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 two fully modified-type estimators, as well as Wald-type hypothesis tests based upon them, for systems of seemingly unrelated cointegrating polynomial regressions. 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 six early industrialized countries. Group-wise pooled estimation leads to almost the same fit as unrestricted country-by-country or seemingly unrelated 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, Group-Wise Pooling, 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, is 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 for EKC analysis include Esteve and Tamarit

[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 have been 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), Fosten *et al.* (2012), Friedl and Getzner (2003), He and Richard (2010), Jalil and Mahmud (2009) and Lindmark (2002). Panel data EKC 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 using standard FM-OLS treating unit root processes and their powers all as unit root processes is discussed in Stypka *et al.* (2016).[5]

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)

[3]Prior to the estimation of these relationships, testing for nonlinear cointegration in EKC-type relationships need to be performed, 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]In the example of a quadratic EKC this means that log GDP per capita and its square are treated as two integrated regressors and standard FM-OLS is performed in the two regressor case. The above-listed papers employing cointegration methods all use cointegration techniques this way, as also discussed in Wagner (2015).

or Pedroni (2000), is clearly often unrealistic.[6] Also, the tacit assumption of these studies that 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.[7] 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, if the corresponding restrictions are not rejected, 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 square 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 *group-wise* pooled settings of a form relevant for our application in detail in Section 2.2.

The application of our methodology to study the EKC for CO_2 emissions for six early industrialized countries over the period 1870–2013 highlights the usefulness of the SUCPR approach. Group-wise pooled estimation of the EKC leads 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 a situation- or problem-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 for data sets with a small

[6] 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.

[7] In terms of econometric methodology Hong and Wagner (2014) discuss 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).

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 regressions including a discussion of group-wise pooling – both with respect to testing for poolability as well estimation imposing the corresponding pooling restrictions – 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 is divided in two subsections. The first contains some additional material and results concerning the three variants (P), (S) and (T) of pooled estimation and the second provides 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 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 and integration bounds for notational simplicity.

2 Seemingly Unrelated Cointegrating Polynomial Regressions

For the discussion in this paper it suffices to consider the special case of a system of seemingly unrelated quadratic polynomial regressions, where 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 country i in year t:

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

with $Z_{i,t} := [D_{i,t}', X_{i,t}']'$, where $D_{i,t} := [1, t]'$ and $X_{i,t} := [x_{i,t}, x_{i,t}^2]'$, 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

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

with $\sum_{j=0}^{\infty} j\|\Phi_j\| < \infty$ and $\sum_{j=0}^{\infty} j\|\Psi_j\| < \infty$. Furthermore, we assume $|\Phi(1)| \neq 0$, which excludes cointegration in the I(1) vector process $\{x_t\}$, 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 $\{\xi_t^0\}_{t\in\mathbb{Z}} := \{[\zeta_t', \epsilon_t']'\}_{t\in\mathbb{Z}}$ 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 a functional central limit theorem to hold, i.e.

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor rT\rfloor}\xi_t = \frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor rT\rfloor}\begin{bmatrix}u_t\\ v_t\end{bmatrix} \Rightarrow B(r) = \begin{bmatrix}B_u(r)\\ B_v(r)\end{bmatrix} := \Omega^{1/2}W(r), \quad 0 \leq r \leq 1, \tag{3}$$

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 ξ_t:

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

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

$$y_t = Z_t'\theta + u_t, \quad t = 1, \ldots, T \tag{5}$$

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 θ in (6) is consistent with a limiting distribution contaminated by second order bias terms, just as in the linear seemingly unrelated cointegration 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 "classical" SUR estimation, i.e., in GLS estimation, is an estimate of the contemporaneous error variance matrix. In the cointegration setting we allow for both error serial correlation and endogenous regressors. To take these two generalizations 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 the OLS and MSUR estimators is derived in Hong and Wagner (2014, Proposition 1) for the SUCPR case. The nuisance parameter dependent limiting distributions of these two estimators provide guidance for the construction of appropriate two-part FM-type corrections.[8] One of the corrections is as in the linear case, i.e., the dependent variable 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.[9] 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

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

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 some additional quantities. We define, again for our special case, the weighting matrix $G = G(T) := I_N \otimes G_\bullet(T)$, with $G_\bullet(T) := \mathrm{diag}(T^{-1/2}, T^{-3/2}, T^{-1}, T^{-3/2})$ and a stochastic process $J(r) := \mathrm{diag}\,(J_1(r), \ldots, J_N(r))$ with $J_i(r) := [1, r, B_{v_i}(r), B_{v_i}^2(r)]'$, where $B_{v_i}(r)$ denotes 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:*

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

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

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

[9]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. Throughout the paper we assume these conditions on bandwidth and kernel to be in place.

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

$$G^{-1}\left(\hat{\theta} - \theta\right) \Rightarrow \left(\int JJ'\right)^{-1} \int J dB_{u.v}, \tag{10}$$

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

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 $B_v(r)$ and consequently 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 $\hat{\theta}$ and $\tilde{\theta}$ contain elements that converge at different rates, obtaining formal results for the Wald-type test statistics requires a condition on the restriction matrix (in case of linear hypotheses) 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 with all estimated coefficients converging at the same rate arise.

Proposition 2 (Hong and Wagner 2014, Proposition 3) *Let* y_t *be generated by* (1) *with the given assumptions in place. Consider* s *linearly independent restrictions collected in* $H_0 : R\theta = r$ *with* $R \in \mathbb{R}^{s \times 4N}$ *of 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 row rank* s.

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

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

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

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 the 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 yet 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), based on the seminal work of Phillips and Moon (1999), provide theory for the cross-sectionally independent case for the cubic formulation with one- and two-way fixed effects.[10]

If the considered null hypothesis is not rejected, then pooled estimation, as described for these three cases in Appendix A.1, of a smaller number of parameters allows to lift some efficiency gains in estimation. For our data, the above-given three "global" hypotheses (P), (S) and (T) are rejected.[11] A more detailed analysis, see Section 3, of the FM-SUR results reveals that the coefficient corresponding to the linear time trend can be pooled in three subgroups (of sizes three, two and one). For the coefficients to GDP and its square, the stochastic regressors, group-wise pooling analysis identifies one group of size three for which pooling is not rejected.

Exploiting the possibilities of group-wise pooling just indicated necessitates formulating the corresponding Wald-type statistics as well as the corresponding 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 testing the null hypothesis that the coefficients for the linear time trend are group-wise pooled over a partition of k subsets $I_{n_j},\ j = 1, \ldots, k$ with $I := \{1, \ldots, N\} =$

[10] 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.

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

$\bigcup_{j=1}^{k} I_{n_j}$. Similarly, we consider a partition over l subsets $I_{m_j},\ j=1,\ldots,l$ for the regressors $[x_{i,t}, x_{i,t}^2]'$, i.e., $I=\bigcup_{j=1}^{l} I_{m_j}$. Without loss of generality we 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 with $|S|$ here 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:

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

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 $M_j = |I_{m_j}|,\ j=1,\ldots,l$. 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^{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. Equivalently, including only the subsets of size larger than one in the restrictions matrix R^{GW} would require to define another index, n_k^* say, 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:

$$\check{D}_t := [\check{D}_{1,t}', \ldots, \check{D}_{k,t}']' \in \mathbb{R}^{k \times N}, \tag{19}$$

where

$$\check{D}_{j,t} := \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{n_j}\}} \cdot t \cdot e_{i,N}', \quad j = 1, \ldots, k. \tag{20}$$

For the stochastic regressors we similarly have

$$\check{X}_t := [\check{X}_{1,t}', \ldots, \check{X}_{l,t}']' \in \mathbb{R}^{2l \times N}, \tag{21}$$

with

$$\check{X}_{j,t} := \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{m_j}\}} \cdot \left(e_{i,N}' \otimes X_{i,t}\right), \quad j = 1, \ldots, l. \tag{22}$$

With these quantities the group-wise pooled model can be compactly written as

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

with $y_t := [y_{1,t}, \ldots, y_{N,t}]'$, $u_t := [u_{1,t}, \ldots, u_{N,t}]'$, $\check{Z}_t := [I_N, \check{D}_t', \check{X}_t']' \in \mathbb{R}^{(N+k+2l) \times N}$ and the parameter vector $\theta^{\mathrm{GW}} := [c_1, \ldots, c_N, \delta_1, \ldots, \delta_k, \beta_1', \ldots, \beta_l']' \in \mathbb{R}^{N+k+2l}$, where $\beta_j = [\beta_{1,j}, \beta_{2,j}]'$ for $j = 1, \ldots, l$. Finally, stacking the quantities over time gives

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

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

The correction terms for the group-wise pooled FM-SOLS and FM-SUR estimators are defined as $A^{\mathrm{GW}*} := [0_{1 \times (N+n_k)}, A_1^{\mathrm{GW}*\prime}, \ldots, A_l^{\mathrm{GW}*\prime}]'$, $\tilde{A}^{\mathrm{GW}*} := [0_{1 \times (N+n_k)}, \tilde{A}_1^{\mathrm{GW}*\prime}, \ldots, \tilde{A}_l^{\mathrm{GW}*\prime}]'$, with

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

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

For group-wise pooled estimation the weighting matrix is given by $\check{G} := \text{diag}(\check{G}_c, \check{G}_D, \check{G}_X) = \text{diag}(T^{-1/2} \cdot I_N, T^{-3/2} \cdot I_k, I_l \otimes \text{diag}(T^{-1}, T^{-3/2}))$. The limit stochastic process is now given by $\check{J}(r) := [I_N, \check{J}_D', \check{J}_X']'$, with $\check{J}_D(r) := [\check{J}_{D_1}(r)', \ldots, \check{J}_{D_k}(r)']'$ and $\check{J}_{D_j}(r) := \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{n_j}\}} \cdot r \cdot e_{i,N}'$ for $j = 1, \ldots, k$. The process $\check{J}_X(r) := [\check{J}_{X_1}(r)', \ldots, \check{J}_{X_l}(r)']'$ is composed of $\check{J}_{X_j}(r) := \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{m_j}\}} \cdot \left(e_{i,N}' \otimes \begin{pmatrix} B_{v_i}(r) \\ B_{v_i}^2(r) \end{pmatrix} \right)$ for $j = 1, \ldots, l$.

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 again 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:*

$$\hat{\theta}^{GW} := \left(\check{Z}'\check{Z}\right)^{-1} \left(\check{Z}'y^{+} - A^{GW*}\right), \tag{27}$$

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

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

$$\check{G}^{-1}\left(\hat{\theta}^{GW} - \theta^{GW}\right) \Rightarrow \left(\int \check{J}\check{J}'\right)^{-1} \int \check{J} dB_{u \cdot v}, \tag{29}$$

$$\check{G}^{-1}\left(\tilde{\theta}^{GW} - \theta^{GW}\right) \Rightarrow \left(\int \check{J}\Omega_{u \cdot v}^{-1}\check{J}'\right)^{-1} \int \check{J}\Omega_{u \cdot v}^{-1} dB_{u \cdot v}. \tag{30}$$

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 of the EKC for CO_2 emissions 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 1870–2013 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.[12] The GDP data, measured in 1990 Geary-Khamis Dollars, are from the Maddison project at the University of Groningen and from The Conference Board

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

Total Economy Database.[13] 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).[14] 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 six countries: Austria (AT), Belgium (BE), Finland (FI), the Netherlands (NL), Switzerland (CH) and the UK.[15]

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-SOLS and FM-SUR, for the six countries listed above. In addition, the lower left block of the table contains the results when estimating the EKC "fully" pooled, allowing only for country specific intercepts (the form of pooling referred to as (P) in Section 2.1).[16] The following messages 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 more than twice as large as the FM-SOLS and FM-SUR turning points. For Switzerland, the turning point is estimated far outside the sample range, with values ranging from 1.3 to 3.1 millions, by all three estimators. This finding is related to the fact that, see Figure 2, per capita CO_2 emissions are essentially constant in Switzerland since about 1980. 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 on the FM-SUR estimator in the discussion from now on.[17] Third, the estimated coefficients and consequently the estimated turning points differ substantially across countries and this heterogeneity can – by

[13]See Bolt and van Zanden (2014), `http://www.ggdc.net/maddison/maddison-project/home.htm` and `http://www.conference-board.org/data/economydatabase`.

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

[15]This is slightly different from Wagner (2015) who finds evidence for a quadratic EKC for CO_2 emissions for only four out of the six 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.

[16]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 $i = 1, \ldots, 6$. Note also that we obtain very similar results for the cubic specification, both with respect to cointegration testing and estimation results. The coefficient to the third power of GDP is not significant throughout and it thus suffices to consider the quadratic specification.

[17]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.

	$\hat{\delta}$	$\hat{\beta}_1$	$\hat{\beta}_2$	TP	$\hat{\delta}$	$\hat{\beta}_1$	$\hat{\beta}_2$	TP
	Austria				**Belgium**			
FM-CPR	-0.017	6.247	-0.277	78,059	-0.004	11.358	-0.599	13,142
(t-values)	(-3.713)	(2.510)	(-2.019)		(-2.727)	(10.121)	(-10.159)	
FM-SOLS	-0.018	10.033	-0.486	30,515	-0.005	12.313	-0.649	13,230
(t-values)	(-4.750)	(4.634)	(-4.073)		(-3.629)	(13.325)	(-13.384)	
FM-SUR	-0.013	8.278	-0.403	28,699	-0.004	10.687	-0.562	13,556
(t-values)	(-4.095)	(4.891)	(-4.182)		(-3.935)	(14.073)	(-13.856)	
	Finland				**Netherlands**			
FM-CPR	-0.029	15.610	-0.737	39,523	0.001	9.437	-0.481	18,280
(t-values)	(-3.260)	(9.356)	(-8.796)		(0.614)	(8.438)	(-8.076)	
FM-SOLS	-0.039	16.162	-0.746	50,845	0.001	9.823	-0.502	17,783
(t-values)	(-4.974)	(10.600)	(-9.721)		(0.585)	(9.334)	(-8.970)	
FM-SUR	-0.029	15.892	-0.752	38,892	0.002	10.185	-0.524	16,524
(t-values)	(-5.863)	(14.140)	(-12.276)		(1.053)	(11.511)	(-10.878)	
	Switzerland				**UK**			
FM-CPR	-0.024	7.755	-0.273	1.5×10^6	-0.008	8.657	-0.446	16,287
(t-values)	(-6.421)	(6.312)	(-4.031)		(-3.406)	(6.532)	(-6.794)	
FM-SOLS	-0.024	6.933	-0.232	3.1×10^6	-0.007	9.887	-0.516	14,496
(t-values)	(-7.981)	(7.399)	(-4.463)		(-3.448)	(8.539)	(-9.001)	
FM-SUR	-0.022	7.441	-0.265	1.3×10^6	-0.007	8.402	-0.437	15,068
(t-values)	(-7.743)	(7.665)	(-4.941)		(-4.035)	(8.667)	(-9.001)	
	Pooled							
FM-SOLS	-0.015	13.572	-0.667	26,053				
(t-values)	(-8.344)	(20.474)	(-18.326)					
FM-SUR	-0.013	13.594	-0.677	23,002				
(t-values)	(-15.293)	(35.246)	(-32.226)					

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 $\exp\left(-\frac{\hat{\beta}_1}{2\hat{\beta}_2}\right)$.

construction – not be captured by the pooled, i.e., almost panel-type, estimation results in the lower left block. This finding highlights that commonly used panel methods need to be considered very carefully, or maybe not used at all for situations as considered here.[18]

The results from Table 1 are displayed graphically in Figures 1 and 2. The first figure displays the estimated EKCs, given by using 144 equidistant values for the explanatory variable from the range of log GDP per capita associated with values of the time trend ranging from 1,. . . ,144 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

[18] As already mentioned, 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.

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.

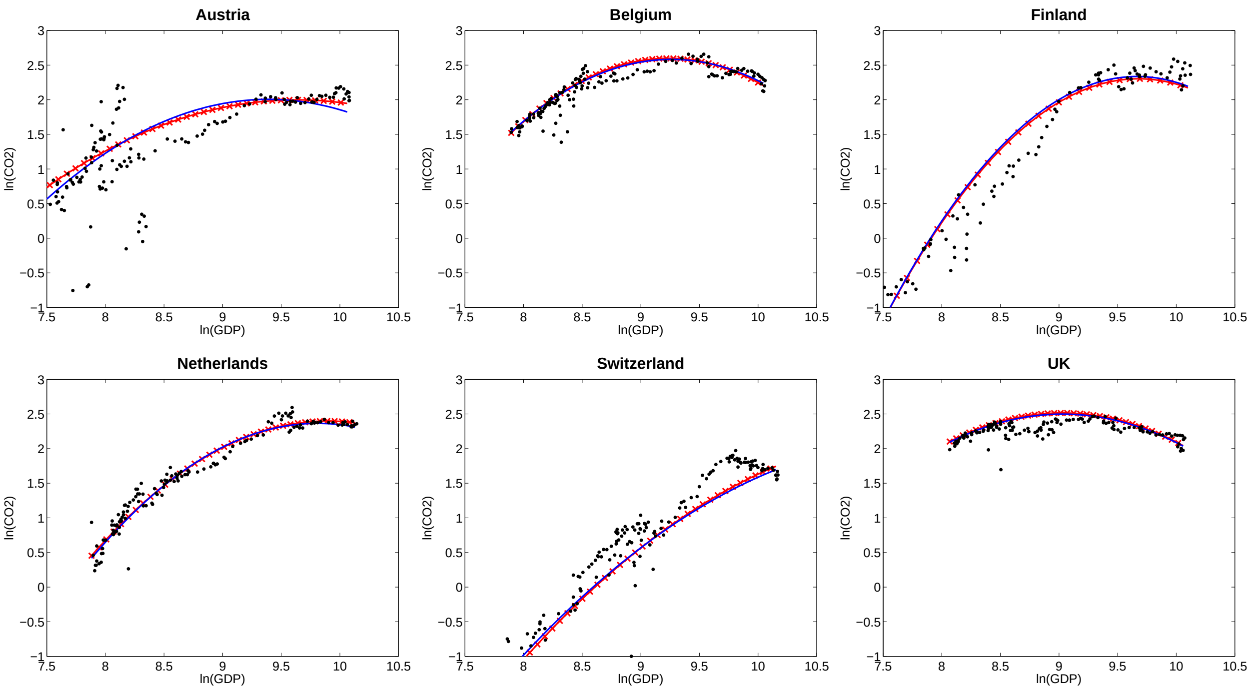

Figure 1: EKC estimation results for Equation (1): scatter plot and EKC. The dots show the pairs of observations of ln(GDP) per capita and ln(CO_2) emissions per capita. The lines show results based on inserting 144 equidistant points from the sample range of ln(GDP) per capita, with corresponding values of the linear trend given by $t = 1, \ldots, 144$ 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.

Performing the poolability tests (P), (S) and (T) described in Section 2.1 and in more detail in Appendix A.1 for the six considered 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-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

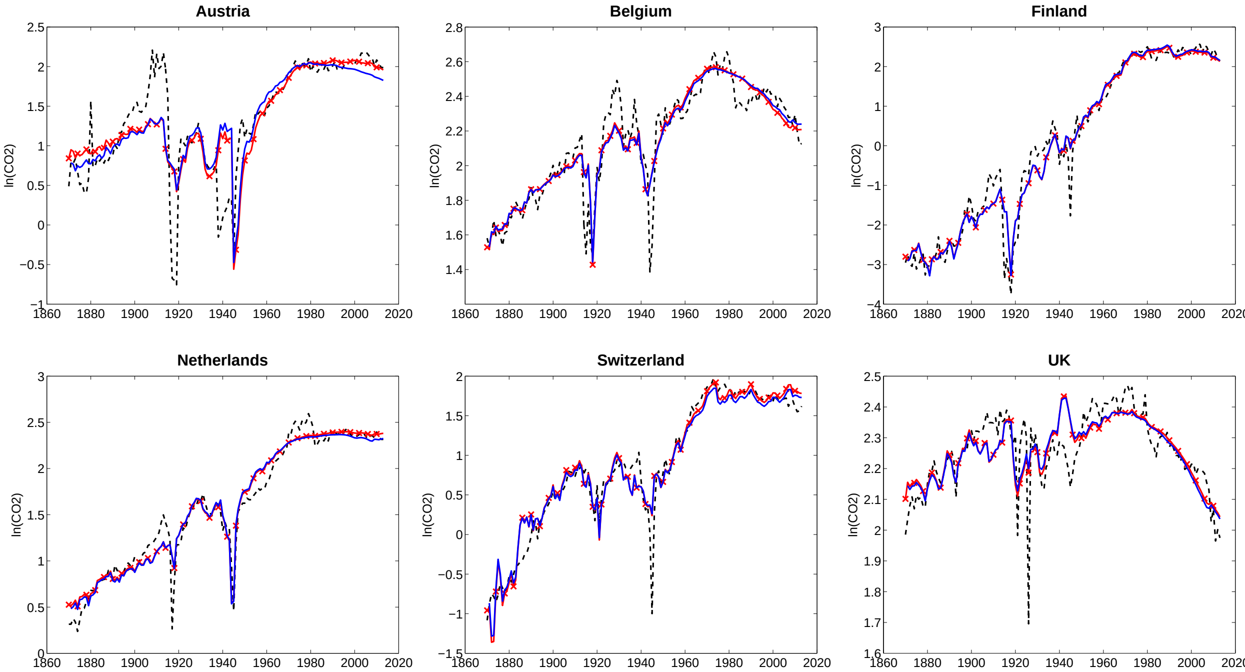

Figure 2: EKC estimation results for Equation (1): actual and fitted values. The dashed lines show the actual values of $\ln(CO_2)$ per capita emissions, the solid lines with x-marks the FM-CPR fitted values and the solid lines the FM-SUR fitted values.

in structure to Figures 1 and 2. For all six countries the differences are quite huge, 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 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. We start by testing 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 15 possible country-pairs, 20 country-triples and so on.[19] The results are given in Table 2 and

[19] 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 six 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. Note that there is a recent literature to identify (coefficient) structure in panel data, see Ke *et al.* (2016) or Su *et al.* (2016). However, our problem does not fit that literature either, since we have small (to medium) N and cointegration in the SUCPR setting, whereas this literature is to date concerned with standard stationary settings.

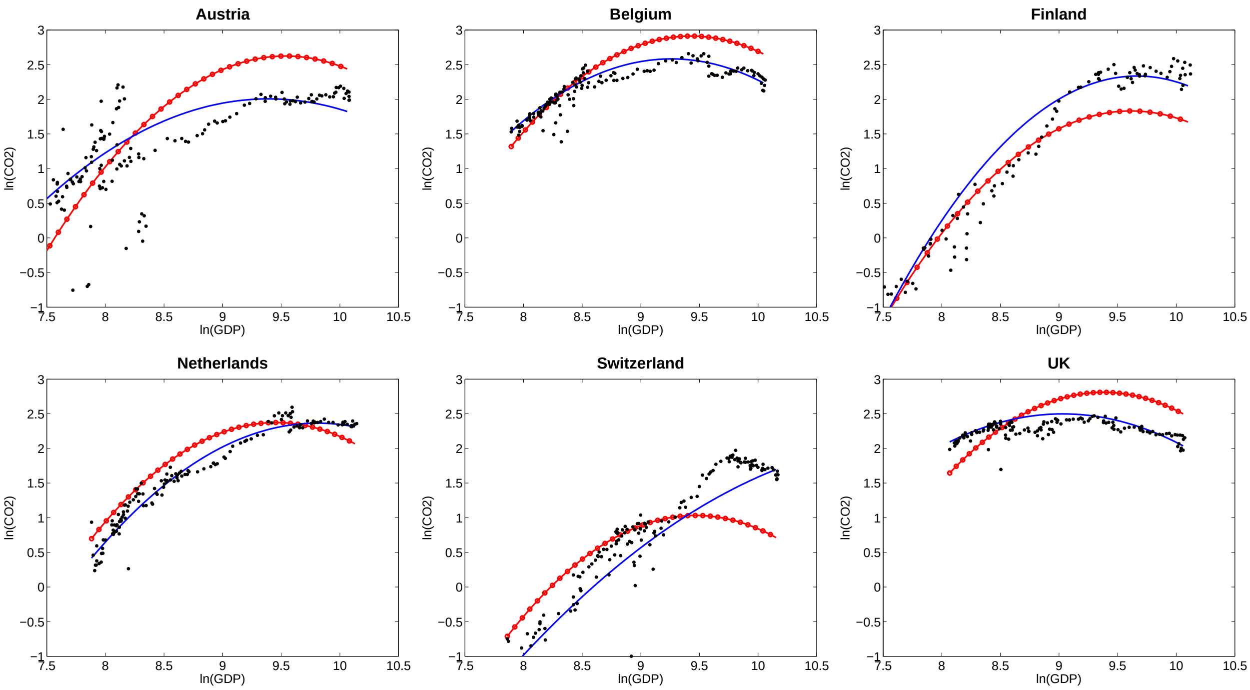

Figure 3: EKC estimation results for Equation (1): scatter plot and EKC. The solid lines correspond to the FM-SUR estimates and the solid lines with o-marks to the pooled FM-SUR estimates. For further explanations see notes to Figure 1.

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 focus again on the results obtained with FM-SUR. The full pooling hypothesis (P) is rejected throughout, even for all pairs. With FM-SUR, the slope parameters β_1 and β_2 can be pooled for (i.e., the pooling hypothesis (S) is not rejected for) four country-pairs, two country-triples and one group of size four (containing AT, BE, NL and UK). With respect to the trend parameters there are three country groups of size three, for which the trend slope can be pooled. Austria, Finland and the UK are each present in two of the three groups.

We take the above results as starting point to estimate the EKC for the six countries in a group-wise pooled fashion. In particular we consider: the trend slope pooled in three groups, comprising Austria, Finland and Switzerland; Belgium and the UK; and the Netherlands (as group of size one) respectively. The slope parameters are pooled in four groups, given by Belgium, the Netherlands and

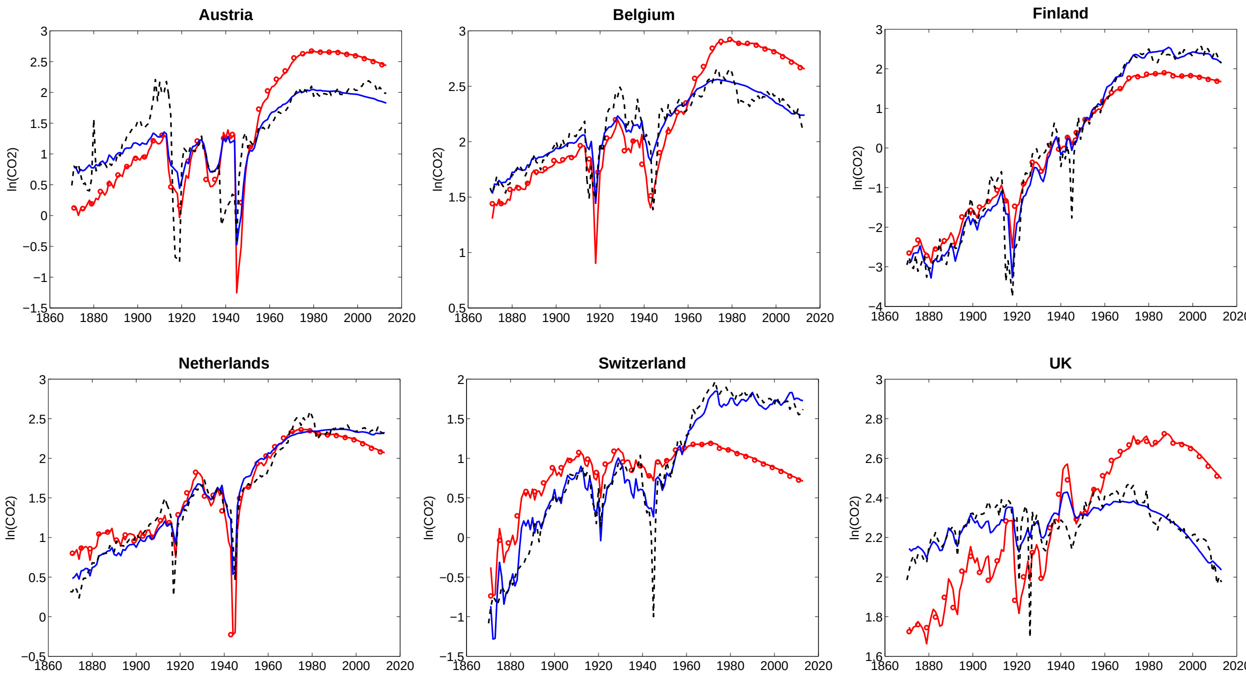

Figure 4: EKC estimation results for Equation (1): actual and fitted values. The dashed lines show the actual values, the solid lines the FM-SUR fitted values and the solid lines with o-marks the pooled FM-SUR fitted values.

the UK; and the three single member groups Austria; Finland; and Switzerland.[20] Table 3 displays the estimation results. As observed up to now, the estimates are also very similar for the now group-wise pooled FM-SOLS and FM-SUR estimates. Looking at the coefficients in the individual groups clearly shows 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 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 group-wise poolability restrictions supported by hypothesis testing in group-wise pooled

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

Group size k	2	3	4	5	6	2	3	4	5	6
Total nr. of groups of size k	15	20	15	6	1	15	20	15	6	1
	FM-SOLS					**FM-SUR**				
Linear Trend & Stochastic Regressors (P)										
Stochastic Regressors (S)	3	2				4	2	1		
Linear Trend (T)	6	2				7	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 for which the respective null hypotheses are not rejected 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.

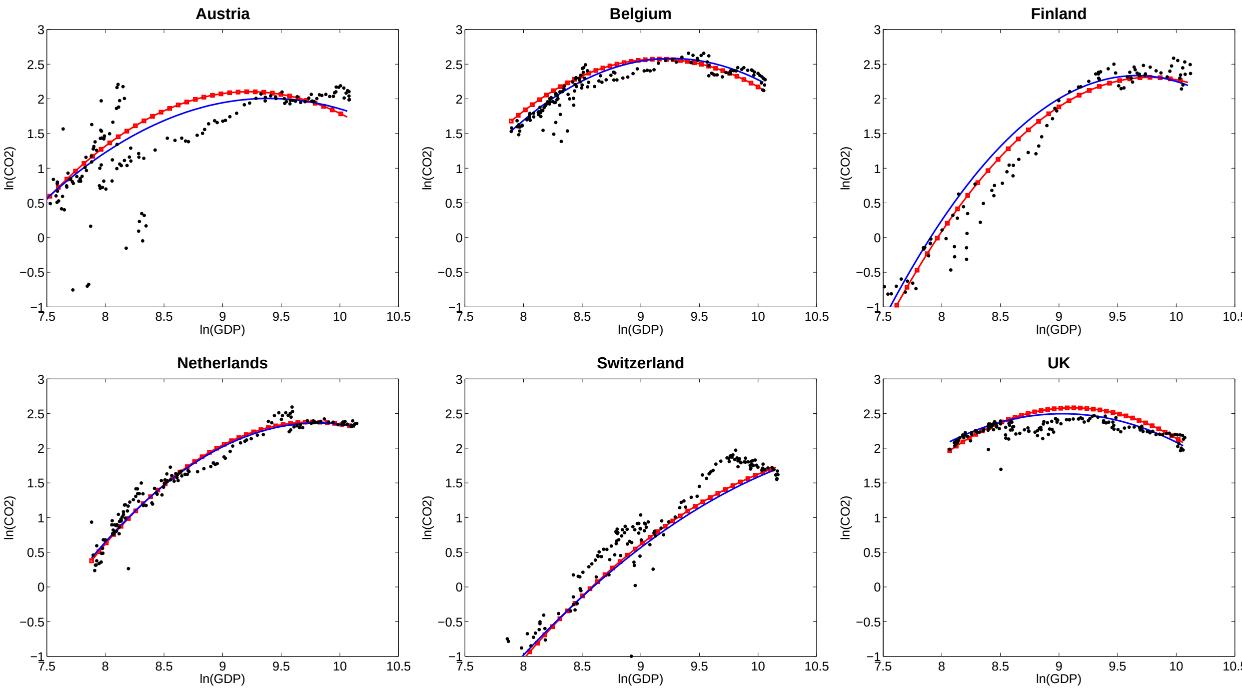

Figure 5: EKC estimation results for Equation (1): scatter plot and EKC. The solid lines correspond to the FM-SUR estimates and the solid lines with the square symbols to the group-wise pooled FM-SUR estimates. For further explanations see notes to Figure 1.

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

	$\hat{\delta}_{n_1}$	$\hat{\delta}_{n_2}$	$\hat{\delta}_{n_3}$					
Countries	**AT-FI-CH**	**BE-UK**	**NL**					
FM-SOLS	-0.022	-0.009	0.001					
(t-values)	(-6.825)	(-7.827)	(0.883)					
FM-SUR	-0.019	-0.009	0.002					
(t-values)	(-9.443)	(-11.122)	(2.017)					
	$\hat{\beta}_{1,m_1}$	$\hat{\beta}_{2,m_1}$	$\hat{\beta}_{1,m_2}$	$\hat{\beta}_{2,m_2}$	$\hat{\beta}_{1,m_3}$	$\hat{\beta}_{2,m_3}$	$\hat{\beta}_{1,m_4}$	$\hat{\beta}_{2,m_4}$
Countries	**BE-NL-UK**		**AT**		**FI**		**CH**	
FM-SOLS	11.580	-0.600	11.054	-0.534	13.907	-0.654	6.991	-0.242
(t-values)	(16.445)	(-16.355)	(5.372)	(-4.691)	(12.850)	(-10.435)	(6.980)	(-4.310)
TP	15,514		31,304		41,480		1.9×10^6	
FM-SUR	10.852	-0.562	10.656	-0.521	14.649	-0.704	8.261	-0.319
(t-values)	(21.370)	(-20.896)	(6.370)	(-5.524)	(15.528)	(-12.645)	(8.781)	(-6.109)
TP	15,677		27,646		32,942		4.2×10^5	

Table 3: Group-wise pooled estimation results for Equation (1) using FM-SOLS and FM-SUR. The trend parameter δ is pooled in three groups (of sizes three, two and one) and the slope parameters β_1, β_2 are pooled in four groups (of sizes three and thrice one). The estimated turning points TP are computed as $\exp\left(-\frac{\hat{\beta}_1}{2\hat{\beta}_2}\right)$.

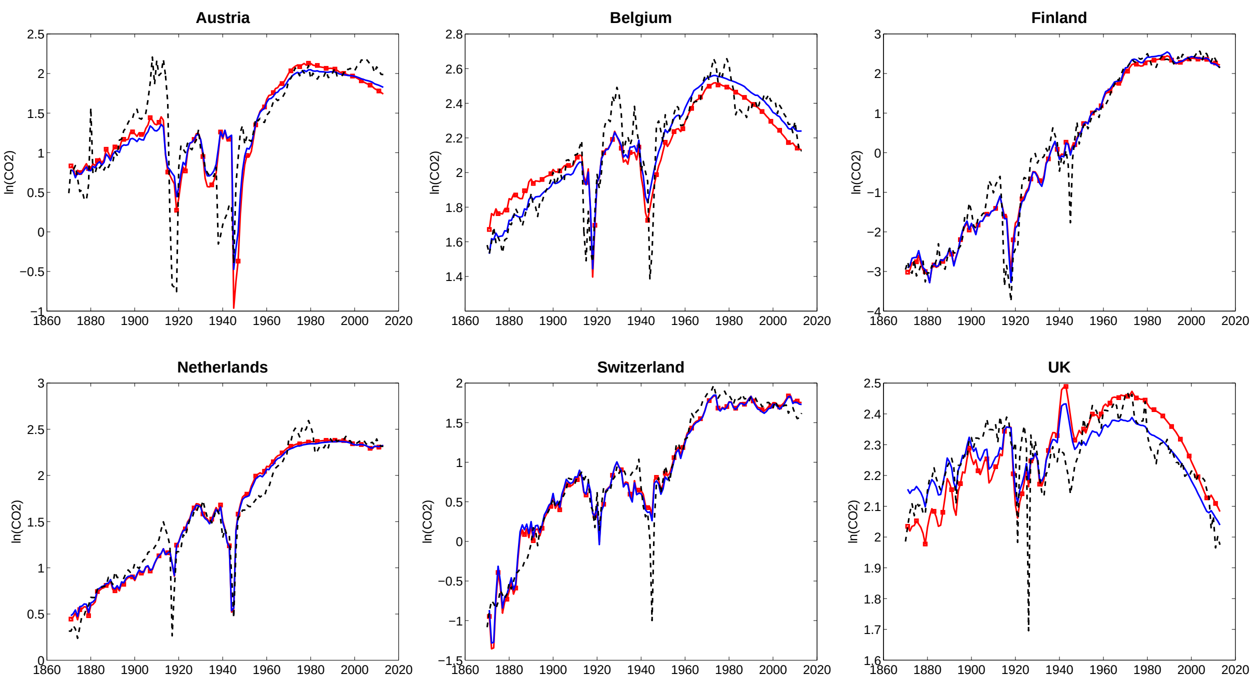

Figure 6: EKC estimation results for Equation (1): actual and fitted values. The dashed lines show the actual values, the solid lines the FM-SUR fitted values and the solid lines with square symbols the group-wise pooled FM-SUR fitted values.

displayed in Figures 3 and 4.[21] 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.

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 also be applied 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 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, need not be identical (potentially up to country specific individual effects) across countries. This, however, is the the key assumption underlying pooling which panel data analysis rests upon. Our SUCPR approach addresses these three issues and provides new tools for group-wise poolability testing and, in case the restrictions are not rejected, corresponding group-wise 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 six early industrialized

[21] 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.

countries over the period 1870–2013. It turns out that the trend respectively slope parameters can 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).

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.[22] 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.[23] 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.

[22] 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.

[23] 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.

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. We furthermore acknowledge the comments of seminar participants at the Friedrich-Alexander-University Erlangen-Nürnberg and the Vienna University of Technology. The usual disclaimer applies.

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*, Ch. 28, 1749–1821. 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*, 179–222. Elsevier, Amsterdam.

Ke, Y., J. Li, and W. Zhang (2016). Structure Identification in Panel Data Analysis. *Annals of Statistics* **44**, 1193–1233.

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*, 97–130. 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., P. Grabarczyk, R. Kawka, and M. Wagner (2016). “Standard” Fully Modified OLS Estimation of Cointegrating Polynomial Regressions. Mimeo.

Su, L., Z. Shi, and P.C.B. Phillips (2016). Identifying Latent Structures in Panel Data. *Econometrica* **84**, 2215–2264.

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, Tokyo.

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)

We consider the three cases of pooling mentioned in the main text and start with defining the quantities corresponding to the three cases. First, we define the three restriction matrices and then we present the correspondingly pooled estimators and their asymptotic distributions.

The Restriction Matrices

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

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

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

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

$$R^{(\mathrm{P})} = \begin{bmatrix} (0_{3\times 1}, I_3) & (0_{3\times 1}, -I_3) & 0_{3\times 4} & \dots & 0_{3\times 4} \\ \vdots & 0_{3\times 4} & (0_{3\times 1}, -I_3) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0_{3\times 4} \\ (0_{3\times 1}, I_3) & 0_{3\times 4} & \dots & 0_{3\times 4} & (0_{3\times 1}, -I_3) \end{bmatrix} \in \mathbb{R}^{3(N-1)\times 4N},\ r = 0_{3(N-1)\times 1}$$

$$R^{(\mathrm{S})} = \begin{bmatrix} (0_{2\times 2}, I_2) & (0_{2\times 2}, -I_2) & 0_{2\times 4} & \dots & 0_{2\times 4} \\ \vdots & 0_{2\times 4} & (0_{2\times 2}, -I_2) & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0_{2\times 4} \\ (0_{2\times 2}, I_2) & 0_{2\times 4} & \dots & 0_{2\times 4} & (0_{2\times 2}, -I_2) \end{bmatrix} \in \mathbb{R}^{2(N-1)\times 4N},\ r = 0_{2(N-1)\times 1}$$

$$R^{(\mathrm{T})} = \begin{bmatrix} (0, 1, 0_{1\times 2}) & (0, -1, 0_{1\times 2}) & 0_{1\times 4} & \dots & 0_{1\times 4} \\ \vdots & 0_{1\times 4} & (0, -1, 0_{1\times 2}) & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0_{1\times 4} \\ (0, 1, 0_{1\times 2}) & 0_{1\times 4} & \dots & 0_{1\times 4} & (0, -1, 0_{1\times 2}) \end{bmatrix} \in \mathbb{R}^{(N-1)\times 4N},\ r = 0_{(N-1)\times 1}$$

Pooled Estimation

In case the respective null hypotheses are not rejected, correspondingly pooled estimation is the 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 limit processes. We discuss the three given cases in turn and start by defining the necessary adapted quantities:

(P) :

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

$$X_t^{(\mathrm{P})} := \begin{bmatrix} t & t & \dots & t \\ x_{1,t} & x_{2,t} & \dots & x_{N,t} \\ x_{1,t}^2 & x_{2,t}^2 & \dots & x_{N,t}^2 \end{bmatrix}$$

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

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

(S) :

$$Z^{(\mathrm{S})} := \begin{bmatrix} Z_1^{(\mathrm{S})\prime} \\ Z_2^{(\mathrm{S})\prime} \\ \vdots \\ Z_T^{(\mathrm{S})\prime} \end{bmatrix}, \quad Z_t^{(\mathrm{S})} := \begin{bmatrix} D_{1,t} & 0_{2\times 1} & \dots & 0_{2\times 1} \\ 0_{2\times 1} & D_{2,t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_{2\times 1} \\ 0_{2\times 1} & \dots & 0_{2\times 1} & D_{N,t} \\ x_{1,t} & x_{2,t} & \dots & x_{N,t} \\ x_{1,t}^2 & x_{2,t}^2 & \dots & x_{N,t}^2 \end{bmatrix}, \quad \theta^{(\mathrm{S})} := \begin{bmatrix} c_1 \\ \delta_1 \\ \vdots \\ c_N \\ \delta_N \\ \beta_1 \\ \beta_2 \end{bmatrix},$$

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

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

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

(T) :

$$
Z^{(\mathrm{T})} := \begin{bmatrix} Z_1^{(\mathrm{T})\prime} \\ Z_2^{(\mathrm{T})\prime} \\ \vdots \\ Z_T^{(\mathrm{T})\prime} \end{bmatrix}, \quad Z_t^{(\mathrm{T})} := \begin{bmatrix} X_{1,t}^{(\mathrm{T})} & 0_{4\times 1} & \dots & 0_{4\times 1} \\ 0_{4\times 1} & X_{2,t}^{(\mathrm{T})} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0_{4\times 1} \\ 0_{4\times 1} & \dots & 0_{4\times 1} & X_{N,t}^{(\mathrm{T})} \\ t & t & \dots & t \end{bmatrix}, \quad \theta^{(\mathrm{T})} := \begin{bmatrix} c_1 \\ \beta_{1,1} \\ \beta_{2,1} \\ \vdots \\ c_N \\ \beta_{1,N} \\ \beta_{2,N} \\ \delta \end{bmatrix},
$$

$$
X_{i,t}^{(\mathrm{T})} := \begin{bmatrix} 1 \\ x_{i,t} \\ x_{i,t}^2 \end{bmatrix}, \quad A^{(\mathrm{T})*} := \begin{bmatrix} A_1^{(\mathrm{T})*} \\ A_2^{(\mathrm{T})*} \\ \vdots \\ A_N^{(\mathrm{T})*} \end{bmatrix}, \quad \tilde{A}^{(\mathrm{T})*} := \begin{bmatrix} \tilde{A}_1^{(\mathrm{T})*} \\ \tilde{A}_2^{(\mathrm{T})*} \\ \vdots \\ \tilde{A}_N^{(\mathrm{T})*} \end{bmatrix},
$$

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

$$
G^{(\mathrm{T})} := \mathrm{diag}\left(I_N \otimes G_1^{(\mathrm{T})}, T^{-3/2}\right), \quad G_1^{(\mathrm{T})} := \mathrm{diag}\left(T^{-1/2}, T^{-1}, T^{-3/2}\right)
$$

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

Corollary 1 (Based on 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 correspondingly pooled FM-SOLS and FM-SUR estimators are, using the quantities defined above, given by:*

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

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

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

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

$$\hat{\theta}^{(T)} := \left(Z^{(T)\prime}Z^{(T)}\right)^{-1}\left(Z^{(T)\prime}y^{+} - \begin{bmatrix} A^{(T)*} \\ 0\end{bmatrix}\right), \tag{38}$$

$$\tilde{\theta}^{(T)} := \left(Z^{(T)\prime}\left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right)Z^{(T)}\right)^{-1}\left(Z^{(T)\prime}\left(I_T \otimes \hat{\Omega}_{u.v}^{-1}\right)y^{+} - \begin{bmatrix} \tilde{A}^{(T)*} \\ 0\end{bmatrix}\right). \tag{39}$$

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

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

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

$$\left(G^{(S)}\right)^{-1}\left(\hat{\theta}^{(S)} - \theta^{(S)}\right) \Rightarrow \left(\int J^{(S)}J^{(S)\prime}\right)^{-1}\int J^{(S)}dB_{u.v}, \tag{42}$$

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

$$\left(G^{(T)}\right)^{-1}\left(\hat{\theta}^{(T)} - \theta^{(T)}\right) \Rightarrow \left(\int J^{(T)}J^{(T)\prime}\right)^{-1}\int J^{(T)}dB_{u.v}, \tag{44}$$

$$\left(G^{(T)}\right)^{-1}\left(\tilde{\theta}^{(T)} - \theta^{(T)}\right) \Rightarrow \left(\int J^{(T)}\Omega_{u.v}^{-1}J^{(T)\prime}\right)^{-1}\int J^{(T)}\Omega_{u.v}^{-1}dB_{u.v}. \tag{45}$$

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

Proof of Proposition 3:

Deriving the limiting distribution of FM-type estimators commences from the limiting distribution of the underlying OLS and – in the SUR case additionally – the MSUR estimator. We start with the group-wise pooled OLS estimator, which is defined as:

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

After centering around the true value θ^{GW} and pre-multiplying with the scaling matrix $\check{G}$ defined in the main text we arrive at:

$$\begin{aligned}\check{G}^{-1}\left(\hat{\theta}_{\text{OLS}}^{\text{GW}} - \theta^{\text{GW}}\right) &= \left(\check{G}\check{Z}'\check{Z}\check{G}\right)^{-1}\left(\check{G}\check{Z}'u\right) \\ &= \left(\sum_{t=1}^{T}\check{G}\check{Z}_t\check{Z}_t'\check{G}\right)^{-1}\left(\sum_{t=1}^{T}\check{G}\check{Z}_t u_t\right)\end{aligned} \tag{47}$$

It follows that $\sum_{t=1}^{T}\check{G}\check{Z}_t\check{Z}_t'\check{G} = \frac{1}{T}\sum_{t=1}^{T}\sqrt{T}\check{G}\check{Z}_t\check{Z}_t'\check{G}\sqrt{T} \Rightarrow \int \check{J}\check{J}'$, using $\lim_{T\to\infty}\sqrt{T}\check{G}\check{Z}_{\lfloor rT\rfloor} = \check{J}(r)$ for $0 \leq r \leq 1$ and the continuous mapping theorem. For the second term similar arguments as in Hong and Wagner (2014, Propositions 1 and 4) – without group-wise pooling in that paper – can be used to establish:

$$\sum_{t=1}^{T}\check{G}\check{Z}_t u_t = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\sqrt{T}\check{G}\check{Z}_t u_t = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\begin{bmatrix}\sqrt{T}\check{G}_c u_t \\ \sqrt{T}\check{G}_D\check{D}_t u_t \\ \sqrt{T}\check{G}_X\check{X}_t u_t\end{bmatrix} \Rightarrow \int \check{J}dB_u + F_u^{\text{GW}}, \tag{48}$$

with $F_u^{\text{GW}} := [0_{1\times(N+n_k)}, F_{u,1}^{\text{GW}}{}', \ldots, F_{u,l}^{\text{GW}}{}']'$, where

$$F_{u,j}^{\text{GW}} := \sum_{i=1}^{N}\mathbb{1}_{\{i\in I_{m_j}\}}\cdot\Delta_{vu}^{i,i}\cdot\begin{pmatrix}1 \\ 2\int B_{v_i}\end{pmatrix}, \quad j = 1,\ldots,l. \tag{49}$$

Altogether this implies that

$$\begin{aligned}\check{G}^{-1}\left(\hat{\theta}_{\text{OLS}}^{\text{GW}} - \theta^{\text{GW}}\right) &\Rightarrow \left(\check{J}\check{J}'\right)^{-1}\left(\int \check{J}dB_u + F_u^{\text{GW}}\right) \\ &= \left(\check{J}\check{J}'\right)^{-1}\left(\int \check{J}dB_{u\cdot v} + \int \check{J}\Omega_{uv}\Omega_{vv}^{-1}dB_v + F_u^{\text{GW}}\right).\end{aligned} \tag{50}$$

We now turn to the group-wise pooled MSUR estimator:

$$\tilde{\theta}_{\text{MSUR}}^{\text{GW}} := \left(\check{Z}'\left(I_T\otimes\hat{\Omega}_{uu}^{-1}\right)\check{Z}\right)^{-1}\check{Z}'\left(I_T\otimes\hat{\Omega}_{uu}^{-1}\right)u, \tag{51}$$

which after centering and scaling can be written as:

$$\check{G}^{-1}\left(\tilde{\theta}^{\text{GW}}_{\text{MSUR}} - \theta^{\text{GW}}\right) = \left(\sum_{t=1}^{T} \check{G}\check{Z}_t\hat{\Omega}_{uu}^{-1}\check{Z}_t'\check{G}\right)^{-1}\left(\sum_{t=1}^{T} \check{G}\check{Z}_t\hat{\Omega}_{uu}^{-1}u_t\right). \tag{52}$$

Since $\hat{\Omega}_{uu} \to \Omega_{uu}$ by assumption, it immediately follows that $\sum_{t=1}^{T} \check{G}\check{Z}_t\hat{\Omega}_{uu}^{-1}\check{Z}_t\check{G}' \Rightarrow \int \check{J}\Omega_{uu}^{-1}\check{J}'$ and for the second term we now obtain using again similar arguments as in Hong and Wagner (2014, Propositions 1 and 4):

$$\sum_{t=1}^{T} \check{G}\check{Z}_t\hat{\Omega}_{uu}^{-1}u_t \Rightarrow \int \check{J}\Omega_{uu}^{-1}dB_u + \tilde{F}_u^{\text{GW}}, \tag{53}$$

with $\tilde{F}_u^{\text{GW}} := [0_{1\times(N+n_k)}, \tilde{F}_{u,1}^{\text{GW}\prime}, \ldots, \tilde{F}_{u,l}^{\text{GW}\prime}]'$, where

$$\tilde{F}_{u,j}^{\text{GW}} := \sum_{i=1}^{N} \mathbb{1}_{\{i\in I_{m_j}\}} \cdot \Delta_{vu}^{i,\cdot} \cdot \left(\Omega_{uu}^{-1}\right)^{\cdot,i} \cdot \begin{pmatrix} 1 \\ 2\int B_{v_i} \end{pmatrix}, \quad j = 1, \ldots, l. \tag{54}$$

Combining the two terms gives the asymptotic distribution of the MSUR estimator:

$$\begin{aligned} \check{G}^{-1}\left(\tilde{\theta}^{\text{GW}}_{\text{MSUR}} - \theta^{\text{GW}}\right) &\Rightarrow \left(\int \check{J}\Omega_{uu}^{-1}\check{J}'\right)^{-1}\left(\int \check{J}\Omega_{uu}^{-1}dB_u + \tilde{F}_u^{\text{GW}}\right) \\ &= \left(\int \check{J}\Omega_{uu}^{-1}\check{J}'\right)^{-1}\left(\int \check{J}\Omega_{uu}^{-1}dB_{u\cdot v} + \int \check{J}\Omega_{uu}^{-1}\Omega_{uv}\Omega_{vv}^{-1}dB_v + \tilde{F}_u^{\text{GW}}\right) \end{aligned} \tag{55}$$

Having the asymptotic distributions of the OLS and MSUR estimators at hand allows to next derive the asymptotic distribution of the FM-OLS and FM-SUR estimators for the group-wise pooled case. We start with FM-SOLS. From the definition of the estimator in (27) and the definition of y^+ it is clear that the essential term to be understood, after centering and scaling, is

$$\sum_{t=1}^{T} \check{G}\check{Z}_t\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_t \Rightarrow \int \check{J}\Omega_{uv}\Omega_{vv}^{-1}dB_v + F_v^{\text{GW}}, \tag{56}$$

with $F_v^{\text{GW}} := [0_{1\times(N+n_k)}, F_{v,1}^{\text{GW}\prime}, \ldots, F_{v,l}^{\text{GW}\prime}]'$, where

$$F_{v,j}^{\text{GW}} := \sum_{i=1}^{N} \mathbb{1}_{\{i\in I_{m_j}\}} \cdot \Omega_{uv}^{i,\cdot} \cdot \Omega_{vv}^{-1} \cdot \Delta_{vv}^{\cdot,i} \cdot \begin{pmatrix} 1 \\ 2\int B_{v_i} \end{pmatrix}, \quad j = 1, \ldots, l. \tag{57}$$

The above result together with (50) implies that

$$\begin{aligned} \sum_{t=1}^{T} \check{G}\check{Z}_t u_t^+ &= \sum_{t=1}^{T} \check{G}\check{Z}_t u_t - \sum_{t=1}^{T} \check{G}\check{Z}_t\hat{\Omega}_{uv}\hat{\Omega}_{vv}^{-1}v_t \\ &\Rightarrow \int \check{J}dB_{u\cdot v} + \int \check{J}\Omega_{uv}\Omega_{vv}^{-1}dB_v + F_u^{\text{GW}} - \int \check{J}\Omega_{uv}\Omega_{vv}^{-1}dB_v - F_v^{\text{GW}} \\ &= \int \check{J}dB_{u\cdot v} + F_u^{\text{GW}} - F_v^{\text{GW}} = \int \check{J}dB_{u\cdot v} + A^{\text{GW}}. \end{aligned} \tag{58}$$

Observing now that the non-zero blocks of A^{GW} are given by

$$A_j^{\text{GW}} = \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{m_j}\}} \cdot \left(\Delta_{vu}^{i,i} - \Omega_{uv}^{i,\cdot} \cdot \Omega_{vv}^{-1} \cdot \Delta_{vv}^{\cdot,i}\right) \cdot \begin{pmatrix} 1 \\ 2\int B_{v_i} \end{pmatrix}, \quad j = 1, \ldots, l, \tag{59}$$

$$= \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{m_j}\}} \cdot \left(\Delta_{vu}^{+}\right)^{i,i} \cdot \begin{pmatrix} 1 \\ 2\int B_{v_i} \end{pmatrix}, \quad j = 1, \ldots, l,$$

shows that $\check{G}A^{\text{GW}*} \Rightarrow A^{\text{GW}}$, which establishes the result (29) for the group-wise FM-SOLS estimator.

A similar reasoning also applies to the group-wise pooled FM-SUR estimator defined in (28). Here the relevant term to consider is

$$\sum_{t=1}^{T} \check{G}\check{Z}_t \hat{\Omega}_{u\cdot v}^{-1} \hat{\Omega}_{uv} \hat{\Omega}_{vv}^{-1} v_t \Rightarrow \int \check{J} \Omega_{u\cdot v}^{-1} \Omega_{uv} \Omega_{vv}^{-1} dB_v + \tilde{F}_v^{\text{GW}}, \tag{60}$$

with $\tilde{F}_v^{\text{GW}} := [0_{1\times(N+n_k)}, \tilde{F}_{v,1}^{\text{GW}}{}', \ldots, \tilde{F}_{v,l}^{\text{GW}}{}']'$, where

$$\tilde{F}_{v,j}^{\text{GW}} := \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{m_j}\}} \cdot \left(\Omega_{u\cdot v}^{-1}\right)^{i,\cdot} \cdot \Omega_{uv} \cdot \Omega_{vv}^{-1} \cdot \Delta_{vv}^{\cdot,i} \cdot \begin{pmatrix} 1 \\ 2\int B_{v_i} \end{pmatrix}, \quad j = 1, \ldots, l. \tag{61}$$

The above result together with (55) implies that

$$\begin{aligned}
\sum_{t=1}^{T} \check{G}\check{Z}_t \hat{\Omega}_{u\cdot v}^{-1} u_t^{+} &= \sum_{t=1}^{T} \check{G}\check{Z}_t \Omega_{u\cdot v}^{-1} u_t - \sum_{t=1}^{T} \check{G}\check{Z}_t \Omega_{u\cdot v}^{-1} \hat{\Omega}_{uv} \hat{\Omega}_{vv}^{-1} v_t \\
&\Rightarrow \int \check{J} \Omega_{u\cdot v}^{-1} dB_{u\cdot v} + \int \check{J} \Omega_{u\cdot v}^{-1} \Omega_{uv} \Omega_{vv}^{-1} dB_v + \tilde{F}_u^{\text{GW}} - \int \check{J} \Omega_{u\cdot v}^{-1} \Omega_{uv} \Omega_{vv}^{-1} dB_v - \tilde{F}_v^{\text{GW}} \\
&= \int \check{J} dB_{u\cdot v} + \tilde{F}_u^{\text{GW}} - \tilde{F}_v^{\text{GW}} = \int \check{J} \Omega_{u\cdot v}^{-1} dB_{u\cdot v} + \tilde{A}^{\text{GW}}.
\end{aligned} \tag{62}$$

The non-zero blocks of $\tilde{A}^{\text{GW}}$ are given by

$$\tilde{A}_j^{\text{GW}} = \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{m_j}\}} \cdot \left(\Delta_{vu}^{i,\cdot} \cdot \left(\Omega_{u\cdot v}^{-1}\right)^{\cdot,i} - \left(\Omega_{u\cdot v}^{-1}\right)^{i,\cdot} \cdot \Omega_{uv} \cdot \Omega_{vv}^{-1} \cdot \Delta_{vv}^{\cdot,i}\right) \cdot \begin{pmatrix} 1 \\ 2\int B_{v_i} \end{pmatrix}, \quad j = 1, \ldots, l, \tag{63}$$

$$= \sum_{i=1}^{N} \mathbb{1}_{\{i \in I_{m_j}\}} \cdot \left(\Delta_{vu}^{+}\right)^{i,\cdot} \cdot \left(\Omega_{u\cdot v}^{-1}\right)^{\cdot,i} \cdot \begin{pmatrix} 1 \\ 2\int B_{v_i} \end{pmatrix}, \quad j = 1, \ldots, l,$$

which implies that $\check{G}\tilde{A}^{\text{GW}*} \Rightarrow \tilde{A}^{\text{GW}}$. □

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	0.945	0.951	0.380	-1.280	-1.304	-1.378
Austria	0.004	0.098	-0.115	-1.908	-1.813	-1.878
Belgium	0.795	0.620	0.225	-1.419	-1.554	-1.587
Canada	-0.255	-0.348	-0.458	-2.496	*-3.057*	*-3.056*
Denmark	0.038	0.062	-0.270	-2.293	-2.270	-2.273
Finland	0.727	0.586	0.169	-2.321	-2.422	-2.412
France	-0.044	-0.159	-0.286	-1.958	-2.204	-2.214
Germany	-0.256	-0.338	-0.317	-2.356	-2.582	-2.598
Italy	0.524	0.234	0.004	-1.665	-1.825	-1.835
Japan	0.222	0.105	-0.134	-1.719	-1.880	-1.893
Netherlands	0.163	0.092	-0.007	-2.184	-2.310	-2.334
New Zealand	-0.100	-0.101	-0.139	-2.606	-2.686	-2.688
Norway	0.909	0.838	0.116	-2.142	-2.177	-2.179
Portugal	1.271	0.997	0.352	-1.872	-1.879	-1.869
Spain	0.704	0.428	0.050	-1.005	-1.192	-1.235
Sweden	0.137	0.131	-0.276	-2.339	-2.401	-2.407
Switzerland	-1.004	-1.053	-1.001	-2.807	-2.449	-2.481
UK	1.118	1.383	0.425	-1.567	-1.418	-1.625
USA	-0.336	-0.308	-0.447	-2.981	*-2.911*	*-2.915*

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.

	PO_t	Shin	$P_{\hat{u}}$	CT
Australia	-2.727	**0.129**	11.330	**0.108**
Austria	-3.785	0.077	**55.997**	0.056
Belgium	**-5.586**	0.054	*50.269*	0.062
Canada	-3.349	**0.189**	12.420	**0.145**
Denmark	**-4.937**	0.053	40.613	0.052
Finland	**-5.656**	0.045	**75.016**	0.050
France	**-4.948**	0.060	28.847	0.066
Germany	**-7.895**	**0.411**	**68.343**	**0.111**
Italy	**-4.163**	**0.182**	34.141	**0.146**
Japan	**-5.829**	**0.155**	8.127	**0.152**
Netherlands	**-5.688**	**0.108**	**96.172**	0.074
New Zealand	**-5.375**	**0.132**	13.337	**0.115**
Norway	-3.530	*0.097*	20.644	*0.095*
Portugal	**-9.111**	**0.101**	19.959	**0.111**
Spain	-3.343	*0.091*	42.578	0.086
Sweden	**-4.268**	*0.085*	28.679	0.085
Switzerland	**-6.282**	0.065	**85.979**	0.084
UK	**-7.673**	*0.085*	**97.169**	0.073
USA	-2.448	**0.576**	13.920	**0.156**
Critical Values ($\alpha = 0.1$)	-4.1567	0.081	45.237	0.086
Critical Values ($\alpha = 0.05$)	-3.8429	0.101	52.952	0.106

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.

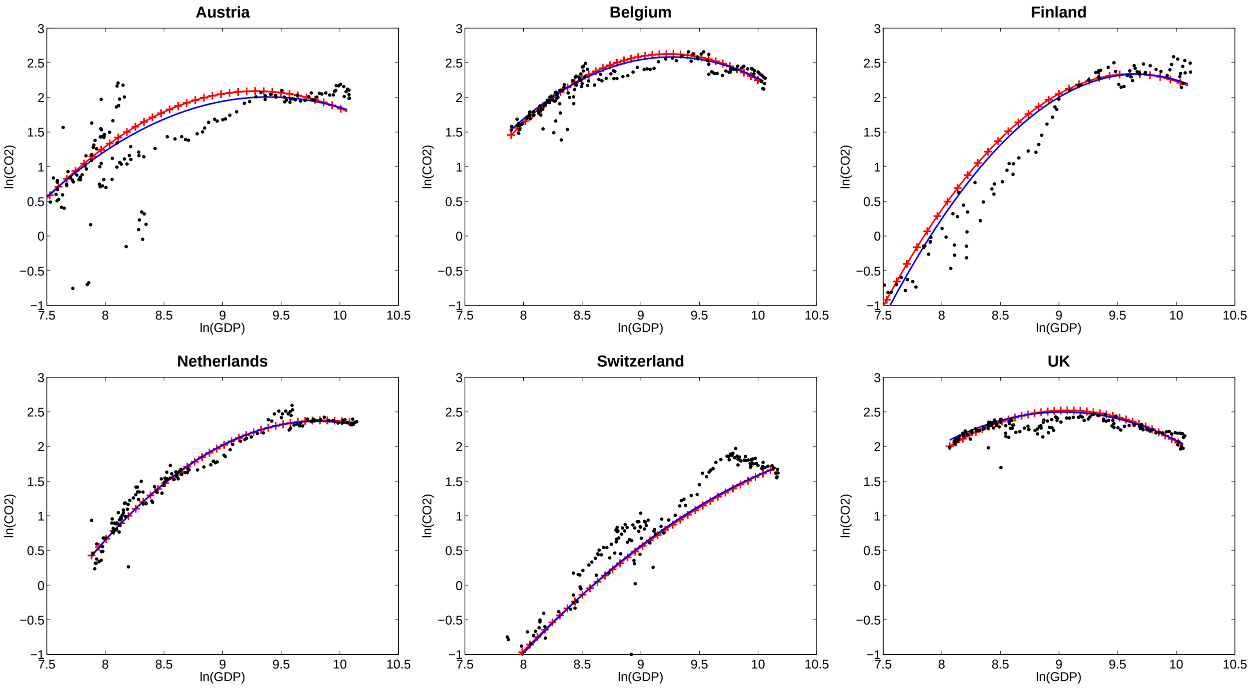

Figure 7: EKC estimation results for Equation (1): scatter plot and EKC. The solid lines correspond to the FM-SUR estimates and the solid lines with +-marks to the FM-SOLS estimates. For further explanations see notes to Figure 1.

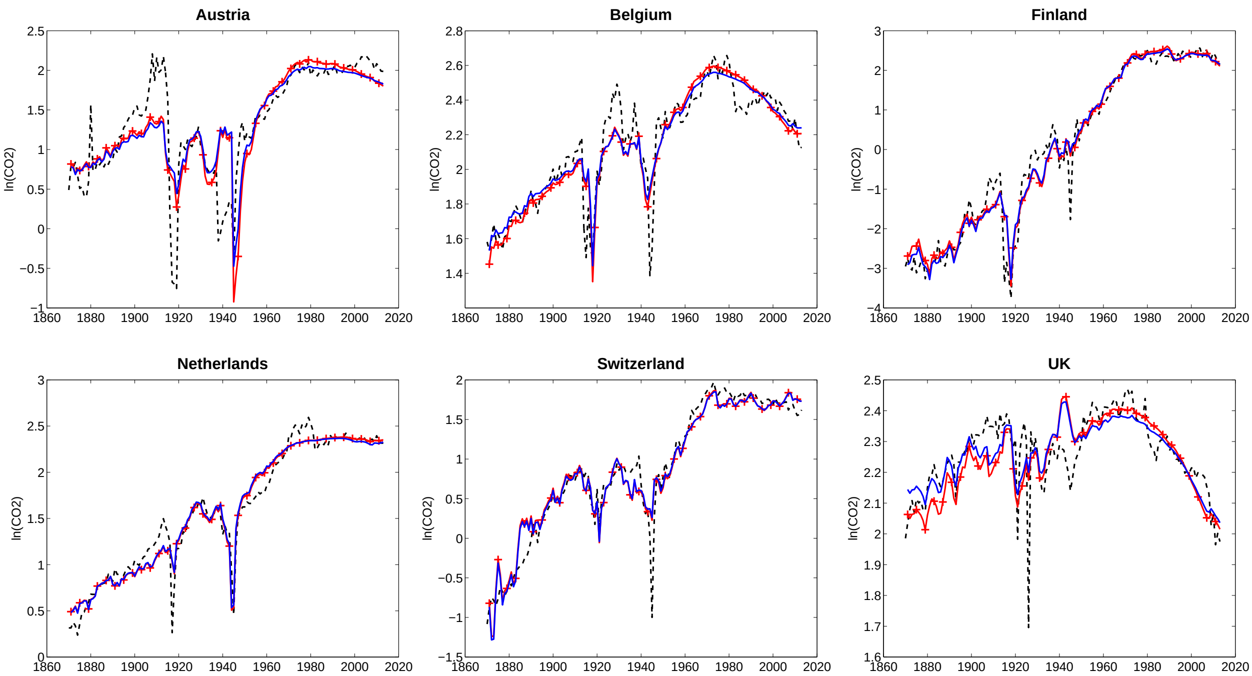

Figure 8: EKC estimation results for Equation (1): actual and fitted values. The dashed lines show the actual values, the solid lines the FM-SUR fitted values and the solid lines with +-marks the FM-SOLS fitted values.

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

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

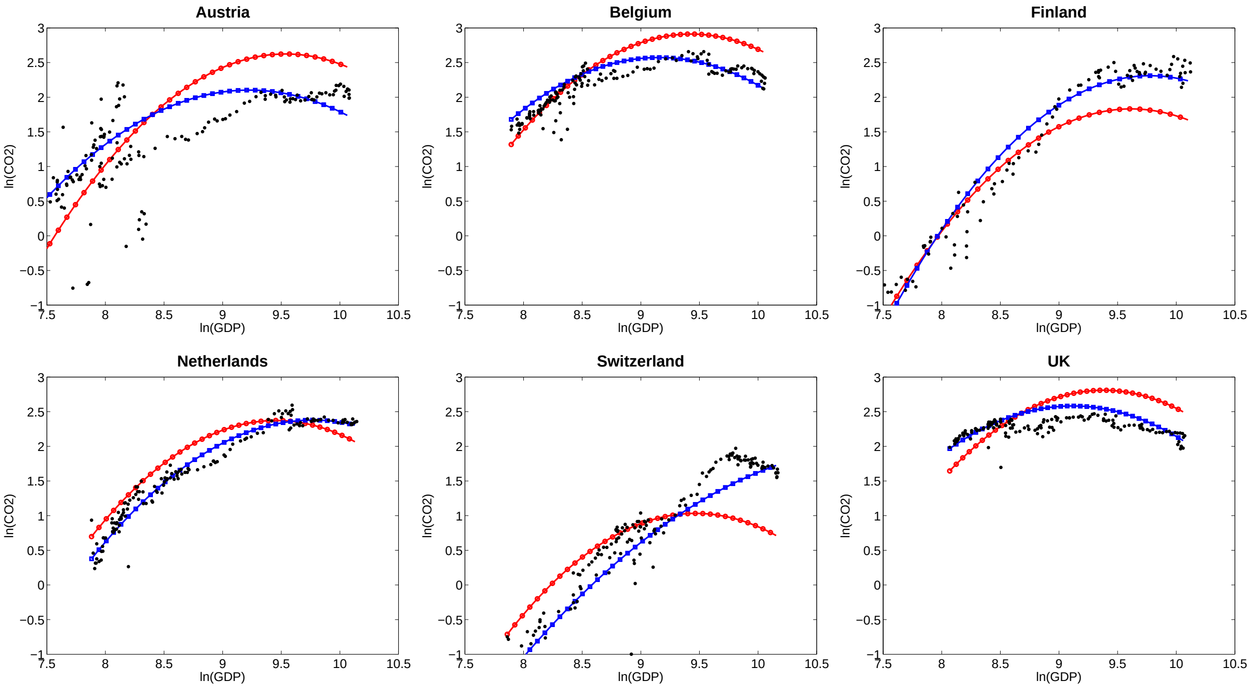

Figure 9: EKC estimation results for Equation (1): scatter plot and EKC. The solid lines with square symbols correspond to the group-wise pooled FM-SUR estimates and the solid lines with o-marks to the pooled FM-SUR estimates. For further explanations see notes to Figure 1.

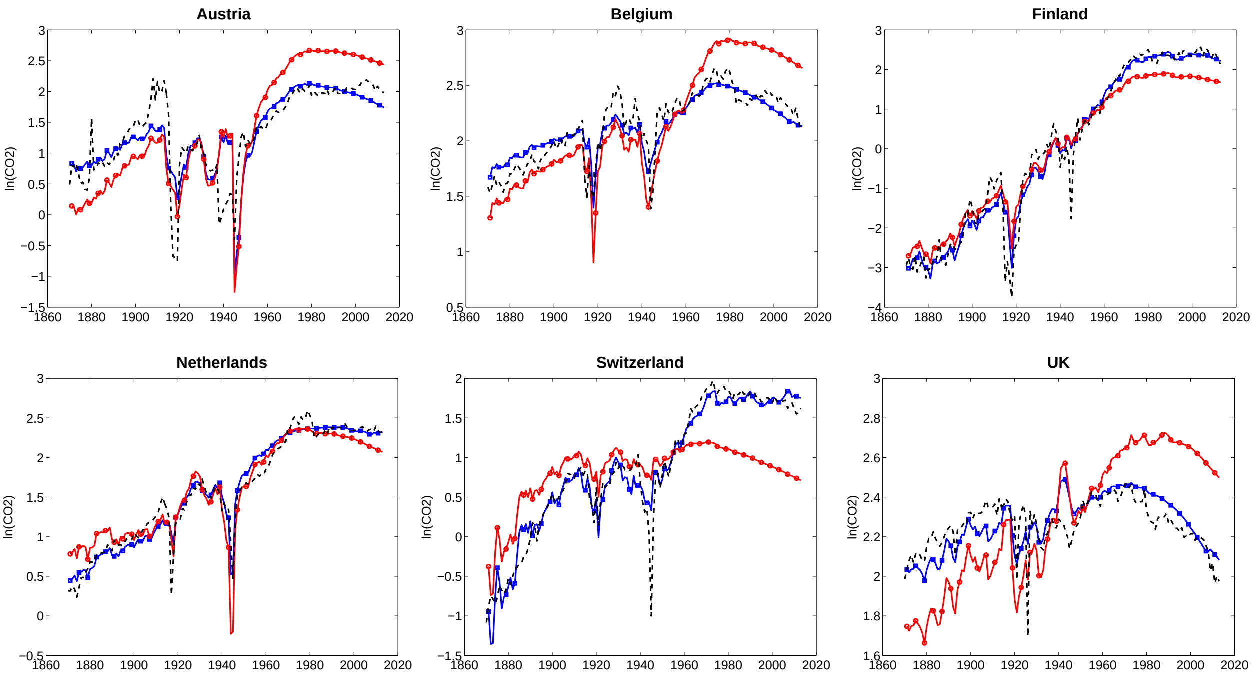

Figure 10: EKC estimation results for Equation (1): actual and fitted values. The dashed lines show the actual values, the solid lines with square symbols the group-wise pooled FM-SUR fitted values and the solid lines with o-marks the pooled FM-SUR fitted values.