad \theta^P := \begin{bmatrix} c_1 \\ \vdots \\ c_N \\ \delta \\ \beta_1 \\ \beta_2 \end{bmatrix},$$

$$X_t^* := \begin{bmatrix} t & t & \dots & t \\ x_{1,t} & x_{2,t} & \dots & x_{N,t} \\ x_{1,t}^2 & x_{2,t}^2 & \dots & x_{N,t}^2 \end{bmatrix}$$

$$G^* := \operatorname{diag}\left(T^{-1/2} \cdot I_N, T^{-3/2}, T^{-1}, T^{-3/2}\right)$$

$$J^*(r) := \begin{bmatrix} I_N \\ \mathbf{B}_N^*(r) \end{bmatrix}, \quad \mathbf{B}_N^*(r) := \begin{bmatrix} r & \dots & r \\ B_{v_1}(r) & \dots & B_{v_N}(r) \\ B_{v_1}^2(r) & \dots & B_{v_N}^2(r) \end{bmatrix}$$

(S) :

$$\tilde{Z} := \begin{bmatrix} \tilde{Z}_1' \\ \tilde{Z}_2' \\ \vdots \\ \tilde{Z}_T' \end{bmatrix}, \quad \tilde{Z}_t := \begin{bmatrix} D_{1,t} & 0_{2\times 1} & \dots & 0_{2\times 1} \\ 0_{2\times 1} & D_{2,t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_{2\times 1} \\ 0_{2\times 1} & \dots & 0_{2\times 1} & D_{N,t} \\ x_{1,t} & x_{2,t} & \dots & x_{N,t} \\ x_{1,t}^2 & x_{2,t}^2 & \dots & x_{N,t}^2 \end{bmatrix}, \quad \theta^S := \begin{bmatrix} c_1 \\ \delta_1 \\ \vdots \\ c_N \\ \delta_N \\ \beta_1 \\ \beta_2 \end{bmatrix},$$

$$A_i^S := \left(\hat{\Delta}_{vu}^{+}\right)^{i,i} \begin{bmatrix} T \\ 2\sum_{t=1}^{T} x_{i,t} \end{bmatrix}, \quad \tilde{A}_i^S := \left(\hat{\Delta}_{vu}^{+}\right)^{i,\cdot} \left(\hat{\Omega}_{u.v}^{-1}\right)^{\cdot,i} \begin{bmatrix} T \\ 2\sum_{t=1}^{T} x_{i,t} \end{bmatrix}$$

$$\tilde{G} := \operatorname{diag}\left(I_N \otimes G_D, G_X\right), \quad G_D = \operatorname{diag}\left(T^{-1/2}, T^{-3/2}\right), \quad G_X = \operatorname{diag}\left(T^{-1}, T^{-3/2}\right)$$

$$\tilde{J}(r) := \begin{bmatrix} D_N(r) \\ \mathbf{B}_N(r) \end{bmatrix}, \quad D_N(r) := I_N \otimes \begin{bmatrix} 1 \\ r \end{bmatrix}, \quad \mathbf{B}_N(r) := \begin{bmatrix} B_{v_1}(r) & \dots & B_{v_N}(r) \\ B_{v_1}^2(r) & \dots & B_{v_N}^2(r) \end{bmatrix}$$

(T) :

$$\check{Z} := \begin{bmatrix} \check{Z}_1' \\ \check{Z}_2' \\ \vdots \\ \check{Z}_T' \end{bmatrix}, \quad \check{Z}_t := \begin{bmatrix} \check{X}_{1,t} & 0_{4\times 1} & \dots & 0_{4\times 1} \\ 0_{4\times 1} & \check{X}_{2,t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_{4\times 1} \\ 0_{4\times 1} & \dots & 0_{4\times 1} & \check{X}_{N,t} \\ t & t & \dots & t \end{bmatrix}, \quad \theta^T := \begin{bmatrix} c_1 \\ \beta_{1,1} \\ \beta_{2,1} \\ \vdots \\ c_N \\ \beta_{1,N} \\ \beta_{2,N} \\ \delta \end{bmatrix},$$

$$\check{X}_{i,t} := \begin{bmatrix} 1 \\ x_{i,t} \\ x_{i,t}^2 \end{bmatrix}, \quad A^T := \begin{bmatrix} A_1^T \\ A_2^T \\ \vdots \\ A_N^T \end{bmatrix}, \quad \check{A}^T := \begin{bmatrix} \check{A}_1^T \\ \check{A}_2^T \\ \vdots \\ \check{A}_N^T \end{bmatrix},$$

$$A_i^T := \left(\hat{\Delta}_{vu}^+\right)^{i,i} \begin{bmatrix} 0 \\ T \\ 2\sum_{t=1}^{T} x_{i,t} \end{bmatrix}, \quad \check{A}_i^T := \left(\hat{\Delta}_{vu}^+\right)^{i,\cdot} \left(\hat{\Omega}_{u.v}^{-1}\right)^{\cdot,i} \begin{bmatrix} 0 \\ T \\ 2\sum_{t=1}^{T} x_{i,t} \end{bmatrix}$$

$$\check{G} := \operatorname{diag}\left(I_N \otimes \check{G}_1, T^{-3/2}\right), \quad \check{G}_1 := \operatorname{diag}\left(T^{-1/2}, T^{-1}, T^{-3/2}\right)$$

$$\check{J}(r) := \begin{bmatrix} \check{\mathbf{B}}_{v_1}(r) & & \\ & \ddots & \\ & & \check{\mathbf{B}}_{v_N}(r) \\ r & \dots & r \end{bmatrix}, \quad \check{\mathbf{B}}_{v_i}(r) := \begin{bmatrix} 1 \\ B_{v_i}(r) \\ B_{v_i}^2(r) \end{bmatrix}$$

**Corollary 1 (Essentially Hong and Wagner 2014, Corollaries 1 and 2)** *Let $y_t$ be generated by* (1) *with the assumptions listed in place and where the pooling restrictions considered in either* $(P)$, $(S)$ *or* $(T)$ *are valid. Furthermore, assume again that long run variance estimation is performed consistently. Then for the three considered cases the FM-SOLS and FM-SUR estimators are, using the quantities defined above, given by:*

$$\theta^{*P}_{\text{FM-SOLS}} := \left(Z^{*\prime} Z^{*}\right)^{-1} \left( Z^{*\prime} y^{+} - \begin{bmatrix} 0_{N\times 1} \\ \sum_{i=1}^{N} A_i^{*} \end{bmatrix} \right), \tag{34}$$

$$\theta^{*P}_{\text{FM-SUR}} := \left(Z^{*\prime} \left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right) Z^{*}\right)^{-1} \left( Z^{*\prime} \left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right) y^{+} - \begin{bmatrix} 0_{N\times 1} \\ \sum_{i=1}^{N} \tilde{A}_i^{*} \end{bmatrix} \right), \tag{35}$$

$$\tilde{\theta}^{S}_{\text{FM-SOLS}} := \left(\tilde{Z}' \tilde{Z}\right)^{-1} \left( \tilde{Z}' y^{+} - \begin{bmatrix} 0_{2N\times 1} \\ \sum_{i=1}^{N} A_i^{S} \end{bmatrix} \right), \tag{36}$$

$$\tilde{\theta}^{S}_{\text{FM-SUR}} := \left(\tilde{Z}' \left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right) \tilde{Z}\right)^{-1} \left( \tilde{Z}' \left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right) y^{+} - \begin{bmatrix} 0_{2N\times 1} \\ \sum_{i=1}^{N} \tilde{A}_i^{S} \end{bmatrix} \right), \tag{37}$$

$$\check{\theta}^{T}_{\text{FM-SOLS}} := \left(\check{Z}' \check{Z}\right)^{-1} \left( \check{Z}' y^{+} - \begin{bmatrix} A^{T} \\ 0 \end{bmatrix} \right), \tag{38}$$

$$\check{\theta}^{T}_{\text{FM-SUR}} := \left(\check{Z}' \left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right) \check{Z}\right)^{-1} \left( \check{Z}' \left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right) y^{+} - \begin{bmatrix} \check{A}^{T} \\ 0 \end{bmatrix} \right). \tag{39}$$

*For $T \to \infty$ the estimators are consistent with the following limiting distributions:*

$$G^{*-1}\left(\theta^{*P}_{\text{FM-SOLS}} - \theta^{P}\right) \Rightarrow \left(\int J^{*} J^{*\prime}\right)^{-1} \int J^{*} dB_{u.v}, \tag{40}$$

$$G^{*-1}\left(\theta^{*P}_{\text{FM-SUR}} - \theta^{P}\right) \Rightarrow \left(\int J^{*} \Omega_{u.v}^{-1} J^{*\prime}\right)^{-1} \int J^{*} \Omega_{u.v}^{-1} dB_{u.v}, \tag{41}$$

$$\tilde{G}^{-1}\left(\tilde{\theta}^{S}_{\text{FM-SOLS}} - \theta^{S}\right) \Rightarrow \left(\int \tilde{J} \tilde{J}'\right)^{-1} \int \tilde{J} dB_{u.v}, \tag{42}$$

$$\tilde{G}^{-1}\left(\tilde{\theta}^{S}_{\text{FM-SUR}} - \theta^{S}\right) \Rightarrow \left(\int \tilde{J} \Omega_{u.v}^{-1} \tilde{J}'\right)^{-1} \int \tilde{J} \Omega_{u.v}^{-1} dB_{u.v}, \tag{43}$$

$$\check{G}^{-1}\left(\check{\theta}^{T}_{\text{FM-SOLS}} - \theta^{T}\right) \Rightarrow \left(\int \check{J} \check{J}'\right)^{-1} \int \check{J} dB_{u.v}, \tag{44}$$

$$\check{G}^{-1}\left(\check{\theta}^{T}_{\text{FM-SUR}} - \theta^{T}\right) \Rightarrow \left(\int \check{J} \Omega_{u.v}^{-1} \check{J}'\right)^{-1} \int \check{J} \Omega_{u.v}^{-1} dB_{u.v}. \tag{45}$$

**Appendix A.2: Pooling the Trend Coefficient and the Coefficients of the Stochastic Regressors Over Different Subsets: Group-Wise Pooling**

**Proof of Proposition 3**:

Deriving the limiting distribution of FM-type estimators always commences from the limiting distribution of the underlying OLS and in the SUR case additionally the MSUR estimators. In our group-wise pooled setting these two estimators are defined as

$$\ddot{\theta}^{\mathrm{GW}}_{\mathrm{OLS}} := \left(\ddot{Z}'\ddot{Z}\right)^{-1}\ddot{Z}'y, \tag{46}$$

$$\ddot{\theta}^{\mathrm{GW}}_{\mathrm{MSUR}} := \left(\ddot{Z}'\left(I_T \otimes \hat{\Omega}_{uu}^{-1}\right)\ddot{Z}\right)^{-1}\ddot{Z}'\left(I_T \otimes \hat{\Omega}_{uu}^{-1}\right)y. \tag{47}$$

We start with the regressor cross-product matrix of the group-wise pooled OLS estimator $\ddot{\theta}^{\mathrm{GW}}_{\mathrm{OLS}}$:

$$\ddot{Z}_t\ddot{Z}_t' = \begin{bmatrix} I_N & \ddot{D}_t' & \ddot{X}_t' \\ \ddot{D}_t & \ddot{D}_t\ddot{D}_t' & \ddot{D}_t\ddot{X}_t' \\ \ddot{X}_t & \ddot{X}_t\ddot{D}_t' & \ddot{X}_t\ddot{X}_t' \end{bmatrix}, \tag{48}$$

where $\ddot{D}_t\ddot{D}_t' = \mathrm{diag}(\sum_{j\in I_{n_1}} D_jD_j', \ldots, \sum_{j\in I_{n_k}} D_jD_j')$, $\ddot{X}_t\ddot{X}_t' = \mathrm{diag}(\sum_{j\in I_{m_1}} X_jX_j', \ldots, \sum_{j\in I_{m_l}} X_jX_j')$ and

$$\ddot{D}_t\ddot{X}_t' = \begin{bmatrix} \sum_{i\in I_{n_1}}\sum_{j\in I_{m_1}} \delta_{ij}D_{i,t}X_{j,t}' & \cdots & \sum_{i\in I_{n_1}}\sum_{j\in I_{m_l}} \delta_{ij}D_{i,t}X_{j,t}' \\ \vdots & \ddots & \vdots \\ \sum_{i\in I_{n_k}}\sum_{j\in I_{m_1}} \delta_{ij}D_{i,t}X_{j,t}' & \cdots & \sum_{i\in I_{n_k}}\sum_{j\in I_{m_l}} \delta_{ij}D_{i,t}X_{j,t}' \end{bmatrix}, \tag{49}$$

with $\delta_{ij}$ denoting the Kronecker delta. The cross-product of the regressor matrix and the error term is given by

$$\ddot{Z}_tu_t = \left[u_t', \sum_{j\in I_{n_1}} D_{j,t}'u_{j,t}, \ldots, \sum_{j\in I_{n_k}} D_{j,t}'u_{j,t}, \sum_{j\in I_{m_1}} X_{j,t}'u_{j,t}, \ldots, \sum_{j\in I_{m_l}} X_{j,t}'u_{j,t}\right]'. \tag{50}$$

Similar arguments as used in Hong and Wagner (2014, Propositions 1 and 4) imply for the group-wise pooled OLS estimator $\ddot{\theta}^{\mathrm{GW}}_{\mathrm{OLS}}$ that

$$\begin{aligned} \ddot{G}^{-1}\left(\ddot{\theta}^{\mathrm{GW}}_{\mathrm{OLS}} - \theta^{\mathrm{GW}}\right) &= \ddot{G}^{-1}\left(\sum_{t=1}^T \ddot{Z}_t\ddot{Z}_t'\right)^{-1}\left(\sum_{t=1}^T \ddot{Z}_tu_t\right) \\ &= \left(\sum_{t=1}^T \begin{bmatrix} \ddot{G}_cI_N\ddot{G}_c & \ddot{G}_c\ddot{D}_t'\ddot{G}_D & \ddot{G}_c\ddot{X}_t'\ddot{G}_X \\ \ddot{G}_D\ddot{D}_t\ddot{G}_c & \ddot{G}_D\ddot{D}_t\ddot{D}_t'\ddot{G}_D & \ddot{G}_D\ddot{D}_t\ddot{X}_t'\ddot{G}_X \\ \ddot{G}_X\ddot{X}_t\ddot{G}_c & \ddot{G}_X\ddot{X}_t\ddot{D}_t'\ddot{G}_D & \ddot{G}_X\ddot{X}_t\ddot{X}_t'\ddot{G}_X \end{bmatrix}\right)^{-1}\left(\sum_{t=1}^T \begin{bmatrix} \ddot{G}_cu_t \\ \ddot{G}_D\ddot{D}_tu_t \\ \ddot{G}_X\ddot{X}_tu_t \end{bmatrix}\right) \\ &\Rightarrow \left(\int \ddot{J}\ddot{J}'\right)^{-1}\left(\int \ddot{J}dB_u + \mathring{A}^{\mathrm{GW}}\right), \end{aligned} \tag{51}$$

with

$$\mathring{A}^{\mathrm{GW}}_{m_j} := \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{m_j}\}} \cdot (\Delta_{vu})^{i,i} \cdot \begin{pmatrix} 1 \\ 2\int B_{v_i}(r) \end{pmatrix}. \tag{52}$$

For the group-wise pooled MSUR estimator $\ddot{\theta}^{\mathrm{GW}}_{\mathrm{MSUR}}$ we obtain

$$\ddot{G}^{-1}\left(\ddot{\theta}^{\mathrm{GW}}_{\mathrm{MSUR}} - \theta^{\mathrm{GW}}\right) = \left(\sum_{t=1}^{T} \ddot{G}\ddot{Z}_t\hat{\Omega}_{uu}^{-1}\ddot{Z}_t'\ddot{G}\right)^{-1} \sum_{t=1}^{T} \ddot{G}\ddot{Z}_t\hat{\Omega}_{uu}^{-1}u_t. \tag{53}$$

Denote $W = (w_{ij})_{i,j} \in \mathbb{R}^{N\times N}$ a symmetric matrix. This leads to

$$\ddot{D}_t W \ddot{D}_t' = \begin{bmatrix} \sum_{i\in I_{n_1}}\sum_{j\in I_{n_1}} w_{ij}D_{i,t}D_{j,t}' & \dots & \sum_{i\in I_{n_1}}\sum_{j\in I_{n_k}} w_{ij}D_{i,t}D_{j,t}' \\ \vdots & \ddots & \vdots \\ \sum_{i\in I_{n_k}}\sum_{j\in I_{n_1}} w_{ij}D_{i,t}D_{j,t}' & \dots & \sum_{i\in I_{n_k}}\sum_{j\in I_{n_k}} w_{ij}D_{i,t}D_{j,t}' \end{bmatrix} \tag{54}$$

and analogous expressions for $\ddot{X}_t W \ddot{X}_t'$ and $\ddot{D}_t W \ddot{X}_t'$. Furthermore, $\ddot{Z}_t W u_t$ is given by

$$\left[\sum_{j\in I} w_{1j}u_{j,t}, \dots, \sum_{j\in I} w_{Nj}u_{j,t}, \sum_{i\in I_{n_1}}\sum_{j\in I} w_{ij}D_{i,t}'u_{j,t}, \dots, \sum_{i\in I_{n_k}}\sum_{j\in I} w_{ij}D_{i,t}'u_{j,t}, \sum_{i\in I_{m_1}}\sum_{j\in I} w_{ij}X_{i,t}'u_{j,t}, \dots, \sum_{i\in I_{m_l}}\sum_{j\in I} w_{ij}X_{i,t}'u_{j,t}\right]'. \tag{55}$$

The quantities $\ddot{Z}_t\hat{\Omega}_{uu}^{-1}\ddot{Z}_t'$ and $\ddot{Z}_t\hat{\Omega}_{uu}^{-1}u_t$ in (53) are now given in more detail in (54) and (55) by setting $W = \hat{\Omega}_{uu}^{-1}$. Therefore, the limit of the first term of the above product is, when using a consistent long run variance estimator, given by

$$\sum_{t=1}^{T} \ddot{G}\ddot{Z}_t\hat{\Omega}_{uu}^{-1}\ddot{Z}_t'\ddot{G} \Rightarrow \begin{bmatrix} \int \Omega_{uu}^{-1} & \int \Omega_{uu}^{-1}\ddot{J}_D' & \int \Omega_{uu}^{-1}\ddot{J}_X' \\ \int \ddot{J}_D\Omega_{uu}^{-1} & \int \ddot{J}_D\Omega_{uu}^{-1}\ddot{J}_D' & \int \ddot{J}_D\Omega_{uu}^{-1}\ddot{J}_X' \\ \int \ddot{J}_X\Omega_{uu}^{-1} & \int \ddot{J}_X\Omega_{uu}^{-1}\ddot{J}_D' & \int \ddot{J}_X\Omega_{uu}^{-1}\ddot{J}_X' \end{bmatrix} = \int \ddot{J}\Omega_{uu}^{-1}\ddot{J}'. \tag{56}$$

Next consider the $m_i$ element of $\sum_{t=1}^{T} \ddot{G}_X\ddot{X}_t\hat{\Omega}_{uu}^{-1}u_t$, given by

$$\sum_{t=1}^{T} \ddot{G}_{X_{m_i}} \sum_{k\in I_{m_i}}\sum_{j\in I} (\hat{\Omega}_{uu}^{-1})^{k,j} X_{k,t}u_{j,t}. \tag{57}$$

For fixed $j \in I$ it follows that

$$(\hat{\Omega}_{uu}^{-1})^{k,j}\sum_{t=1}^{T} \ddot{G}_{X_{m_i}} \sum_{k\in I_{m_i}} X_{k,t}u_{j,t} \Rightarrow (\Omega_{uu}^{-1})^{k,j}\left(\int J_{X_{m_i}}dB_{u_j} + \sum_{k\in I_{m_i}} (\Delta_{vu})^{k,j}\begin{pmatrix} 1 \\ 2\int B_{v_k} \end{pmatrix}\right) \tag{58}$$

and therefore summation over $j = 1, \dots, N$ leads to

$$\sum_{t=1}^{T} \ddot{G}_{X_{m_i}} X_{m_i,t}\hat{\Omega}_{uu}^{-1}u_t \Rightarrow \int J_{X_{m_i}}\Omega_{uu}^{-1}dB_u + \sum_{k\in I_{m_i}} (\Delta_{vu})^{k,\cdot}(\Omega_{uu}^{-1})^{\cdot,k}\begin{pmatrix} 1 \\ 2\int B_{v_k} \end{pmatrix}. \tag{59}$$

Finally, stacking $i = m_1, \ldots, m_l$ gives

$$\sum_{t=1}^{T} \ddot{G}_X \ddot{X}_t \hat{\Omega}_{uu}^{-1} u_t \Rightarrow \int \ddot{J}_X \Omega_{uu}^{-1} dB_u + [\tilde{\mathring{A}}_{m_1}^{\text{GW}\prime}, \ldots, \tilde{\mathring{A}}_{m_l}^{\text{GW}\prime}]', \tag{60}$$

with

$$\tilde{\mathring{A}}_{m_j}^{\text{GW}} := \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{m_j}\}} \cdot (\Delta_{vu})^{i,\cdot} \cdot \left(\Omega_{uu}^{-1}\right)^{\cdot,i} \cdot \begin{pmatrix} 1 \\ 2 \int B_{v_i}(r) \end{pmatrix} \tag{61}$$

for $j = 1, \ldots, l$. The limits for the terms involving the deterministic components are derived in a similar way, with a difference being that they, of course, do not contain additive bias terms. Combining the terms gives the limiting distribution of the MSUR estimator $\ddot{\theta}_{\text{MSUR}}^{\text{GW}}$.

Let us now consider the group-wise pooled FM-SOLS estimator $\ddot{\theta}_{\text{FM-SOLS}}^{\text{GW}}$. Consider the term $\sum_t \ddot{G} \ddot{Z}_t u_t^+$ first, with the $m_i$ element given by

$$\sum_{t=1}^{T} \ddot{G}_{X_{m_i}} \sum_{k \in I_{m_i}} X_{k,t} \left( u_{k,t} - v_t' \hat{\Omega}_{vv}^{-1} \hat{\Omega}_{vu}^{\cdot,k} \right). \tag{62}$$

The limit of $\sum\limits_{t=1}^{T} \ddot{G}_{X_{m_i}} \sum\limits_{k \in I_{m_i}} X_{k,t} u_{k,t}$ is already known from the analysis of the group-wise pooled OLS estimator $\ddot{\theta}_{\text{OLS}}^{\text{GW}}$. Thus, it remains to consider

$$\sum_{t=1}^{T} \ddot{G}_{X_{m_i}} \sum_{k \in I_{m_i}} X_{k,t} \hat{\Omega}_{uv}^{k,\cdot} \hat{\Omega}_{vv}^{-1} v_t \Rightarrow \int J_{X_{m_i}} \Omega_{uv} \Omega_{vv}^{-1} dB_v + \sum_{k \in I_{m_i}} (\Delta_{vv})^{k,\cdot} \Omega_{uu}^{-1} \Omega_{vu}^{\cdot,k} \begin{pmatrix} 1 \\ 2 \int B_{v_k} \end{pmatrix}. \tag{63}$$

Combining the results for the two parts yields

$$\sum_{t=1}^{T} \ddot{G}_{X_{m_i}} X_{m_i,t} u_t^+ \Rightarrow \int J_{X_{m_i}} dB_{u \cdot v} + \sum_{k \in I_{m_i}} (\Delta_{vu}^+)^{k,k} \begin{pmatrix} 1 \\ 2 \int B_{v_k} \end{pmatrix}. \tag{64}$$

The result for $\sum_{t=1}^{T} \ddot{G}_X \ddot{X}_t u_t^+$ follows from stacking the results for $i = m_1, \ldots, m_l$. A similar result holds again for the terms involving the deterministic components, again without resultant bias terms. Next, note that by construction

$$\ddot{G}_{X_{m_i}} A_{m_i}^{\text{GW}} \Rightarrow \sum_{k \in I_{m_i}} (\Delta_{vu}^+)^{k,k} \begin{pmatrix} 1 \\ 2 \int B_{v_k} \end{pmatrix}. \tag{65}$$

Combining these expressions and inserting them in the definition of the group-wise pooled FM-SOLS estimator $\ddot{\theta}_{\text{FM-SOLS}}^{\text{GW}}$ gives the result.

For the group-wise pooled FM-SUR estimator $\ddot{\theta}^{\text{GW}}_{\text{FM-SUR}}$ the result follows from combining the results already obtained for $\ddot{\theta}^{\text{GW}}_{\text{MSUR}}$ and $\ddot{\theta}^{\text{GW}}_{\text{FM-SOLS}}$. First, we obtain

$$\sum_{t=1}^{T} \ddot{G}\ddot{Z}_t\hat{\Omega}^{-1}_{u\cdot v}\ddot{Z}_t'\ddot{G} \Rightarrow \int \ddot{J}\Omega^{-1}_{u\cdot v}\ddot{J}', \tag{66}$$

replacing $\hat{\Omega}^{-1}_{uu}$ by $\hat{\Omega}^{-1}_{u\cdot v}$ in (56). For the second term we obtain using similar arguments as in the FM-SOLS case

$$\sum_{t=1}^{T} \ddot{G}_{X_{m_i}} X_{m_i,t}\hat{\Omega}^{-1}_{u\cdot v}u_t^+ \Rightarrow \int J_{X_{m_i}}\Omega^{-1}_{u\cdot v}dB_{u\cdot v} + \sum_{k\in I_{m_i}} (\Delta^+_{vu})^{k,\cdot}(\Omega^{-1}_{u\cdot v})^{\cdot,k}\begin{pmatrix}1\\ 2\int B_{v_k}\end{pmatrix}. \tag{67}$$

For the correction term we have by construction that

$$\ddot{G}_{X_{m_i}}\ddot{A}^{\text{GW}}_{m_i} \Rightarrow \sum_{k\in I_{m_i}} (\Delta^+_{vu})^{k,\cdot}(\Omega^{-1}_{u\cdot v})^{\cdot,k}\begin{pmatrix}1\\ 2\int B_{v_k}\end{pmatrix}, \tag{68}$$

which leads again by inserting all the components in the definition of the estimator to the stated asymptotic distribution. □

# Appendix B: Additional Empirical Results

| | Intercept | | | Intercept and Linear Trend | | |
|---|---|---|---|---|---|---|
| | PP | $PP(fb)_1$ | $PP(fb)_2$ | PP | $PP(fb)_1$ | $PP(fb)_2$ |
| Australia | 0.963 | 0.969 | 0.395 | -1.232 | -1.256 | -1.340 |
| Austria | 0.113 | 0.247 | -0.022 | -1.836 | -1.707 | -1.801 |
| Belgium | 1.106 | 0.931 | 0.430 | -1.261 | -1.389 | -1.443 |
| Canada | -0.218 | -0.320 | -0.438 | -2.492 | *-3.048* | *-3.047* |
| Denmark | 0.162 | 0.184 | -0.195 | -2.357 | -2.376 | -2.340 |
| Finland | 1.003 | 0.917 | 0.342 | -2.299 | -2.324 | -2.393 |
| France | -0.003 | -0.121 | -0.259 | -1.939 | -2.177 | -2.189 |
| Germany | -0.253 | -0.338 | -0.322 | -2.323 | -2.548 | -2.566 |
| Italy | 0.829 | 0.505 | 0.198 | -1.683 | -1.802 | -1.823 |
| Japan | 0.302 | 0.181 | -0.080 | -1.708 | -1.858 | -1.873 |
| Netherlands | 0.285 | 0.225 | 0.099 | -2.119 | -2.227 | -2.270 |
| New Zealand | -0.174 | -0.176 | -0.204 | -2.608 | -2.689 | -2.690 |
| Norway | 1.336 | 1.316 | 0.318 | -2.116 | -2.144 | -2.134 |
| Portugal | 1.802 | 1.599 | 0.641 | -1.759 | -1.772 | -1.761 |
| Spain | 1.114 | 0.837 | 0.303 | -0.836 | -1.000 | -1.067 |
| Sweden | 0.480 | 0.473 | -0.128 | -2.545 | -2.588 | -2.599 |
| Switzerland | -1.001 | -1.047 | -1.001 | -2.774 | -2.420 | -2.450 |
| United Kingdom | 1.418 | 1.843 | 0.588 | -1.318 | -1.102 | -1.457 |
| United States | -0.310 | -0.275 | -0.428 | -2.971 | *-2.880* | *-2.886* |

Table 4: Unit root test results for log GDP per capita. The tests employed are the Phillips-Perron (1988) test, PP, as well as the one- and two-step detrended fixed-$b$ versions, $PP(fb)_1$ and $PP(fb)_2$, of this test developed in Vogelsang and Wagner (2013). The specifications of the deterministic components are intercept only and intercept and linear trend. The results are based on the Bartlett kernel with bandwidth chosen according to Newey and West (1994). *Italic* entries denote rejection of the null hypothesis at the 10% level and **bold** entries indicate rejection at the 5% level.

| | $\text{PO}_t$ | Shin | $P_{\hat{u}}$ | CT |
|---|---|---|---|---|
| Australia | -2.498 | **0.130** | 11.581 | **0.107** |
| Austria | -3.702 | 0.077 | **54.886** | 0.056 |
| Belgium | **-5.464** | 0.059 | *49.352* | 0.066 |
| Canada | -2.821 | **0.189** | 15.918 | **0.148** |
| Denmark | **-5.077** | 0.054 | *45.320* | 0.055 |
| Finland | **-5.585** | 0.049 | **74.872** | 0.053 |
| France | **-4.984** | 0.061 | 30.775 | 0.067 |
| Germany | **-7.714** | **0.413** | **68.790** | **0.116** |
| Italy | **-4.524** | **0.186** | 35.780 | **0.150** |
| Japan | **-6.304** | **0.163** | 9.198 | **0.156** |
| Netherlands | **-5.668** | **0.105** | **94.376** | 0.072 |
| New Zealand | **-5.314** | **0.138** | 12.481 | **0.121** |
| Norway | -3.684 | *0.094* | 19.327 | *0.092* |
| Portugal | **-11.272** | *0.092* | **75.219** | *0.098* |
| Spain | -3.595 | **0.101** | 41.730 | *0.090* |
| Sweden | **-4.469** | *0.083* | 30.697 | 0.085 |
| Switzerland | **-6.147** | 0.053 | **86.006** | 0.073 |
| United Kingdom | **-7.826** | *0.088* | **100.674** | 0.076 |
| United States | -2.585 | **0.605** | 12.747 | **0.157** |

Table 5: Cointegration and non-cointegration test results for (1). The left block-column presents the results for the "linear" non-cointegration test $PO_t$ of Phillips and Ouliaris (1990) and the "linear" cointegration test of Shin (1994). Linear here refers to an application of these tests treating log GDP per capita and its square as two integrated processes. The right block-column presents the results for the modifications of these two tests to the CPR setting discussed in Wagner (2013, 2015). These are labelled $P_{\hat{u}}$ (non-cointegration test) and $CT$ (cointegration test). *Italic* entries denote rejection of the null hypothesis at the 10% level and **bold** entries indicate rejection at the 5% level.

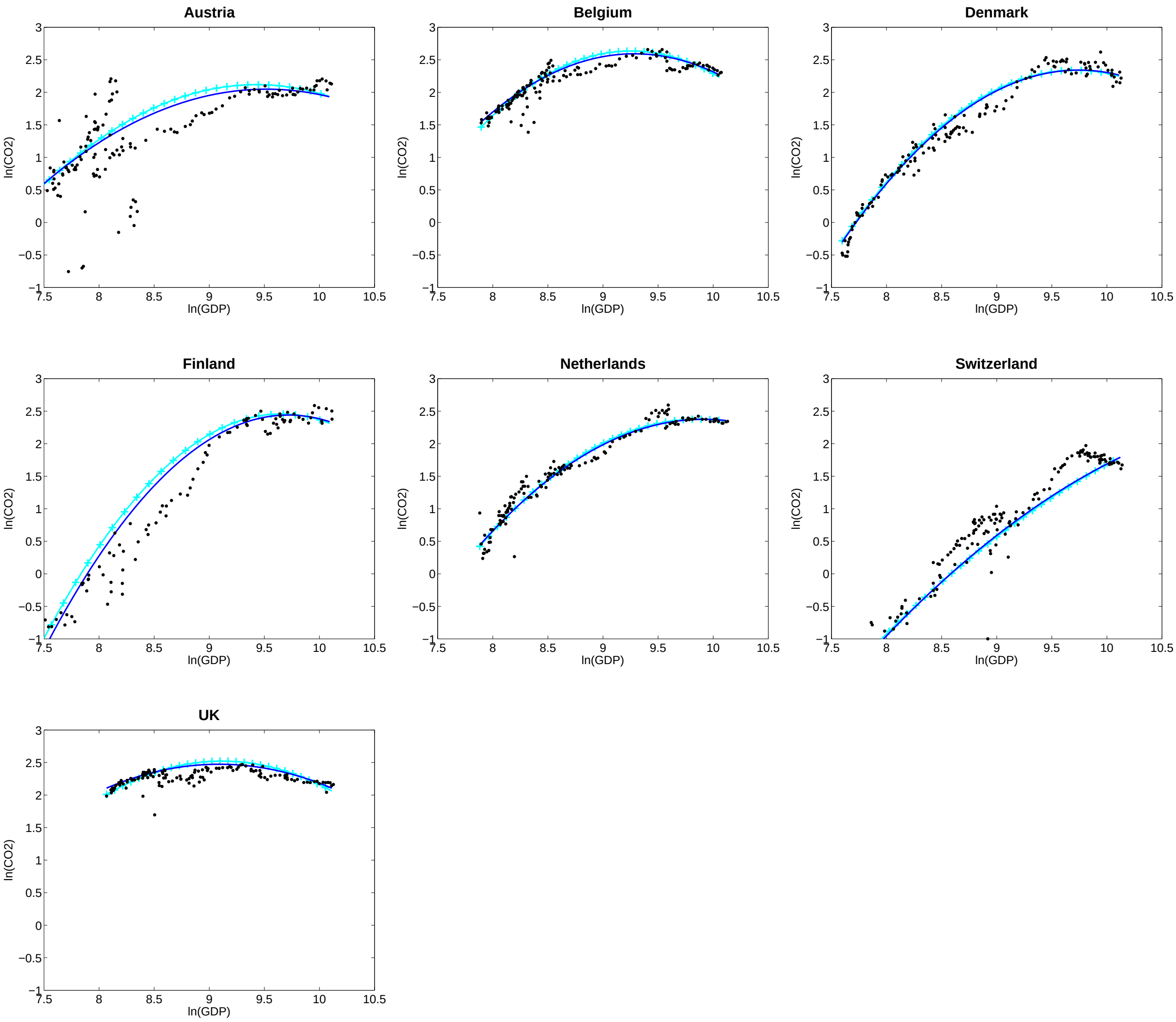


Figure 7: EKC estimation results for Equation (1): scatter plot and EKC. The solid lines correspond to the FM-SUR estimates and the solid lines with +-marks to the FM-SOLS estimates. For further explanations see notes to Figure 1.

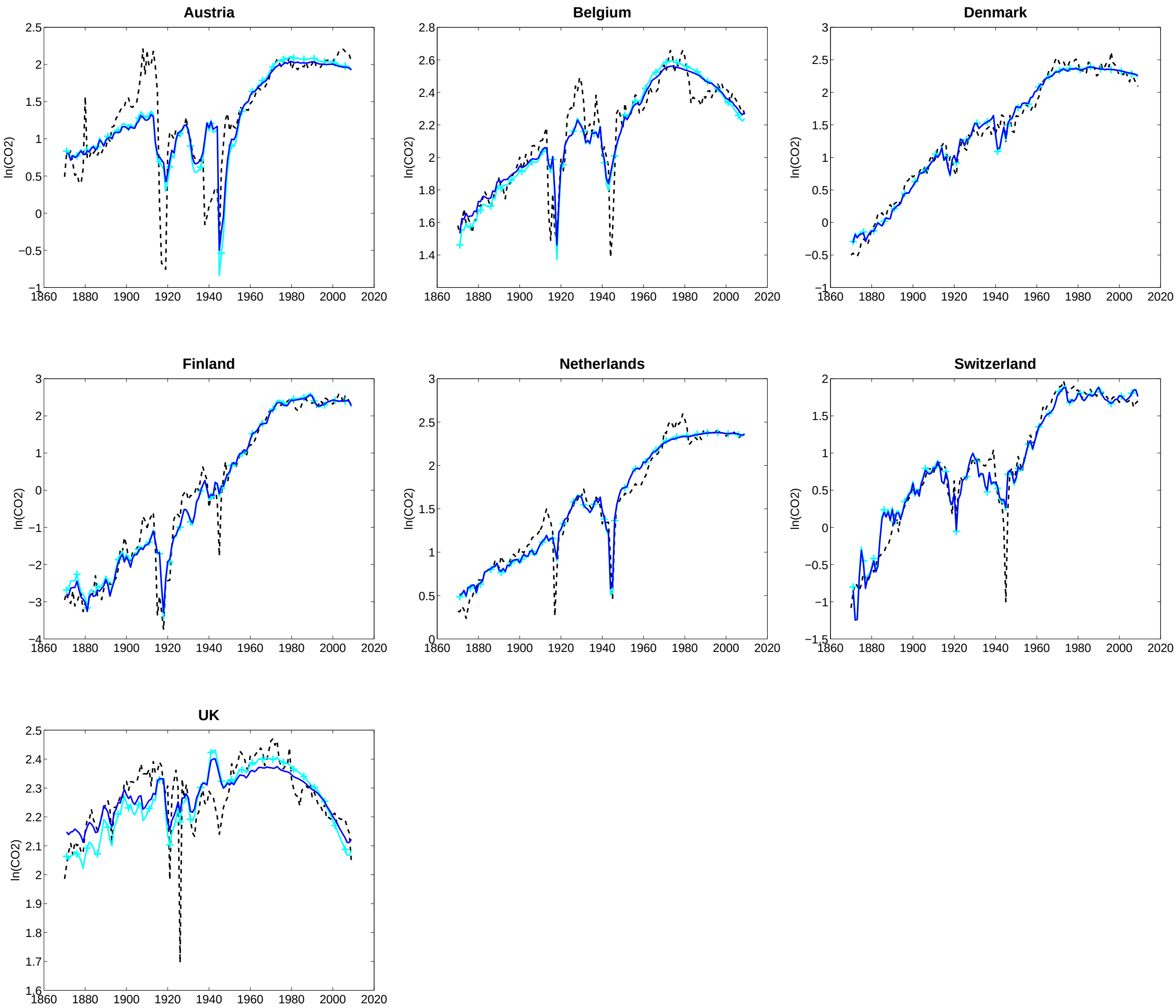


Figure 8: EKC estimation results for Equation (1): actual and fitted values. The dashed lines show the actual values, the solid lines the FM-SUR fitted values and the solid lines with +-marks the FM-SOLS fitted values.

| | | FM-SOLS |
|---|---|---|
| Linear Trend & Stochastic Regressors (P) | 2 | DK-FI |
| Stochastic Regressors (S) | 2<br>3 | AT-DK, AT-NL, BE-UK, DK-FI, NL-UK,<br>AT-BE-NL, BE-NL-UK |
| Linear Trend (T) | 2<br><br>3<br><br>4 | AT-DK, AT-FI, AT-CH, AT-UK, BE-DK, BE-NL,<br>BE-UK, DK-FI, DK-CH, DK-UK, FI-CH,<br>AT-BE-UK, AT-DK-FI, AT-DK-CH, AT-DK-UK,<br>AT-FI-CH, BE-NL-UK, BE-DK-UK, DK-FI-CH,<br>AT-DK-FI-CH, AT-BE-DK-UK |
| | | **FM-SUR** |
| Linear Trend & Stochastic Regressors (P) | 2 | DK-FI |
| Stochastic Regressors (S) | 2<br>3 | AT-DK, AT-NL, BE-NL, BE-UK, DK-FI, NL-UK,<br>AT-BE-NL, AT-NL-UK, BE-NL-UK, DK-NL-UK |
| Linear Trend (T) | 2<br><br>3<br><br>4 | AT-DK, AT-FI, AT-CH, AT-UK, BE-DK, BE-NL,<br>BE-UK, DK-FI, DK-CH, DK-UK, FI-CH, FI-UK,<br>AT-BE-UK, AT-DK-FI, AT-DK-CH, AT-DK-UK, AT-FI-CH,<br>AT-FI-UK, BE-DK-UK, DK-FI-UK, DK-FI-CH,<br>AT-DK-FI-UK, AT-DK-FI-CH, AT-BE-DK-UK |

Table 6: List of group members corresponding to the tests described in Table 2. For more details see notes to Table 2.

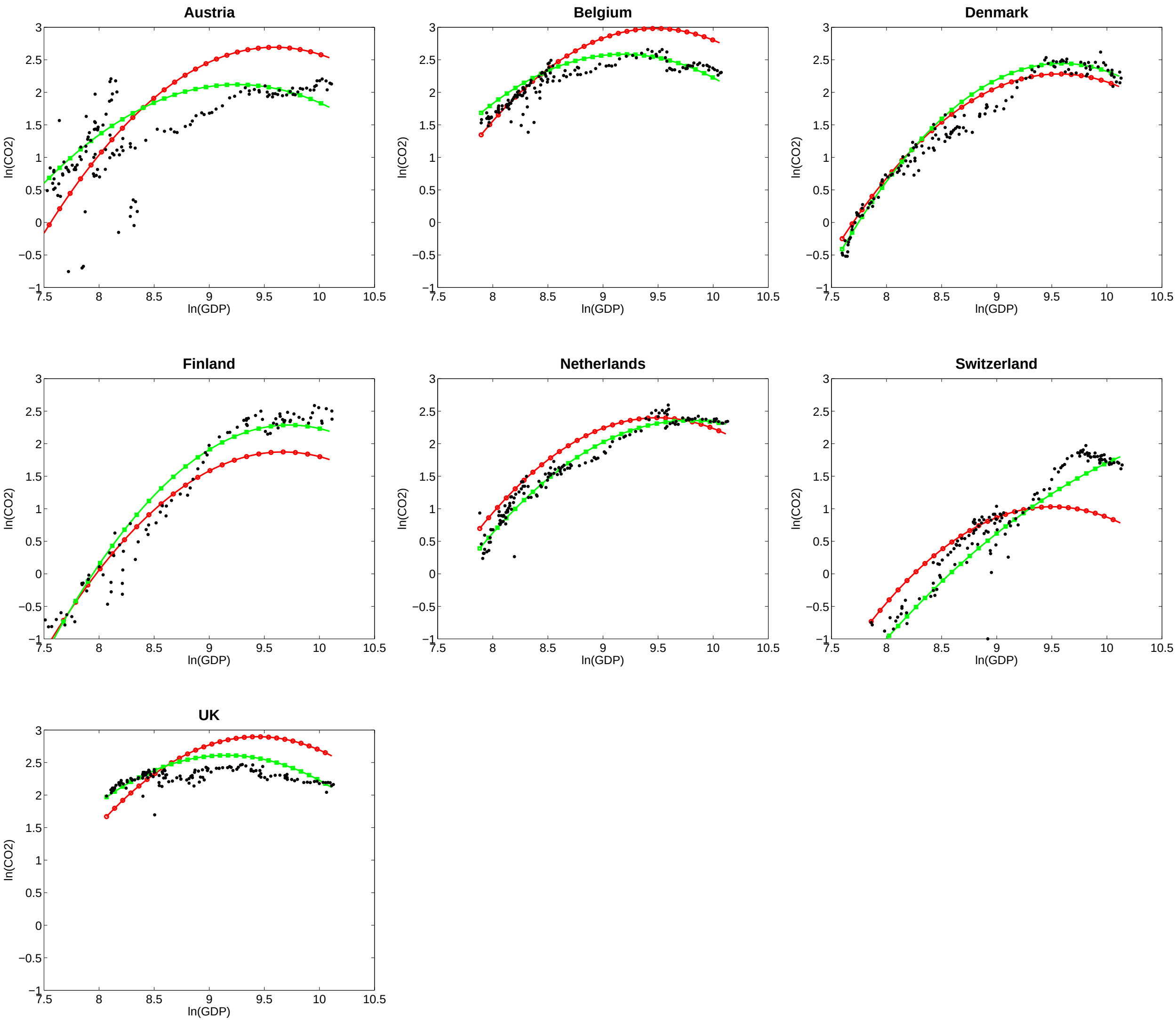


Figure 9: EKC estimation results for Equation (1): scatter plot and EKC. The solid lines with square symbols correspond to the group-wise pooled FM-SUR estimates and the solid lines with o-marks to the pooled FM-SUR estimates. For further explanations see notes to Figure 1.

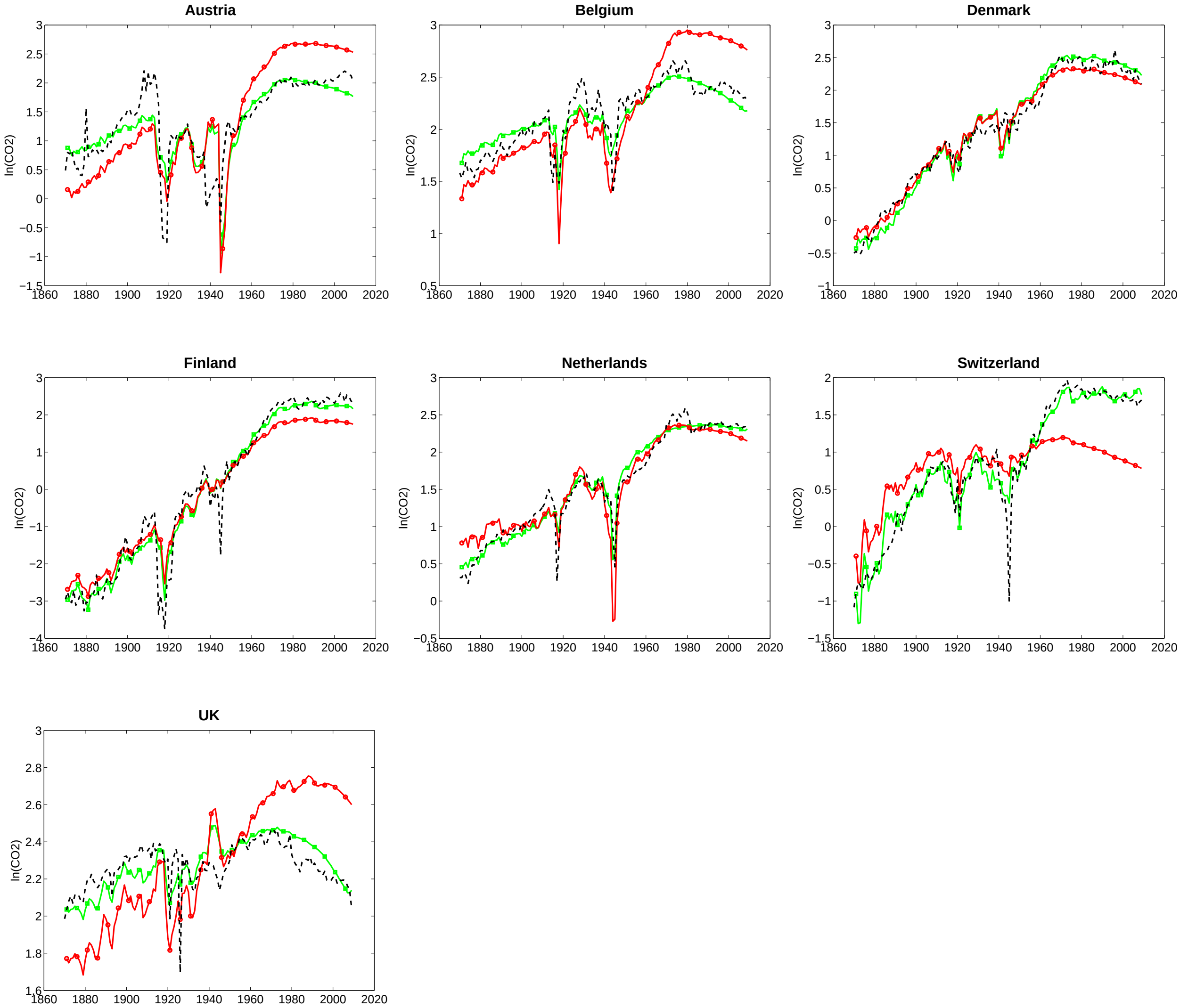


Figure 10: EKC estimation results for Equation (1): actual and fitted values. The dashed lines show the actual values, the solid lines with square symbols the group wise pooled FM-SUR fitted values and the solid lines with o-marks the pooled FM-SUR fitted values.