Banka Slovenije Working Papers

Integrated Modified Least Squares Estimation and (Fixed-b) Inference for Systems of Cointegrating Multivariate Polynomial Regressions

Authors: Sebastian Veldhuis, Martin Wagner

December 2024

BANKA SLOVENIJE

EVROSISTEM

Collection: Banka Slovenije Working Papers

Title: Integrated Modified Least Squares Estimation and (Fixed-b) Inference for Systems of Cointegrating Multivariate Polynomial Regressions

Authors: Sebastian Veldhuis, Martin Wagner

Issue: December 2024

Year: 2024

Place of publication: Ljubljana

Issued by:
Banka Slovenije
Slovenska 35,
1505 Ljubljana, Slovenija
www.bsi.si

Electronic edition:
https://www.bsi.si/en/publications/research/banka-slovenije-working-papers

Kataložni zapis o publikaciji (CIP) pripravili v Narodni in univerzitetni knjižnici v Ljubljani
COBISS.SI-ID 219690499
ISBN 978-961-7230-11-6 (PDF)

Integrated Modified Least Squares Estimation and (Fixed-b) Inference for Systems of Cointegrating Multivariate Polynomial Regressions

Sebastian Veldhuis[1] and Martin Wagner[1,2,3]

[1]Department of Economics, University of Klagenfurt
[2]Bank of Slovenia, Ljubljana
[3]Institute for Advanced Studies, Vienna

Abstract

We consider integrated modified ordinary and generalized least squares estimation for systems of cointegrating multivariate polynomial regressions, i. e., systems of regressions that include deterministic variables, integrated processes and products of non-negative integer powers of these variables as regressors. The stationary errors are allowed to be correlated across equations, over time and with the regressors. The necessity to consider integrated modified generalized least squares estimation arises in case of estimation under restrictions, which in general implies that ordinary and generalized least squares estimators cease to be identical. We discuss in detail hypothesis testing for the unrestricted and restricted estimators. Furthermore, we develop asymptotically pivotal fixed-b inference, which is shown to be available only in the case of *full design* for up-to-the-intercept-identical hypotheses tested in all equations in systems with identical regressors in all equations.

JEL Classification: C12, C13, C32

Keywords: Integrated Modified Estimation, Cointegrating Multivariate Polynomial Regression, Fixed-b Inference, Generalized Least Squares

Povzetek

Obravnavamo integrirano modificirano navadno in posplošeno ocenjevanje najmanjših kvadratov za sisteme kointegrirajočih multivariatnih polinomskih regresij, tj. sisteme regresij, ki kot regresorje vključujejo deterministične spremenljivke, integrirane procese in produkte nenegativnih celoštevilskih moči teh spremenljivk. Dovoljeno je, da so stacionarne napake korelirane med enačbami, skozi čas in z regresorji. Potreba po upoštevanju integriranega modificiranega posplošenega ocenjevanja po metodi najmanjših kvadratov se pojavi v primeru ocenjevanja z omejitvami, kar na splošno pomeni, da običajne in posplošene cenilke po metodi najmanjših kvadratov niso več identične. Podrobno obravnavamo testiranje hipotez za neomejene in omejene cenilke. Poleg tega razvijamo asimptotično pivotalno inferenco s fiksno-b, ki je na voljo le v primeru popolnega načrta za hipoteze, identične do prestreznega člena, testirane v vseh enačbah v sistemih z enakimi regresorji v vseh enačbah.

1 Introduction

This paper considers integrated modified (IM) least squares estimation – both OLS and GLS – for systems of cointegrating multivariate polynomial regressions (SCMPRs). These are systems of regressions that include deterministic variables, integrated processes and products of (non-negative) integer powers of these variables as regressors. The error terms, assumed to be jointly stationary across equations, are allowed to be correlated – both serially and across equations – and the stochastic regressors are allowed to be endogenous. Thus, the paper extends the analysis of *univariate* cointegrating multivariate polynomial regressions (CMPRs) of Vogelsang and Wagner (2024) to the systems setting.[1]

Integrated modified ordinary least squares (IM-OLS) estimation, introduced for cointegrating linear regressions in Vogelsang and Wagner (2014), has several key advantages compared to other modified least squares estimators used in cointegrating regression analysis: First, estimation is tuning-parameter-free. IM-OLS requires the estimation of a conditional long-run covariance matrix for inference only. Second, IM estimation allows performing fixed-b inference in cointegrating regressions (see Vogelsang and Wagner, 2014, Section 5). Fixed-b inference is designed to capture the impact of kernel and bandwidth choices, required for estimating the above-mentioned conditional long-run covariance matrix, on the sampling distributions of test statistics.[2] Third, and this is a very important conceptual advantage, IM estimation can be extended straightforwardly to allow for the inclusion not only of powers of integrated processes as regressors (the CPR case), but also of arbitrary non-negative integer products of integrated processes as regressors.[3]

[1] This paper fulfills a similar (but more encompassing) extension-to-systems-of-equations role as Wagner (2023) fulfills for Wagner and Hong (2016). These two earlier papers – discussing fully modified least squares estimation – only consider (systems of) cointegrating polynomial regressions (CPRs), in which cross-products (of non-negative integer powers) of integrated processes are not included as regressors. Note that Wagner (2023) contains a typo in the definition of $\hat{M}^+$ below equation (7). The u and v subscripts in the $\hat{\Delta}$-terms in the definition of $\hat{M}_u$ need to be switched, e. g., in the i-th row $\hat{\Delta}_{u_i v_j}$ needs to be replaced by $\hat{\Delta}_{v_j u_i}$ for $j = 1, \ldots, m$. The sentence should then continue with: "and $\hat{M}_v$ defined analogously, with $\hat{\Delta}_{v_j u_i}$, $i = 1, \ldots, n$, $j = 1, \ldots, m$ replaced by $\hat{\Delta}_{v_j v_i}$, $i, j = 1, \ldots, m$."

[2] Fixed-b analysis of spectral estimators has been introduced by Neave (1970). It has been developed into an alternative framework for (robust) inference for stationary regressions in Kiefer and Vogelsang (2005).

[3] Note that this overcomes, admittedly only for the case of multivariate polynomials, the additive separability between integrated regressors that is most often assumed in the nonlinear cointegration literature. For a more encompassing discussion of this aspect see Vogelsang and Wagner (2024, Section 1).

One important (economic) application of CMPRs, i. e., of regressions involving cross-products of integrated regressors, are so-called *Translog* functions, see, e. g., Christensen *et al.* (1971) or the example in Remark 1. A second important application is RESET-type specification testing (originally introduced by Ramsey, 1969) of cointegrating (multivariate polynomial) regressions using the version of Thursby and Schmidt (1977), where (non-negative integer) powers and cross-products of powers of the regressors in the original regression are included in an augmented test regression. See the discussion in Vogelsang and Wagner (2024, Sections 2.4 and 3) for the single-equation case.[4]

Considering systems of equations, useful to, e. g., analyze Translog cost or production function systems with several outputs, adds some further aspects compared to the single-equation setting discussed in Vogelsang and Wagner (2024). First, see also the corresponding discussion in Wagner (2023), systems of equations necessitate a detailed consideration of *generalized* least squares estimators, in this paper integrated modified generalized least squares (IM-GLS). This stems from the well-known fact that OLS- and GLS-type estimators coincide in general only (for any positive definite symmetric weighting matrix) in systems with identical regressors in all equations and without parameter constraints. Second, the scope of fixed-b inference needs to be investigated in more detail than in the single-equation case. It turns out that – in addition to *full design*, required also in Vogelsang and Wagner (2024) – fixed-b inference is only available for up-to-the-intercept-identical hypotheses tested in all equations in systems with identical regressors in all equations. Whilst this is, of course, restrictive it includes, e. g., fixed-b RESET-type specification testing for systems of equations with identical regressors in all equations under both the null specification and in the augmented test regression.[5]

This short paper is organized as follows: Section 2 discusses, organized in four subsections, the setup and assumptions, unrestricted estimation and inference, estimation and inference under restrictions and fixed-b inference. Section 3 briefly summarizes and concludes. All proofs are relegated to the appendix. `MATLAB` code for IM-OLS/GLS estimation and inference, including fixed-b inference – which necessitates (the generation of) specification-dependent critical values (that additionally depend upon kernel function and bandwidth) – is available upon request.

[4] Note that, as already discussed in Wagner and Hong (2016, Section 2.3) and Vogelsang and Wagner (2024, Section 2.4) additional integrated regressors (and their non-negative integer powers and cross-products) can be included in the augmented test regression.

[5] The limited scope of fixed-b inference implies that there is no need to consider fixed-b inference for IM-GLS estimators, since IM-OLS and IM-GLS coincide in settings where fixed-b inference is available.

2 Theory

2.1 Setup and Assumptions

We start with considering unrestricted systems of cointegrating multivariate polynomial regressions (SCMPRs) where all equations include the same set of regressors:

$$\begin{aligned} y_t &= \Theta Z_t + u_t, \qquad t = 1, \dots, T, \\ x_t &= x_{t-1} + v_t, \end{aligned} \tag{1}$$

with $y_t := (y_{1t}, \dots, y_{nt})'$, $Z_t := (z_{1t}, \dots, z_{|\mathcal{I}|t})'$, with $z_{it} = t^{i_0} x_{1t}^{i_1} \cdots x_{mt}^{i_m}$ for $i = 1, \dots, |\mathcal{I}|$ and i_j non-negative integers for $j = 0, \dots, m$, $\Theta = [\theta_{h,\mathbf{i}}]_{h=1,\dots,n,\mathbf{i}\in\mathcal{I}} \in \mathbb{R}^{n\times|\mathcal{I}|}$ and $x_t := (x_{1t}, \dots, x_{mt})'$. The regressors z_{it}, $i = 1, \dots, |\mathcal{I}|$ are ordered, e. g., by lexicographic ordering of the multi-indices $\mathbf{i} := (i_0, \dots, i_m)$ from a multi-index set $\mathcal{I}$ indexing all regressors. To avoid perfect multi-collinearity by construction, we assume that no multi-index $\mathbf{i}$ is contained more than once in $\mathcal{I}$.

The results in this paper can be established under the same assumptions, adapted to multivariate y_t, as used in, e. g., Vogelsang and Wagner (2024, Footnote 10) and we, therefore, abstain from positing a detailed set of assumptions. As is common in the cointegrating regression literature, we also exclude cointegration amongst the m integrated regressors $\{x_t\}_{t\in\mathbb{Z}}$ as well as multi-cointegration in the system. Defining $\{\eta_t\}_{t\in\mathbb{Z}} := \{(u_t', v_t')'\}_{t\in\mathbb{Z}}$, a functional central limit theorem holds:

$$T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} \eta_t \Rightarrow B(r) = \begin{pmatrix} B_u(r) \\ B_v(r) \end{pmatrix} = \Omega^{1/2} W(r), \tag{2}$$

for $0 \leq r \leq 1$, with $W(r)$ denoting $(n+m)$-dimensional standard Brownian motion and *by assumption* positive definite long-run covariance matrix:

$$\Omega = \begin{pmatrix} \Omega_{uu} & \Omega_{uv} \\ \Omega_{vu} & \Omega_{vv} \end{pmatrix} := \sum_{j=-\infty}^{\infty} \mathbb{E}(\eta_{t-j}\eta_t'), \tag{3}$$

partitioned conformably with η_t. When $\Omega_{uv} \neq 0$, the regressors are endogenous and the setting also allows for relatively unrestricted forms of serial correlation of the errors $\{\eta_t\}_{t\in\mathbb{Z}}$. Using, e. g., the Cholesky decomposition of $\Omega_{vv} = \Omega_{vv}^{1/2}(\Omega_{vv}^{1/2})'$, one can write (2) more specifically as:

$$\begin{pmatrix} B_u(r) \\ B_v(r) \end{pmatrix} := \begin{pmatrix} \Omega_{u\cdot v}^{1/2} & \Omega_{uv}(\Omega_{vv}^{-1/2})' \\ 0 & \Omega_{vv}^{1/2} \end{pmatrix} \begin{pmatrix} W_{u\cdot v}(r) \\ W_v(r) \end{pmatrix}, \tag{4}$$

with $\Omega_{u\cdot v} := \Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}$ the (innovation) covariance matrix of $B_{u\cdot v}(r) := B_u(r) - \Omega_{uv}\Omega_{vv}^{-1}B_v(r)$.[6]

Remark 1. To exemplify the setting and notation, consider the following simple example (with $n = m = 2$) of a two-output firm – with output "proper" Y_t and (unwanted output) emissions E_t – using a Translog production function with the two input factors capital K_t and labor L_t.[7] Including additionally (equation-specific) intercepts leads to:

$$\begin{aligned}
\ln Y_t &= \theta_{1,(0,0,0)} + \theta_{1,(0,1,0)} \ln K_t + \theta_{1,(0,0,1)} \ln L_t + \theta_{1,(0,2,0)}(\ln K_t)^2 \\
&\quad + \theta_{1,(0,0,2)}(\ln L_t)^2 + \theta_{1,(0,1,1)} \ln K_t \ln L_t + u_{1t}, \\
\ln E_t &= \theta_{2,(0,0,0)} + \theta_{2,(0,1,0)} \ln K_t + \theta_{2,(0,0,1)} \ln L_t + \theta_{2,(0,2,0)}(\ln K_t)^2 \\
&\quad + \theta_{2,(0,0,2)}(\ln L_t)^2 + \theta_{2,(0,1,1)} \ln K_t \ln L_t + u_{2t},
\end{aligned}$$

with the coefficients double-indexed by first the equation index $h = 1, 2$ and second the multi-index $\mathbf{i} = (i_0, i_1, i_2) \in \mathcal{I}$ corresponding to regressor $z_{it} = t^{i_0}(\ln K_t)^{i_1}(\ln L_t)^{i_2}$ with $\mathcal{I} = \{(0,0,0),(0,1,0),(0,0,1),(0,2,0),(0,0,2),(0,1,1)\}$ and $|\mathcal{I}| = 6$. In matrix notation, the above system can be written more compactly as in (1):

[6] Reduced rank of $\Omega_{u\cdot v}$ corresponds to multi-cointegration, excluded in this paper due to the assumption that Ω is positive definite.

[7] When considering emissions as one of the outputs, it is usual to also include energy as production factor. We abstain from including additional production factors as well as a measure of the technology level *merely* for algebraic brevity. In some applications, the inputs are disaggregated into five input factors: capital, labor, energy, materials and services, labelled as KLEMS. For the 27 member countries of the European Union, the *EU-KLEMS* project provides corresponding annual data over the period 1995–2020 for 23 industries. The short sample period makes this data set potentially unsuitable for cointegration analysis (of the long-run behavior), but this question will be investigated in detail elsewhere.

$$\underbrace{\begin{bmatrix} \ln Y_t \\ \ln E_t \end{bmatrix}}_{=y_t} = \underbrace{\begin{bmatrix} \theta_{1,(0,0,0)} & \theta_{1,(0,1,0)} & \theta_{1,(0,0,1)} & \theta_{1,(0,2,0)} & \theta_{1,(0,0,2)} & \theta_{1,(0,1,1)} \\ \theta_{2,(0,0,0)} & \theta_{2,(0,1,0)} & \theta_{2,(0,0,1)} & \theta_{2,(0,2,0)} & \theta_{2,(0,0,2)} & \theta_{2,(0,1,1)} \end{bmatrix}}_{=\Theta} \underbrace{\begin{bmatrix} 1 \\ \ln K_t \\ \ln L_t \\ (\ln K_t)^2 \\ (\ln L_t)^2 \\ \ln K_t \ln L_t \end{bmatrix}}_{=Z_t} + \underbrace{\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}}_{=u_t},$$

$$\underbrace{\begin{bmatrix} \ln K_t \\ \ln L_t \end{bmatrix}}_{=x_t} = \underbrace{\begin{bmatrix} \ln K_{t-1} \\ \ln L_{t-1} \end{bmatrix}}_{=x_{t-1}} + \underbrace{\begin{bmatrix} v_{1t} \\ v_{2t} \end{bmatrix}}_{=v_t}.$$

This two-output Translog example thus corresponds to a system without restrictions on the parameter matrix, since all regressors are included in both equations.

Remark 2. The setting considered in this paper is closely related to the seemingly unrelated cointegrating polynomial regression (SUCPR) setting considered in Wagner *et al.* (2020) and Knorre and Wagner (2024).[8] A difference to those papers is the consideration of cointegrating *multivariate* polynomial regressions, i. e., the inclusion of cross-product terms (and that the two papers mentioned both consider fully modified rather than integrated modified estimation). Given our setting, a system of seemingly unrelated multivariate cointegrating polynomial regressions would be of the form:

$$\begin{aligned} y_{ht} &= z_{ht}'\theta_h + u_{ht}, \quad h = 1, \ldots, n, \\ x_{ht} &= x_{h,t-1} + v_{ht}, \end{aligned} \tag{5}$$

where $z_{ht} := (z_{h1t}, \ldots, z_{h|\mathcal{I}_h|t})'$ with components of the form $z_{hit} = t^{i_0} x_{h1t}^{i_1} \cdots x_{hm_ht}^{i_{m_h}}$ for $h = 1, \ldots, n$ and $i = 1, \ldots, |\mathcal{I}_h|$. While in the "classical" seemingly unrelated regressions (SUR) setting of, e. g., Zellner (1962), no explicit assumptions are posited about the relationship between the regressors in the different equations, the seemingly unrelated cointegrating regression literature often assumes that the regressors are disjoint between the equations, with exceptions being Moon and Perron (2004) for the cointegrating linear case and Knorre and Wagner (2024) for the cointegrating polynomial case. As Knorre

[8] See also the discussion in Wagner (2023, Remark 2).

and Wagner (2024), we refer to *common* regressors as regressors occurring in at least two equations.[9]

The equation system (5) can be rewritten in the form given in (1): Combine all distinct elements of x_{ht}, $h = 1, \ldots, n$ into a vector $x_t \in \mathbb{R}^m$. In the case of no *common integrated regressors*, $x_t := (x'_{1t}, \ldots, x'_{nt})'$, with $m = \sum_{h=1}^n m_h$, and in the case that all integrated regressors are common – as in the example in Remark 1 $x_t := x_{1t} = \cdots = x_{nt}$, with $m = m_1 = \cdots = m_n$.[10] To define the joint regressor vector combine all distinct elements of $\{z_{hit}\}_{h=1,\ldots,n,i=1,\ldots,|\mathcal{I}_h|}$ into a vector Z_t with elements $z_{it} = t^{i_0} x_{1t}^{i_1} \cdots x_{mt}^{i_m}$ with a correspondingly defined multi-index set $\mathcal{I}$. Clearly, elements of the vector Z_t that occur as regressors only in one or some equations, imply *block-zero* restrictions in rows of Θ.

Consider as a simple example a system of two countries described by Cobb-Douglas production functions with again capital K_t and labor L_t as production factors and with a common technology level A_t, assumed for simplicity here as observable:[11]

[9]Note for completeness that, of course, Zellner (1962) also discusses the case where the regressors are assumed to be identical across all equations. For this setting he shows the important and widely-used result that OLS and GLS estimation coincide in the case that there are no restrictions on the parameters. This result is, as is also well known, e. g., the "basis" for equation-by-equation OLS estimation of unrestricted vector autoregressive models.

[10]One key difference between Wagner *et al.* (2020) and Knorre and Wagner (2024) is that the former paper excludes common regressors and the latter paper closes that gap. Both papers contain detailed discussions of *group-wise* pooling, i. e., of groups of coefficients being identical across groups of equations. This allows testing numerous specification-related hypotheses as well as performing correspondingly restricted estimation. A second difference between the two papers is that Knorre and Wagner (2024) provide a discussion for the general multiple integrated regressors case, whereas the discussion in Wagner *et al.* (2020) considers a more stylized and, thus, algebraically more accessible setting with only one integrated regressor and its powers per equation.
To complete the discussion, panel cointegrating (polynomial) regression settings are also closely related, see, e. g., de Jong and Wagner (2022) for pooled estimation or Wagner and Reichold (2023) for group-mean estimation. In this case, one usually assumes that the regressors in a setting like (5) are disjoint across equations (with some papers allowing for common regressors or factors in the cointegrating linear case) and that the coefficients are *pooled*, i. e., identical across equations. In line with the assumptions posited in classical panel analysis, the literature often assumes cross-sectional independence and potentially cross-sectional i.i.d. behavior (see, e. g., Phillips and Moon, 1999). Furthermore, it is customary, as in standard panel analysis, to include individual- and time-specific (fixed or random) effects.

[11]Commencing from $Y_t = \exp(\theta \ln A_t u_t) K_t^\alpha L_t^\beta$ taking logarithms leads, as is well known, to a *linear* cointegrating relationship, i. e., $\ln Y_t = \theta + \ln A_t + \alpha \ln K_t + \beta \ln L_t + u_t$, if one assumes that $\ln A_t, \ln K_t, \ln L_t$ are integrated (but not cointegrated) and that u_t is stationary.

$$\underbrace{\begin{bmatrix} \ln Y_{1t} \\ \ln Y_{2t} \end{bmatrix}}_{=y_t} = \underbrace{\begin{bmatrix} \theta_{1,1} & 1 & \alpha_1 & \beta_1 & 0 & 0 \\ \theta_{2,1} & 1 & 0 & 0 & \alpha_2 & \beta_2 \end{bmatrix}}_{=\Theta} \underbrace{\begin{bmatrix} 1 \\ \ln A_t \\ \ln K_{1t} \\ \ln L_{1t} \\ \ln K_{2t} \\ \ln L_{2t} \end{bmatrix}}_{=Z_t} + \underbrace{\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}}_{=u_t},$$

$$x_t = x_{t-1} + v_t,$$

with $x_t = (\ln A_t, \ln K_{1t}, \ln L_{1t}, \ln K_{2t}, \ln L_{2t})'$, i. e., $Z_t = (1, x_t')'$.[12] In the notation of (1), the elements of Z_t are of the form $z_{it} = t^{i_0}(\ln A_t)^{i_1}(\ln K_{1t})^{i_2} \cdots (\ln L_{2t})^{i_5}$. Since the Cobb-Douglas system corresponds to a system of (log-)linear cointegrating relationships, the multi-indices $\mathbf{i}$ are given by $(0, 0, \ldots, 0)$ (the intercept) or $i_0 = 0$ and exactly one $i_j = 1$ for $j = 1, \ldots, 5$ (the five integrated regressors), i. e., $\mathcal{I} = \{(0,0,0,0,0,0), (0,1,0,0,0,0), \ldots, (0,0,0,0,0,1)\}$ and $|\mathcal{I}| = 6$.[13]

The example illustrates several important aspects: As mentioned above, rewriting systems of seemingly unrelated cointegrating regressions into the systems format considered in this paper leads to block-zero restrictions in rows of the coefficient matrix Θ, in this case for the two equation-specific variables capital and labor. Common integrated regressors (across all equations), in this example $\ln A_t$, do not lead to row-wise zero restrictions. Since we assume in this example that $\ln A_t$ enters the production function one-to-one, the corresponding elements in Θ are restricted to be equal to one. However, one may want to estimate the corresponding coefficients, $\theta_{1,A}$ and $\theta_{2,A}$ say, to test whether the coefficients are identical (and/or equal to one). Other hypotheses that one may want to test in this example are constant returns to scale, i. e., $\alpha + \beta = 1$, either in one or both equations, or the same (in case of constant returns to scale) factor shares in both countries, i. e., $\alpha_1 = \alpha_2$.[14] Note that OLS- and GLS-type estimation do, in general,

[12] The ordering of the stochastic regressors in x_t chosen here, with first the common regressor and then the country-specific regressors ordered by equation is, of course, only one possibility. Knorre and Wagner (2024), e. g., order common regressor(s) last.

[13] Linking back to the standard notation defined above, e. g., α_1 corresponds to $\theta_{1,(0,0,1,0,0,0)}$.

[14] In the Translog case, testing for constant returns to scale involves several hypotheses. In the example of Remark 1 these are (omitting the equation index): $\theta_{(0,1,0)} + \theta_{(0,0,1)} = 1$, $\theta_{(0,2,0)} + \theta_{(0,0,2)} + \theta_{(0,1,1)} = 0$ and $\theta_{(0,2,0)} = \theta_{(0,0,2)}$.
One can also perform specification testing, using, e. g., a Translog system as a more general alternative to a Cobb-Douglas system and test the corresponding zero restrictions to the squared production factors and the cross-product of the two production factors. In fact, one interpretation of the Translog

not coincide for this system, if one imposes the restrictions on Θ for estimation. This exemplifies the need for considering both OLS- and GLS-type estimators.

Remark 3. In (1) we allow only for polynomial time trends, i. e., terms corresponding to multi-indices of the form $(i_0, 0, \ldots, 0)$ for $i_0 \geq 0$. However, more general deterministic components can, of course, be included, e. g., in a regression model of the form:

$$\begin{aligned} y_t &= \Theta_D D_t + \Theta Z_t + u_t, \\ x_t &= x_{t-1} + v_t, \end{aligned} \tag{6}$$

with Z_t containing only z_{it} with multi-indices where $\min_{j=1,\ldots,m} i_j > 0$. In this case, it suffices to assume for $D_t \in \mathbb{R}^p$ that there exists a sequence of $p \times p$ scaling matrices A_D and a p-dimensional vector of functions $D(z)$ such that for $0 \leq r \leq 1$ it holds that:

$$\lim_{T\to\infty} T^{1/2} A_D D_{\lfloor rT \rfloor} = D(r) \;\text{ with }\; 0 < \int_0^r D(z)D(z)' dz < \infty. \tag{7}$$

It is also possible to have elements of a more general D_t-vector included in the cross-product terms, provided asymptotic multi-collinearity is excluded.

2.2 Estimation and Inference

IM-OLS estimation is in fact nothing but OLS estimation of the partial sum version of equation (1) that is augmented by the original integrated regressors:

$$\begin{aligned} S_t^y &= \Theta S_t^Z + \Gamma x_t + S_t^u, \qquad t = 1, \ldots, T, \\ &= \Phi \tilde{S}_t^Z + S_t^u, \end{aligned} \tag{8}$$

with $S_t^y := \sum_{j=1}^t y_j$, S_t^Z, S_t^u defined analogously, $\tilde{S}_t^Z := (S_t^{Z\prime},\ x_t')'$ and $\Phi := (\Theta,\ \Gamma) \in \mathbb{R}^{n\times(|\mathcal{I}|+m)}$. Stacking all observations, equation (8) can be written as:

$$\begin{aligned} S^y &= \Theta S^Z + \Gamma X + S^u, \\ &= \Phi \tilde{S}^Z + S^u, \end{aligned} \tag{9}$$

function is to consider it as a second-order Taylor approximation to an unknown more general production function, see, e. g., Christensen *et al.* (1973) or Denny and Fuss (1977). This interpretation, of course, leads to the potential consideration of higher-order Taylor approximations.

with $S^y := (S_1^y, \ldots, S_T^y)$, $S^Z := (S_1^Z, \ldots, S_T^Z)$, $X := (x_1, \ldots, x_T)$, $S^u := (S_1^u, \ldots, S_T^u)$ and $\tilde{S}^Z := (\tilde{S}_1^Z, \ldots, \tilde{S}_T^Z)$. Exactly as discussed in a closely related context in Wagner (2023, Remark 1) and in fact known since Zellner (1962), see Footnote 9, for unrestricted systems of equations (that are linear in parameters) with identical regressors in all equations, OLS estimation coincides (algebraically) with GLS estimation for any (regular) weighting matrix. Consequently, with identical regressors in all equations and without parameter restrictions, it suffices to consider the system version of the *single equation* IM-OLS estimator for CMPRs discussed in Vogelsang and Wagner (2024).[15] The IM-OLS estimator $\hat{\Phi}$ is defined as the OLS estimator of Φ in (9), i. e.,:

$$\hat{\Phi} := (S^y \tilde{S}^{Z\prime})(\tilde{S}^Z \tilde{S}^{Z\prime})^{-1}. \tag{10}$$

The discussion of the asymptotic properties of the IM-OLS estimator requires the definition of two quantities: First, the scaling matrix sequence $A_{\text{IM}} := \text{diag}(A_{\text{IM},\Theta}, I_m)$ with $A_{\text{IM},\Theta}$ a diagonal matrix with the entry corresponding to regressor $t^{i_0} x_{1t}^{i_1} \cdots x_{mt}^{i_m}$ given by $T^{-(i_0 + (\sum_{j=1}^m i_j)/2 + 1/2)}$. Second, the properly scaled partial sum process corresponding to the Second, the limit process corresponding to the regressors Z_t, i. e., $Z(r) := \lim_{T\to\infty} T^{1/2} A_{\text{IM},\Theta} Z_{\lfloor rT \rfloor}$ for $0 \leq r \leq 1$, with $Z(r) := (z_1(r), \ldots, z_{|\mathcal{I}|}(r))'$, $z_i(r) := r^{i_0} B_{v_1}(r)^{i_1} \cdots B_{v_m}(r)^{i_m}$ for $0 \leq r \leq 1$, $i = 1, \ldots, |\mathcal{I}|$ and $B_{v_j}(r)$ denoting the j-th component of $B_v(r)$.

Proposition 1. *Let the data be generated by* (1) *with appropriate assumptions in place. Define* $\Phi^* := (\Theta, \Omega_{uv}\Omega_{vv}^{-1})$, *then as* $T \to \infty$ *it holds that:*[16]

$$\begin{aligned}(\hat{\Phi} - \Phi^*)A_{IM}^{-1} &\Rightarrow \Omega_{u\cdot v}^{1/2} \int_0^1 W_{u\cdot v}(s) f(s)' ds \left(\int_0^1 f(s)f(s)'ds\right)^{-1} \\ &= \Omega_{u\cdot v}^{1/2} \int_0^1 dW_{u\cdot v}(s)[F(1) - F(s)]' \left(\int_0^1 f(s)f(s)'ds\right)^{-1},\end{aligned} \tag{11}$$

where:

$$f(r) := \begin{bmatrix} \int_0^r Z(s)ds \\ B_v(r) \end{bmatrix}, \quad F(r) := \int_0^r f(s)ds. \tag{12}$$

[15]Later, when discussing hypothesis testing and estimation under restrictions, it is convenient to consider vectorized version(s) of (9), either vectorized by *observation*, i. e., $\text{vec}(S^y) = (\tilde{S}^{Z\prime} \otimes I_n)\text{vec}(\Phi) + \text{vec}(S^u)$ or vectorized by *equation*, i. e., $\text{vec}(S^{y\prime}) = (I_n \otimes \tilde{S}^{Z\prime})\text{vec}(\Phi') + \text{vec}(S^{u\prime})$.

[16]To detail notation: The (i,j)-element of $\int_0^1 dW_{u\cdot v}(s)[F(1) - F(s)]'$ is given by $\int_0^1 [F_j(1) - F_j(s)]dW_{u\cdot v,i}(s)$.

As indicated in Footnote 15, for hypothesis testing and estimation under restrictions, it is convenient to consider the vectorized (by equation) version of the IM-OLS estimator $\hat{\Phi}$ defined in (10). Defining $\phi := \text{vec}(\Phi')$ and $\phi^* := \text{vec}(\Phi^{*\prime})$, this leads to:

$$\hat{\phi} := \text{vec}\left((\tilde{S}^Z \tilde{S}^{Z\prime})^{-1}(\tilde{S}^Z S^{y\prime})\right) = (I_n \otimes (\tilde{S}^Z \tilde{S}^{Z\prime})^{-1})(I_n \otimes \tilde{S}^Z)\text{vec}(S^{y\prime}) \tag{13}$$

and:

$$\begin{aligned} &(I_n \otimes A_{\text{IM}}^{-1})\left(\hat{\phi} - \phi^*\right) \\ &\Rightarrow (\Omega_{u\cdot v}^{1/2} \otimes I_{|\mathcal{I}|+m})\text{vec}\left(\left(\int_0^1 f(s)f(s)'ds\right)^{-1} \int_0^1 [F(1) - F(s)]dW_{u\cdot v}(s)'\right). \end{aligned} \tag{14}$$

Conditional upon $W_v(r)$, the limiting distribution given in (14) is normal with zero mean and (conditional) covariance matrix:

$$\begin{aligned} V_{\text{IM}} := \Omega_{u\cdot v} \otimes &\left(\left(\int_0^1 f(s)f(s)'ds\right)^{-1}\right. \\ &\times \left.\left(\int_0^1 \big[F(1) - F(s)\big]\big[F(1) - F(s)\big]'ds\right)\left(\int_0^1 f(s)f(s)'ds\right)^{-1}\right). \end{aligned} \tag{15}$$

Given a consistent estimator $\hat{\Omega}_{u\cdot v}$ of $\Omega_{u\cdot v}$, based on $\hat{\eta}_t := (\hat{u}_t', v_t')'$, with $\hat{u}_t$ the OLS residuals of (1), an – up to scaling – estimator of V_{IM} immediately follows by simply using the sample counterparts of the expressions appearing in the limit given in (15), i. e.,:

$$\hat{V}_{\text{IM}} := \hat{\Omega}_{u\cdot v} \otimes (\tilde{S}^Z \tilde{S}^{Z\prime})^{-1} CC' (\tilde{S}^Z \tilde{S}^{Z\prime})^{-1}, \tag{16}$$

with $C := (c_1, \ldots, c_T)$, $c_t := S_T^{\tilde{S}^Z} - S_{t-1}^{\tilde{S}^Z}$ for $t = 1, \ldots, T$, $S_t^{\tilde{S}^Z} := \sum_{j=1}^t \tilde{S}_j^Z$ and $S_0^{\tilde{S}^Z} = 0$. By construction, when $\hat{\Omega}_{u\cdot v} \to \Omega_{u\cdot v}$ in probability, which holds under standard assumptions on kernels and bandwidths (see, e. g., Jansson, 2002), it follows that $(I_n \otimes A_{\text{IM}}^{-1})\hat{V}_{\text{IM}}(I_n \otimes A_{\text{IM}}^{-1}) \Rightarrow V_{\text{IM}}$.

The limiting distribution given in (14), in conjunction with the estimator $\hat{V}_{\text{IM}}$ given in (16), allows for asymptotic standard inference for testing (linear) restrictions on ϕ under two assumptions on the restrictions matrix, R say, that are detailed (for the single-equation case) in Vogelsang and Wagner (2024, Section 2.2): The first relates to the fact that the parameter vector $\hat{\phi}$ contains elements that converge at different rates, which has some implications for hypotheses that lead to standard inference (encoded in the matrix A_R below). The second assumption on R, not explicitly stated in the proposition, is that none of the hypotheses tested involves elements of Γ, which is not estimated consistently.[17]

Proposition 2. *Let the data be generated by* (1) *with appropriate assumptions in place and assume that long-run covariance estimation is performed consistently. Consider s linearly independent linear restrictions collected in:*

$$H_0: \ Rvec(\Phi') = R\phi = r, \tag{17}$$

with $R \in \mathbb{R}^{s\times(|\mathcal{I}|+m)n}$ of full row rank, $r \in \mathbb{R}^s$ and suppose that there exists a matrix sequence $A_R \in \mathbb{R}^{s\times s}$ such that:

$$\lim_{T\to\infty} A_R^{-1} R(I_n \otimes A_{IM}) = R^*, \tag{18}$$

with $R^* \in \mathbb{R}^{s\times(|\mathcal{I}|+m)n}$ of full row rank s. Then, it holds under the null hypothesis for $T \to \infty$ that the Wald-type statistic:

$$\tau_W := (R\hat{\phi} - r)' \left(R\hat{V}_{IM}R'\right)^{-1} (R\hat{\phi} - r) \Rightarrow \mathcal{O}_s, \tag{19}$$

with $\hat{V}_{IM}$ as defined in (16) *and $\mathcal{O}_s$ denoting a chi-squared distributed random variable with s degrees of freedom.*

In the case that $s = 1$, it holds under the null hypothesis for $T \to \infty$ that the t-type statistic:

$$\tau_t := \frac{R\hat{\phi} - r}{\sqrt{R\hat{V}_{IM}R'}} \Rightarrow \mathcal{Z}, \tag{20}$$

with $\mathcal{Z}$ denoting a univariate standard normally distributed random variable.

2.3 Estimation and Inference Under Restrictions

As discussed in Wagner (2023, Section 2.3), the cointegrating regression literature rarely considers restricted least squares estimation, with one exception being the seemingly

[17] More formally, with K denoting the so-called commutation matrix, this means that for $\phi = \text{vec}(\Phi') = K\text{vec}(\Phi) = K(\text{vec}(\Theta)', \text{vec}(\Gamma)')'$ it has to hold that $R\phi = RK(\text{vec}(\Theta)', \text{vec}(\Gamma)')'$ is of the form $RK = (R_{\text{vec}(\Theta)}, 0_{s\times nm})$.

unrelated regressions (SUR) cointegration literature, see, e. g., Moon (1999), Moon and Perron (2004), Park and Ogaki (1991) or Wagner *et al.* (2020). When not all equations include the same set of regressors, OLS- and GLS-type estimation, in general, cease to be algebraically (and asymptotically) equivalent.[18] Potential choices concerning weighting matrices in seemingly unrelated cointegrating regression systems are discussed in Park and Ogaki (1991), see also Wagner (2023). IM-GLS estimation adds one additional formal aspect to the discussion: The errors in the partial sum regression are integrated and, therefore, weighting matrices cannot be directly related to covariance or long-run covariance matrices of the error process, but rather to the first differences of the errors, motivating the Park and Ogaki (1991) choices $W = \Omega_{uu}^{-1}$ or $W = \Omega_{u \cdot v}^{-1}$ also in the IM setting.[19] Clearly, restricted IM-OLS estimation is contained as the special case with $\hat{W} = W = I_n$.[20]

To obtain a closed-form solution for the restricted estimator we consider, analogously to hypothesis testing above, only affine restrictions on the parameter vector, i. e.,:

$$\phi = \mathrm{vec}(\Phi') = D\varphi + d, \tag{21}$$

with $D \in \mathbb{R}^{(|\mathcal{I}|+m)n \times g}$ of full column rank, $\varphi \in \mathbb{R}^g$, arranged by equation similarly to ϕ, and $d \in \mathbb{R}^{(|\mathcal{I}|+m)n}$.[21] Given the mentioned fact that only the parameters in Θ are estimated consistently, we only consider restrictions on Θ and do not consider restrictions involving elements of Γ. As above, we need once again to posit an asymptotic rank condition on the constraint matrix, i. e., we need to assume that there exists a matrix sequence $A_D \in \mathbb{R}^{g \times g}$ such that:

$$\lim_{T \to \infty} (I_n \otimes A_{\mathrm{IM}}^{-1}) D A_D = D^*, \tag{22}$$

with $D^* \in \mathbb{R}^{(|\mathcal{I}|+m)n \times g}$ of full column rank.

[18] We refer to GLS estimation for any variant of weighted least squares estimation and not – as in, e. g., the classical Zellner (1962) setting – when weighting takes place with the inverse of the error covariance matrix.

[19] We use the notation W for weighting matrices for obvious reasons and are confident that no confusion with Wiener processes, always written with argument, i. e., as $W(r)$, will occur.

[20] To be precise, the $nT \times nT$ weighting matrices considered in these cases are $\Omega_{uu}^{-1} \otimes I_T$, $\Omega_{u \cdot v}^{-1} \otimes I_T$ and $I_n \otimes I_T$, respectively. Note that all GLS results presented in this paper consider weighting matrices of the form $\hat{W} \otimes I_T$. From an algebraic perspective, one could, in principle, consider also "full" $\hat{W} \in \mathbb{R}^{nT \times nT}$ weighting matrices.

[21] As is well known, the explicit formulation of restrictions used in (21) is equivalent to the implicit formulation $R\phi = r$ used in the discussion of the Wald-type test. Starting from the explicit formulation, denote with $D_\perp \in \mathbb{R}^{(|\mathcal{I}|+m)n \times ((|\mathcal{I}|+m)n - g)}$ a matrix of full column rank that fulfills $D_\perp' D = 0$. Then $R = D_\perp'$, $r = D_\perp' d$ and $s = (|\mathcal{I}| + m)n - g$.

Proposition 3. *Let the data be generated by* (1) *with appropriate assumptions in place and* ϕ *fulfilling the (explicit) restrictions posited in* (21). *Furthermore, assume that there exists a matrix sequence* A_D *such that condition* (22) *holds. The* restricted integrated modified generalized least squares (IM-GLS) estimator $\hat{\phi}_R$ *of* ϕ *with symmetric weighting matrix sequence* $\hat{W}$ *is defined as:*

$$\hat{\phi}_R := D\hat{\varphi} + d, \tag{23}$$

with:

$$\hat{\varphi} := \left((D'(\hat{W} \otimes \tilde{S}^Z \tilde{S}^{Z\prime})D\right)^{-1} \times \left(D'\left(vec\left(\tilde{S}^Z S^{y\prime}\hat{W}\right) - (\hat{W} \otimes \tilde{S}^Z \tilde{S}^{Z\prime})d\right)\right). \tag{24}$$

For φ^* *such that* $\phi^* = D\varphi^* + d$*, it holds for* $T \to \infty$ *and* $\hat{W} \to W > 0$ *that:*

$$A_D^{-1}(\hat{\varphi} - \varphi^*) \Rightarrow \left(D^{*\prime}\left(W \otimes \int_0^1 f(s)f(s)'ds\right)D^*\right)^{-1} \times \left(D^{*\prime}\text{vec}\left(\int_0^1 \big[F(1) - F(s)\big]dB_{u\cdot v}(s)'W\right)\right). \tag{25}$$

The limiting distribution of $\hat{\varphi}$ *given in* (25) *is – conditional upon* $W_v(r)$ *– normal with zero mean and covariance matrix:*

$$V_{IM,R} := A^{-1}BA^{-1}, \tag{26}$$

with:

$$A := D^{*\prime}\left(W \otimes \int_0^1 f(s)f(s)'ds\right)D^*, \tag{27}$$

$$B := D^{*\prime}\left(W\Omega_{u\cdot v}W \otimes \int_0^1 \big[F(1) - F(s)\big]\big[F(1) - F(s)\big]'ds\right)D^*. \tag{28}$$

An estimator of $V_{\mathrm{IM,R}}$ is readily available, analogously to (16) and, therefore, asymptotically chi-squared or standard normal inference on φ follows, under conditions (22) and (30), similarly to Proposition 2:[22]

[22]Clearly, as already discussed below (21), since the elements of φ that correspond to elements of Γ are not estimated consistently, the hypotheses are only allowed to involve entries of φ corresponding to elements of Θ. Furthermore, note that the limiting distribution of $\hat{\phi}_{\mathrm{R}}$ is, by construction, singular unless D is a regular matrix.

Proposition 4. *Let the data be generated by* (1) *with appropriate assumptions in place and assume that long-run covariance estimation is performed consistently. Let the parameter vector* φ *in* $\phi = D\varphi + d$ *with condition* (22) *in place fulfill* s_φ *linearly independent restrictions, i. e.,:*

$$H_0 : R_\varphi \varphi = r_\varphi, \tag{29}$$

with $R_\varphi \in \mathbb{R}^{s_\varphi \times g}$ *with full row rank* s_φ *and* $r_\varphi \in \mathbb{R}^{s_\varphi}$*. Furthermore, assume that there exists a matrix sequence* $A_\varphi \in \mathbb{R}^{s_\varphi \times s_\varphi}$ *such that:*

$$\lim_{T\to\infty} A_\varphi^{-1} R_\varphi A_D = R_\varphi^* \tag{30}$$

exists and has full row rank s_φ *and that* $\hat{W} \to W > 0$ *in probability. Then, it holds under the null hypothesis for* $T \to \infty$ *that the Wald-type statistic:*

$$\tau_{W,R} := (R_\varphi \hat{\varphi} - r_\varphi)' \left(R_\varphi \hat{A}^{-1} \hat{B} \hat{A}^{-1} R_\varphi' \right)^{-1} (R_\varphi \hat{\varphi} - r_\varphi) \Rightarrow \mathcal{O}_{s_\varphi}, \tag{31}$$

with $\mathcal{O}_{s_\varphi}$ *denoting a chi-squared distributed random variable with* s_φ *degrees of freedom and:*

$$\hat{A} := D'(\hat{W} \otimes \tilde{S}^Z \tilde{S}^{Z\prime})D, \tag{32}$$

$$\hat{B} := D'(\hat{W}\hat{\Omega}_{u\cdot v}\hat{W} \otimes CC')D. \tag{33}$$

In the case that $s_\varphi = 1$*, a t-type statistic that is asymptotically standard normally distributed can be defined analogously to Proposition 2.*

Remark 4. To illustrate matters, let us consider the Cobb-Douglas system from Remark 2 again. For this example, we have:

$$\Phi = \left[\begin{array}{cccccc|ccc} \theta_{1,1} & 1 & \alpha_1 & \beta_1 & 0 & 0 & \gamma_{1,1} & \cdots & \gamma_{1,5} \\ \theta_{2,1} & 1 & 0 & 0 & \alpha_2 & \beta_2 & \gamma_{2,1} & \cdots & \gamma_{2,5} \end{array}\right] \in \mathbb{R}^{2\times 11},$$

with $\gamma_{h,j}$, $h = 1, 2(= n)$ and $j = 1, \ldots, 5(= m)$, denoting the coefficients in Γ. Since there are 16 free parameters in Φ, using the notation as defined in (21), this implies that $D \in \mathbb{R}^{22\times 16}$, $\varphi \in \mathbb{R}^{16}$ and $d \in \mathbb{R}^{22}$. More specifically, $\varphi = (\theta_{1,1}, \alpha_1, \beta_1, \gamma_{1,1}, \ldots, \gamma_{1,5}, \theta_{2,1}, \alpha_2, \beta_2, \gamma_{2,1}, \ldots, \gamma_{2,5})'$ and $d = (0, 1, 0, \ldots, 0, 1, 0, \ldots, 0)'$ with the two 1-entries in the second and thirteenth coordinate. The rows of the matrix D are all of the form (i) one 1-entry, relating an element of φ, i. e., one of the unknown parameters, to an unrestricted coordinate of ϕ, and all other entries equal to zero, or (ii) zeros only, corresponding to coordinates in ϕ that are fixed (either at zero or one in this example).[23]

Next, consider restrictions on the elements of φ. The joint null hypothesis of constant returns to scale in both countries and equal factor shares corresponds to $H_0 : \alpha_1 + \beta_1 = 1, \alpha_2 + \beta_2 = 1, \alpha_1 = \alpha_2$:[24]

$$R_\varphi \varphi = \begin{bmatrix} 0 & 1 & 1 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 & -1 & 0 & 0 & \dots & 0 \end{bmatrix} \varphi = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = r_\varphi.$$

For this example, condition (22) is satisfied with $A_{\text{IM}} = \text{diag}(T^{-1/2}, T^{-1}I_5, I_5)$, $A_D = I_2 \otimes \text{diag}(T^{-1/2}, T^{-1}I_2, I_5)$ and $D = D^*$. That $D = D^*$ stems directly from the fact that in each of the hypotheses considered only coefficients with the same convergence rate are involved. This is a "standard testing" situation in which the rank constraint (22) is satisfied trivially. The same observation holds true also for testing constant returns to scale of Translog production functions (see Footnote 14).

Remark 5. Note that, exactly as discussed in Wagner (2023, Remark 1), for restrictions of the form $D = I_n \otimes \mathcal{D}$, with an asymptotic rank condition of the form (22) holding for a full rank limiting matrix $D^* = I_n \otimes \mathcal{D}^*$, the IM-GLS estimator coincides with the IM-OLS estimator for any positive definite symmetric weighting matrix $\hat{W}$.

2.4 Fixed-b Inference

As mentioned in the introduction, one advantage of the IM-OLS estimator introduced for single-equation cointegrating linear regressions in Vogelsang and Wagner (2014) and extended to the single-equation CMPR setting in Vogelsang and Wagner (2024) is that it can be used for asymptotically pivotal fixed-b inference. In the CMPR setting, see Vogelsang and Wagner (2024, Corollary 1 and Proposition 3), asymptotically pivotal fixed-b inference requires *full design* of the regression model. Full design means that the limit process $Z(r)$ can be written as $Z(r) = \Pi_Z Z_W(r)$, with Π_Z a regular matrix and $Z_W(r)$ a functional of standard Brownian motions.[25]

[23] This occurs in coordinates (and rows of D) two, five, six and 13 to 15 of ϕ.

[24] Equivalently, one can also take $\beta_1 = \beta_2$ as third linearly independent restriction.

[25] Full design of (S)CMPRs can always be achieved by adding regressors, see the (single-equation) discussion in Vogelsang and Wagner (2024).

The *system* CMPR setting considered in this paper adds another complexity to asymptotically pivotal fixed-b inference: The key quantity in fixed-b inference is a modified estimator $\hat{\Omega}_{u\cdot v,\mathrm{M}}$ of $\Omega_{u\cdot v}$ constructed from *modified* residuals $\hat{S}^u_{t,\mathrm{M}}$ as defined below. In the system case considered, this long-run covariance matrix is now, obviously, an $n \times n$ matrix rather than, as in Vogelsang and Wagner (2014, 2024), a scalar. With respect to V_{IM}, this implies that (using a lower case letter for a scalar quantity) the *variance scaling factor* in the test statistic is not of the form $\omega_{u\cdot v}$ times a matrix but, see (15), given by the Kronecker product of $\Omega_{u\cdot v}$ and a matrix, $\mathcal{M}$ say. This implies, see the proof of Proposition 5, that a sufficient condition for asymptotically pivotal fixed-b inference is that the restrictions matrix $R \in \mathbb{R}^{s\times(|\mathcal{I}|+m)n}$ – in $R\phi = r$ – fulfills $R = I_n \otimes \mathcal{R}$, with $\mathcal{R} \in \mathbb{R}^{s/n\times(|\mathcal{I}|+m)}$, with $s_{\mathcal{R}} := \frac{s}{n}$ a (positive) integer. This allows to perform fixed-b inference for testing up-to-the-intercept-identical hypotheses in all equations in systems with identical regressors in all equations. Whilst this is clearly restrictive, it, e. g., includes RESET-type specification testing for SUCPR systems with identical regressors in all equations in both the null specification and the augmented regression.

Proposition 5. *Let the data be generated by* (1) *and assume that full design prevails. Consider $s = s_{\mathcal{R}}n$ linearly independent restrictions collected in:*[26]

$$H_0 : Rvec(\Phi') = (I_n \otimes \mathcal{R})\phi = r, \tag{34}$$

with $\mathcal{R} \in \mathbb{R}^{s_{\mathcal{R}}\times(|\mathcal{I}|+m)}$ of full row rank $s_{\mathcal{R}}$, $r \in \mathbb{R}^{s_{\mathcal{R}}n}$ and suppose that there exists a matrix sequence $A_{\mathcal{R}} \in \mathbb{R}^{s_{\mathcal{R}}\times s_{\mathcal{R}}}$ such that:

$$\lim_{T\to\infty} A_{\mathcal{R}}^{-1}\mathcal{R}A_{IM} = \mathcal{R}^*, \tag{35}$$

with $\mathcal{R}^* \in \mathbb{R}^{s_{\mathcal{R}}\times(|\mathcal{I}|+m)}$ of full row rank $s_{\mathcal{R}}$. Then, it holds under the null hypothesis for $T \to \infty$ that the fixed-b Wald-type statistic:

$$\tau_{W,b} := (R\hat{\phi} - r)'\left(R\hat{V}_{IM,M}R'\right)^{-1}(R\hat{\phi} - r) \Rightarrow \mathcal{Z}'_{s_{\mathcal{R}}n}(Q(P)^{-1} \otimes I_{s_{\mathcal{R}}})\mathcal{Z}_{s_{\mathcal{R}}n}, \tag{36}$$

with $\hat{V}_{IM,M}$ defined similarly as $\hat{V}_{IM}$ in (16)*, but with $\hat{\Omega}_{u\cdot v}$ replaced by $\hat{\Omega}_{u\cdot v,M}$, defined in* (37) *below, and $\mathcal{Z}_{s_{\mathcal{R}}n}$ an $s_{\mathcal{R}}n$-dimensional standard normally distributed random vec-*

[26] As in Proposition 2, we assume that none of the hypotheses tested involves elements of Γ, compare Footnote 17. This requires the last m columns of $\mathcal{R}$ to be zero.

tor independent of $Q(P)$. The precise form of $Q(P)$ depends on the specification of the SCMPR (1)*, the kernel function $k(\cdot)$ and the bandwidth-to-sample-size ratio* $0 < b \leq 1$.[27]

It is key for asymptotically pivotal fixed-b inference that $\mathcal{Z}_{s_{\mathcal{R}}n}$ and $Q(P)$ in (36) are independent random variables. The necessity to achieve independence implies (for exactly the same reason as discussed in detail in Vogelsang and Wagner, 2014, 2024) that, as indicated above, $\Omega_{u\cdot v}$ cannot be estimated using the IM-OLS residuals $\hat{S}_t^u := S_t^y - \hat{\Phi}\tilde{S}_t^Z$ and $\hat{S}^u := (\hat{S}_1^u, \ldots, \hat{S}_T^u)$. Instead, *orthogonalized* modified residuals, $\hat{S}_{t,\mathrm{M}}^u$, have to be used to annihilate (nuisance-parameter-dependent) correlation. These are given by $\hat{S}_{\mathrm{M}}^u := \hat{S}_u \left(I_T - M^{\perp\prime}(M^{\perp}M^{\perp\prime})^{-1}M^{\perp}\right)$, with $M^{\perp} := M(I_T - \tilde{S}^{Z\prime}(\tilde{S}^Z\tilde{S}^{Z\prime})^{-1}\tilde{S}^Z)$, $M := (M_1, \ldots, M_T)$ and $M_t := t\sum_{j=1}^T \tilde{S}_j^Z - \sum_{j=1}^{t-1}\sum_{s=1}^j \tilde{S}_s^Z$ for $t = 1, \ldots, T$. The required modified estimator of $\Omega_{u\cdot v}$ is now defined as:

$$\begin{aligned}
\hat{\Omega}_{u\cdot v,\mathrm{M}} &:= T^{-1}\sum_{i=2}^T\sum_{j=2}^T k\left(\frac{|i-j|}{B}\right)\Delta\hat{S}_{i,\mathrm{M}}^u \Delta\hat{S}_{j,\mathrm{M}}^{u\prime}, \\
&\Rightarrow \Omega_{u\cdot v}^{1/2} Q(P) \Omega_{u\cdot v}^{1/2\prime},
\end{aligned} \tag{37}$$

with kernel function $k(\cdot)$ and bandwidth $B = bT$ for some $0 < b \leq 1$.

Remark 6. Let us illustrate the scope of fixed-b inference with the Translog example discussed in Remark 1 and the Cobb-Douglas example in Remarks 2 and 4. The two-outputs Translog example is an unrestricted system, i. e., a system with identical regressors with unrestricted coefficients in both equations. It is therefore possible to perform fixed-b inference for testing whether the system is in fact Cobb-Douglas, corresponding to the coefficients to the squared production factors and the cross product being zero in both equations, or whether both outputs are produced with constant returns to scale (with the corresponding hypotheses given in Footnote 14). For the Cobb-Douglas example fixed-b inference is not available since the starting point is, due to country-specific production factors, a restricted system of equations (with block-zero restrictions in the parameter matrix Θ.

[27]Given the comparably limited scope for fixed-b inference in the SCMPR setting, we abstain from explicitly stating and defining all necessary quantities. The stochastic process $P(r)$ is the multivariate analogue of $P(r)$ as defined in Vogelsang and Wagner (2024, Proposition 3). The key difference is that $W_{u\cdot v}(r)$ is now an n-dimensional rather than a scalar process. The other elements constituting $P(r)$ – $g(r)$, $G(r)$, $h(r)$ and $H(r)$ – are exactly as in Vogelsang and Wagner (2024, Corollary 1 and Proposition 3). Furthermore, the form of the functional(s) $Q(P)$ is exactly as given above Vogelsang and Wagner (2024, Proposition 3), conveniently defined there already for the multivariate case.

Remark 7. Note that if the restrictions considered in this subsection are not rejected, the discussion in Footnote 21 clarifies that the corresponding restricted estimation problem is, unsurprisingly, subject to the type of restrictions discussed in Remark 5. This is a situation in which IM-GLS coincides with IM-OLS, or, in other words, the fixed-b discussion in this paper is (algebraically) confined to IM-OLS.

3 Summary and Conclusions

This paper has extended integrated modified least squares estimation theory from the single equation cointegrating multivariate polynomial regression setting considered in Vogelsang and Wagner (2024) to the systems case. It is important to note that the CMPR and SCMPR setting allow to overcome the ubiquitous additive separability assumption between integrated regressors in the nonlinear cointegrating regression literature, admittedly only for polynomial functions. Whilst this is, of course, a very special class of functions it (i) includes the relevant case of Translog-type (systems of) equations, (ii) allows for RESET-type specification testing and (iii) is an important starting point for multivariate Sieve-type (or Taylor) approximation of more general (differentiable) functions.

Considering systems of equations adds two aspects compared to single equation analysis. First, it necessitates the detailed consideration of not only ordinary least squares but also generalized least squares estimators. This stems from the well-known fact that OLS and GLS estimators only necessarily coincide in equation systems with identical regressors across all equations and without parameter constraints (as shown in Zellner, 1962). Whilst such an unrestricted system might be a valid starting point for empirical analysis, the result of, e. g., specification testing or the imposition of restrictions stemming from economic theory will typically be restricted systems. Consequently, we discuss in detail inference for both unrestricted and restricted IM estimators. It furthermore turns out that the scope of fixed-b inference is relatively limited. In addition to *full design* of the regression system – a necessary condition also in the single-equation setting in Vogelsang and Wagner (2024) – it turns out that fixed-b inference is only available for up-to-the-intercept-identical hypotheses tested in all equations in systems with identical regressors in all equations. Despite this being very restrictive, it allows, e. g., for RESET-type specification testing for SUCPR systems with identical regressors in all equations in both the null specification and the augmented regression.

The methods developed in this paper are going to be used in future work to estimate Translog-type production function systems, both at the aggregate (multi-)national level as well as at the sectoral level.[28] In addition, it may also be interesting to extend the SUR-type theory of Knorre and Wagner (2024) from the SCPR to the SCMPR case considered in this paper.

Acknowledgements

The authors gratefully acknowledge financial support from the Jubiläumsfonds of the Oesterreichische Nationalbank (Grant no. 18692). This paper was written whilst the second author gratefully enjoyed the hospitality of the Robert Schuman Centre for Advanced Studies of the European University Institute during a Simone Veil Fellowship in the first half of 2024. The views expressed in this paper are solely ours and not necessarily those of the Bank of Slovenia or the European System of Central Banks. On top of this the usual disclaimer applies.

References

Christensen, L.R., D.W. Jorgenson and L.J. Lau (1971). Conjugate Duality and the Transcendental Logarithmic Production Function. *Econometrica* **39**, 255–256.

Christensen, L.R., D.W. Jorgenson and L.J. Lau (1973). Transcendental Logarithmic Production Frontiers. *Review of Economics and Statistics* **55**, 28–45.

de Jong, R.M. and M. Wagner (2022). Panel Cointegrating Polynomial Regression Analysis and an Illustration with the Environmental Kuznets Curve. *Econometrics and Statistics*, in press: `https://doi.org/10.1016/j.ecosta.2022.03.005`.

Denny, M. and M. Fuss (1977). The Use of Approximation Analysis to Test for Separability and the Existence of Consistent Aggregates. *American Economic Review* **67**, 404–418.

[28] In the context of such an empirical analysis we will also delve into the finite sample performance of the developed methods more systematically. Cursory simulation evidence suggests that in terms of performance effectively the "usual kind of results" can be expected (compare, e. g., Wagner and Hong, 2016; Wagner *et al.*, 2020; Knorre and Wagner, 2024; Vogelsang and Wagner, 2024), also with respect to the impact of the number of equations in relation to the sample size.

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

Kiefer, N.M. and T.J. Vogelsang (2005). A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests. *Econometric Theory* **21**, 1130–1164.

Knorre, F. and M. Wagner (2024). Fully Modified OLS Estimation and Inference for Seemingly Unrelated Cointegrating Polynomial Regressions with Common Integrated Regressors. *Mimeo.*

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

Moon, H.R. and B. Perron (2004). Efficient Estimation of the SUR Cointegration Regression Model and Testing for Purchasing Power Parity. *Econometric Reviews* **23**, 293–323.

Neave, H.R. (1970). An Improved Formula for the Asymptotic Variance of Spectrum Estimates. *Annals of Mathematical Statistics* **41**, 70–77.

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

Phillips, P.C.B. and H.R. Moon (1999). Linear Regression Limit Theory for Nonstationary Panels. *Econometrica* **67**, 1057–1111.

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

Thursby, J.G. and P. Schmidt (1977). Some Properties of Tests for Specification Error in a Linear Regression Model. *Journal of the American Statistical Association* **72**, 635–641.

Vogelsang, T.J. and M. Wagner (2014). Integrated Modified OLS Estimation and Fixed-b Inference for Cointegrated Regressions. *Journal of Econometrics* **178**, 741–760.

Vogelsang, T.J. and M. Wagner (2024). Integrated Modified OLS Estimation and Fixed-b Inference for Cointegrating Multivariate Polynomial Regressions. *IHS Working Paper Series 53.*

Wagner, M. (2023). Fully Modified Least Squares Estimation and Inference for Systems of Cointegrating Polynomial Regressions. *Economics Letters* **228**, 111186.

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

Wagner, M. and K. Reichold (2023). Panel Cointegrating Polynomial Regressions: Group-Mean Fully Modified OLS Estimation and Inference. *Econometric Reviews* **42**, 358–392.

Wagner, M., P. Grabarczyk and S.H. Hong (2020). Fully Modified OLS Estimation and Inference for Seemingly Unrelated Cointegrating Polynomial Regressions and the Environmental Kuznets Curve for Carbon Dioxide Emissions. *Journal of Econometrics* **214**, 216–255.

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: Proofs

Proof of Proposition 1. The result presents the system version of the IM-OLS estimator and its asymptotic properties derived for the single-equation CMPR setting with $n = 1$ in Vogelsang and Wagner (2024, Proposition 1) and follows upon combining the individual equation results. ■

Proof of Proposition 2. Under the null hypothesis and condition (18) on A_R, it holds that:

$$A_R^{-1}(R\hat{\phi} - r) = \left(A_R^{-1}R(I_n \otimes A_{\mathrm{IM}})\right)\Big((I_n \otimes A_{\mathrm{IM}}^{-1})(\hat{\phi} - \phi^*)\Big) \Rightarrow R^*\mathcal{Y},$$

with $\mathcal{Y}$ denoting the random variable (limiting distribution) given in (14). $R^*\mathcal{Y}$ is under the null hypothesis – conditional upon $W_v(r)$ – normally distributed with zero mean and covariance matrix $R^*V_{\mathrm{IM}}R^{*\prime}$. Under condition (18), it furthermore holds that $A_R^{-1}R\hat{V}_{\mathrm{IM}}R'A_R^{-1\prime} \Rightarrow R^*V_{\mathrm{IM}}R^{*\prime}$. Combining the two results now immediately leads to the asymptotic chi-squared distribution for τ_{W} as defined in (19) by noting that conditional convergence to a chi-squared distribution that is (by definition) independent of $W_v(r)$ amounts to unconditional convergence. ■

Proof of Proposition 3. Centering of the IM-OLS estimator, compare Proposition 1, takes place around Φ^*. Therefore, considering:

$$D'\text{vec}\left(\tilde{S}^Z\tilde{S}^{Z\prime}\Phi^{*\prime}\hat{W}\right) = D'(\hat{W}\otimes\tilde{S}^Z\tilde{S}^{Z\prime})\phi^* = D'(\hat{W}\otimes\tilde{S}^Z\tilde{S}^{Z\prime})(D\varphi^* + d),$$

implies:

$$\begin{aligned}\hat{\varphi}-\varphi^* &= \left(D'(\hat{W}\otimes\tilde{S}^Z\tilde{S}^{Z\prime})D\right)^{-1}\left((D'\text{vec}\left(\tilde{S}^Z(S^y-\Phi^*\tilde{S}^Z)'\hat{W}\right)\right) \\ &= \left(D'(\hat{W}\otimes\tilde{S}^Z\tilde{S}^{Z\prime})D\right)^{-1}\left(D'\text{vec}\left(\tilde{S}^Z(S^u-\Omega_{uv}\Omega_{vv}^{-1}X)'\hat{W}\right)\right).\end{aligned} \tag{38}$$

With condition (22) and $\hat{W}\to W>0$ in probability in place, it follows from straightforward calculations that:

$$\begin{aligned}A_D^{-1}(\hat{\varphi}-\varphi^*) \Rightarrow & \left(D^{*\prime}\left(W\otimes\int_0^1 f(s)f(s)'ds\right)D^*\right)^{-1} \\ & \times\left(D^{*\prime}\text{vec}\left(\int_0^1 f(s)B_{u\cdot v}(s)'dsW\right)\right),\end{aligned} \tag{39}$$

with the result as given in the main text in (25) following by partial integration. ■

Proof of Proposition 4. The result follows analogously to the result for the Wald-type statistic for linear hypotheses on ϕ derived in Proposition 2. An additional complication is that two asymptotic full rank conditions, one related to the matrix D relating φ and ϕ, given in (22), and one related to the restrictions matrix R_φ, given in (30), have to be fulfilled. Also, of course, $\hat{A}$ and $\hat{B}$ need to be properly scaled to converge. ■

Proof of Proposition 5. As in the proof of Vogelsang and Wagner (2014, Lemma 2), it is easiest to establish the asymptotic behavior of the modified residuals $\hat{S}^u_{\lfloor rT\rfloor,\mathrm{M}}$ by noting that they are equivalently given as the OLS residuals of the regression of S^y_t on $\tilde{S}^Z_t$ and M_t. Based on this observation, it can be shown that $T^{1/2}\sum_{t=2}^{\lfloor rT\rfloor}\Delta\hat{S}^u_{t,\mathrm{M}}\Rightarrow\Omega_{u\cdot v}^{1/2}P(r)$, with $P(r)$ defined *similarly* to (28) in Vogelsang and Wagner (2024), with the only difference being that $W_{u\cdot v}(r)$ is now an n-dimensional process rather than a scalar process.[29] The

[29] To be precise, $P(r) := \int_0^r dW_{u\cdot v}(s) - \int_0^1 dW_{u\cdot v}(s)[H(1)-H(s)]'\left(\int_0^1 h(s)h(s)'ds\right)^{-1}h(r)$. Note that $H(r)$ is – which requires full design – a functional of standard Brownian motions.

second important ingredient for asymptotically pivotal fixed-b inference is independence of $P(r)$ – as input in $Q(P)$ – and $\mathcal{Z}_{s_{\mathcal{R}}n} = (R^* V_{\text{IM}} R^{*\prime})^{-1/2}(R^*\mathcal{Y})$. This can be shown analogously to the $n = 1$ case in the proof of Vogelsang and Wagner (2024, Proposition 3), in particular (57)–(59).[30] Write V_{IM} as defined in (15) for brevity as $V_{\text{IM}} = \Omega_{u\cdot v} \otimes \mathcal{M}$ and consider – to conclude the proof – the asymptotic behavior of the modified covariance estimator which is the central term in the fixed-b Wald-type statistic $\tau_{\text{W},b}$ defined in (36):

$$\begin{aligned} A_R^{-1} R \hat{V}_{\text{IM,M}} R' A_R^{-1\prime} &\Rightarrow (I_n \otimes \mathcal{R}^*) \left(\Omega_{u\cdot v}^{1/2} Q(P) \Omega_{u\cdot v}^{1/2\prime} \otimes \mathcal{M} \right) (I_n \otimes \mathcal{R}^{*\prime}) \\ &= \Omega_{u\cdot v}^{1/2} Q(P) \Omega_{u\cdot v}^{1/2\prime} \otimes (\mathcal{R}^* \mathcal{M} \mathcal{R}^{*\prime}) \\ &= (\Omega_{u\cdot v} \otimes \mathcal{R}^* \mathcal{M} \mathcal{R}^{*\prime})^{1/2} \left(Q(P) \otimes I_{s_{\mathcal{R}}} \right) (\Omega_{u\cdot v} \otimes \mathcal{R}^* \mathcal{M} \mathcal{R}^{*\prime})^{1/2\prime} \\ &= (R^* V_{\text{IM}} R^{*\prime})^{1/2} \left(Q(P) \otimes I_{s_{\mathcal{R}}} \right) (R^* V_{\text{IM}} R^{*\prime})^{1/2\prime}. \end{aligned} \tag{40}$$

Combining the parts defining $\tau_{\text{W},b}$ establishes the result. ∎

[30] Since $\mathcal{Y}$, as given in (14), can – in the case of full design – be written as $\left(\Omega_{u\cdot v}^{1/2} \otimes \Pi^{-1\prime} (\int_0^1 g(s) g(s)' ds)^{-1}\right) \text{vec}\left(\int_0^1 [G(1) - G(s)] dW_{u\cdot v}(s)'\right)$, with $\Pi := \text{diag}(\Pi_Z, \Omega_{vv}^{1/2})$, the relevant component for showing independence is $\text{vec}\left(\int_0^1 [G(1) - G(s)] dW_{u\cdot v}(s)'\right)$.