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 least squares estimation for systems of cointegrating multivariate polynomial regressions, i. e., systems of regressions that include deterministic variables, integrated processes and products of these variables as regressors. The errors are allowed to be correlated across equations, over time and with the regressors. Since, under restrictions on the parameters or in case of non-identical regressors across equations, integrated modified OLS and GLS estimation do not, in general, coincide, we discuss in detail *restricted* integrated generalized least squares estimators and inference based upon them. Furthermore, we develop asymptotically pivotal fixed-b inference, available only in case of full design and for specific hypotheses.

JEL Classification: C12, C13, C32

Keywords: Integrated modified estimation, cointegrating multivariate polynomial regression, fixed-b inference, generalized least squares

*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.

1 Introduction

We discuss integrated modified least squares estimation for systems of cointegrating multivariate polynomial regressions (SCMPRs), i. e., for systems of regressions that contain deterministic variables, integrated processes and products of (non-negative) integer powers of these variables as regressors. The stationary errors are allowed to be serially correlated and the integrated regressors are allowed to be endogenous. The paper, thus, extends the analysis of Vogelsang and Wagner (2024) from the single equation to the system of equations case.

We use and extend the integrated modified (IM) estimation principle introduced for cointegrating linear regressions in Vogelsang and Wagner (2014). IM estimation has several key advantages: First, estimation is tuning parameter-free and only for inference a conditional long-run covariance (matrix) needs to be estimated. Second, IM estimation allows to perform fixed-b inference, designed to capture the impact of kernel and bandwidth choices in long-run covariance estimation on the sampling distribution of test statistics.[1] Third, as discussed in detail in Vogelsang and Wagner (2024), IM estimation can be straightforwardly extended to include not only integer powers of integrated regressors, but arbitrary non-negative integer products of these regressors.[2] One important application of regressions involving cross-products of integrated regressors are *Translog* functions, see, e. g., Christensen *et al.* (1971). A second important application is performing RESET-type specification testing *including* cross-products of (original) regressors, as discussed in detail for the single equation case in Vogelsang and Wagner (2024, Sections 2.4 and 3).

The consideration of systems of equations adds some additional aspects compared to the single equation case: First (see also the corresponding discussion in Wagner, 2023), systems of equations require a detailed consideration of generalized least squares estimators, here integrated modified generalized least squares (IM-GLS). This stems from the fact that OLS- and GLS-type estimators only necessarily coincide (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

[1] 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).

[2] Effectively, the paper fulfills a similar extension-to-systems role as Wagner (2023) has for Wagner and Hong (2016). These two earlier papers – discussing fully modified least squares estimation – only consider (systems of) cointegrating polynomial regressions where cross-products of the regressors are excluded.

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 feasible for hypotheses that are (essentially) identical across equations. Whilst this is restrictive, it does include, e. g., fixed-b RESET-type specification testing with identical null and auxiliary regressors.[3]

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:

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

with $y_t := (y_{1t}, \ldots, y_{nt})'$, $Z_t := (z_{1t}, \ldots, z_{|\mathcal{I}|t})'$, with $z_{it} = t^{i_0} x_{1t}^{i_1} \cdots x_{mt}^{i_m}$ and i_j non-negative integers for $j = 0, \ldots, m$, $\Theta \in \mathbb{R}^{n \times |\mathcal{I}|}$ and $x_t := (x_{1t}, \ldots, x_{mt})'$. The regressors z_{it}, $i = 1, \ldots, |\mathcal{I}|$ are ordered, e. g., by lexicographic ordering of the multi-indices $\mathbf{i} := (i_0, \ldots, i_m)$ 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 in, e. g., Vogelsang and Wagner (2024) 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 $\{x_t\}_{t \in \mathbb{Z}}$. Defining $\{\eta_t\}_{t \in \mathbb{Z}} := \{(u_t', v_t')'\}_{t \in \mathbb{Z}}$, a functional central limit holds:

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

[3]The restricted scope of fixed-b inference, however, implies that one need not consider fixed-b inference for IM-GLS estimators, since in the considered setting IM-OLS and IM-GLS coincide.

for $0 \leq r \leq 1$, with $W(r)$ denoting 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. In the case $\Omega_{uv} \neq 0$, 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 $\Omega_{vv} = \Omega_{vv}^{1/2}(\Omega_{vv}^{1/2})'$, we 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 $B_{u\cdot v}(r) := B_u(r) - \Omega_{uv}\Omega_{vv}^{-1}B_v(r)$.

2.2 Estimation and Inference

IM-OLS estimation is simply OLS estimation of the partial sum version of equation (1) that is augmented by the original integrated regressors:

$$\begin{aligned} S_t^y &= \Theta S_t^Z + \Gamma x_t + S_t^u, \qquad t = 1, \ldots, T, \\ &= \Phi \tilde{S}_t^Z + S_t^u, \end{aligned} \tag{5}$$

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 (5) can be written as:

$$\begin{aligned} S^y &= \Theta S^Z + \Gamma X + S^u, \\ &= \Phi \tilde{S}^Z + S^u, \end{aligned} \tag{6}$$

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), see, in particular, Remark 1, and known since Zellner (1962), for 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, without parameter restrictions and with identical regressors in all equations, it suffices to consider the system version of the *single equation* IM-OLS estimator for CMPRs discussed in Vogelsang and Wagner (2024).[4] The IM-OLS estimator $\hat{\Phi}$ is defined as the OLS estimator of Φ in (6), i. e.,:

$$\hat{\Phi} := (S^y \tilde{S}^{Z\prime})(\tilde{S}^Z \tilde{S}^{Z\prime})^{-1}. \tag{7}$$

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 limit process corresponding to the regressors $Z(r) := \lim_{T\to\infty} T^{1/2} A_{\text{IM},\Theta} Z_{\lfloor rT \rfloor}$ 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:*[5]

$$\begin{aligned}(\hat{\Phi} - \Phi^*)A_{IM}^{-1} &\Rightarrow \Omega_{u\cdot v}^{1/2} \int_0^1 W_{u\cdot v}(s) f(s)' ds \left(\int_0^1 f(s) f(s)' ds \right)^{-1} \\ &= \Omega_{u\cdot v}^{1/2} \int_0^1 dW_{u\cdot v}(s) [F(1) - F(s)]' \left(\int_0^1 f(s) f(s)' ds \right)^{-1},\end{aligned} \tag{8}$$

where:

$$f(r) := \begin{bmatrix} \int_0^r Z(s) ds \\ B_v(r) \end{bmatrix}, \quad F(r) := \int_0^r f(s) ds. \tag{9}$$

As indicated in Footnote 4, 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 (7). Defining $\phi := \text{vec}(\Phi')$ and $\phi^* := \text{vec}(\Phi^{*\prime})$, this leads to:

$$\hat{\phi} := \text{vec}\left((\tilde{S}^Z \tilde{S}^{Z\prime})^{-1} (\tilde{S}^Z S^{y\prime}) \right) = (I_n \otimes (\tilde{S}^Z \tilde{S}^{Z\prime})^{-1})(I_n \otimes \tilde{S}^Z)\text{vec}(S^{y\prime}) \tag{10}$$

[4] Later, when discussing hypothesis testing and estimation under restrictions, it is convenient to consider vectorized version(s) of (6), 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., $\text{vec}(S^{y\prime}) = (I_n \otimes \tilde{S}^{Z\prime})\text{vec}(\Phi') + \text{vec}(S^{u\prime})$.

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

and:

$$(I_n \otimes A_{\text{IM}}^{-1}) \left(\hat{\phi} - \phi^*\right) \tag{11}$$
$$\Rightarrow (\Omega_{u\cdot v}^{1/2} \otimes I_{|\mathcal{I}|+m})\text{vec}\left(\left(\int_0^1 f(s)f(s)'ds\right)^{-1} \int_0^1 [F(1) - F(s)]dW_{u\cdot v}(s)'\right).$$

Conditional upon $W_v(r)$, the limiting distribution given in (11) is normal with zero mean and (conditional) covariance matrix:

$$V_{\text{IM}} := \Omega_{u\cdot v} \otimes \left(\left(\int_0^1 f(s)f(s)'ds\right)^{-1} \right. \tag{12}$$
$$\left. \times \left(\int_0^1 \big[F(1) - F(s)\big]\big[F(1) - F(s)\big]'ds\right)\left(\int_0^1 f(s)f(s)'ds\right)^{-1}\right).$$

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 (12), i. e.,:

$$\hat{V}_{\text{IM}} := \hat{\Omega}_{u\cdot v} \otimes (\tilde{S}^Z \tilde{S}^{Z\prime})^{-1} CC' (\tilde{S}^Z \tilde{S}^{Z\prime})^{-1}, \tag{13}$$

with $C := (c_1, \dots, c_T)$, $c_t := S_T^{\tilde{S}^Z} - S_{t-1}^{\tilde{S}^Z}$ for $t = 1, \dots, T$, $S_t^{\tilde{S}^Z} := \sum_{j=1}^t \tilde{S}_j^Z$ and $S_0^{\tilde{S}^Z} = 0$. By construction, $(I_n \otimes A_{\text{IM}}^{-1})\hat{V}_{\text{IM}}(I_n \otimes A_{\text{IM}}^{-1}) \Rightarrow V_{\text{IM}}$.

The limiting distribution given in (11) in conjunction with the estimator $\hat{V}_{\text{IM}}$ given in (13) directly allow 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 is that none of the hypotheses tested involves elements of Γ, which is not estimated consistently.[6]

[6]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 $RK = (R_{\text{vec}(\Theta)}, 0_{s\times nm})$.

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:*

$$H_0 : Rvec(\Phi') = R\phi = r, \tag{14}$$

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:*

$$\lim_{T\to\infty} A_R^{-1} R(I_n \otimes A_{IM}) = R^*, \tag{15}$$

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:*

$$T_W := (R\hat{\phi} - r)' \left(R\hat{V}_{IM}R'\right)^{-1} (R\hat{\phi} - r) \Rightarrow \mathcal{O}_s, \tag{16}$$

with $\hat{V}_{IM}$ *as defined in* (13) *and* $\mathcal{O}_s$ *denoting a chi-squared distributed random variable with s degrees of freedom.*

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

$$T_t := \frac{R\hat{\phi} - r}{\sqrt{R\hat{V}_{IM}R'}} \Rightarrow \mathcal{Z}, \tag{17}$$

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 unrelated regressions (SUR) cointegration literature, see, e. g., Moon (1999), Moon and Perron (2005), Park and Ogaki (1991) or Wagner *et al.* (2020). In the case not all equations include the same set of regressors, OLS- and GLS-type estimation in general cease to be algebraically (and asymptotically) equivalent.[7] Potential choices concerning

[7]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.

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. Clearly, restricted IM-OLS estimation is contained as the special case with $\hat{W} = W = I_n$.[8]

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.,:

$$\phi = D\varphi + d, \tag{18}$$

with $D \in \mathbb{R}^{(|\mathcal{I}|+m)n \times g}$ of full column rank, $\varphi \in \mathbb{R}^g$ and $d \in \mathbb{R}^{(|\mathcal{I}|+m)n}$.[9] Given the mentioned fact that only the parameters in Θ are estimated consistently, we consider only restrictions on Θ and do not impose restrictions on Γ. Using, as in Footnote 6, that $\phi = K\text{vec}(\Phi)$, this implies that $D = K\text{diag}(D_{\text{vec}(\Theta)}, I_{nm})$ and $d = K(d'_{\text{vec}(\Theta)}, 0_{1\times nm})'$. Also as above, we need 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:

$$\lim_{T\to\infty} (I_n \otimes A_{\text{IM}}^{-1}) D A_D = D^*, \tag{19}$$

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

Proposition 3. *Let the data be generated by* (1) *with appropriate assumptions in place and ϕ fulfilling the (explicit) restrictions posited in* (18)*. Furthermore, assume that there exists a matrix sequence A_D such that condition* (19) *holds. The* restricted integrated

[8] 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 consider, in principle, also "full" $\hat{W} \in \mathbb{R}^{nT\times nT}$ weighting matrices.

[9] As is well known, the explicit formulation of restrictions used in (18) 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 $s = (|\mathcal{I}| + m)n - g$.

modified generalized least squares (IM-GLS) estimator $\hat{\phi}_R$ *of* ϕ *with symmetric weighting matrix sequence* $\hat{W}$ *is defined as:*

$$\hat{\phi}_R := D\hat{\varphi} + d, \tag{20}$$

with:

$$\begin{aligned}\hat{\varphi} := &\left((D'(\hat{W} \otimes \tilde{S}^Z \tilde{S}^{Z\prime})D\right)^{-1} \\ &\times \left(D'\left(vec\left(\tilde{S}^Z S^{y\prime}\hat{W}\right) - (\hat{W} \otimes \tilde{S}^Z \tilde{S}^{Z\prime})d\right)\right).\end{aligned} \tag{21}$$

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

$$\begin{aligned}A_D^{-1}(\hat{\varphi} - \varphi^*) \Rightarrow &\left(D^{*\prime}\left(W \otimes \int_0^1 f(s)f(s)'ds\right)D^*\right)^{-1} \\ &\times \left(D^{*\prime}\text{vec}\left(\int_0^1 \left[F(1) - F(s)\right]dB_{u\cdot v}(s)'W\right)\right).\end{aligned} \tag{22}$$

The limiting distribution of $\hat{\varphi}$ *given in* (22) *is – conditional upon* $W_v(r)$ *– normal with zero mean and covariance matrix:*

$$V_{IM,R} := A^{-1}BA^{-1}, \tag{23}$$

with:

$$A := D^{*\prime}\left(W \otimes \int_0^1 f(s)f(s)'ds\right)D^*, \tag{24}$$

$$B := D^{*\prime}\left(W\Omega_{u\cdot v}W \otimes \int_0^1 \left[F(1) - F(s)\right]\left[F(1) - F(s)\right]'ds\right)D^*. \tag{25}$$

An estimator of $V_{\text{IM,R}}$ is readily available, analogously to (13) and, therefore, asymptotically chi-squared or standard normal inference on φ follows, under conditions (19) and (27), similarly to Proposition 2:[10]

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 pa-*

[10]Note that the limiting distribution of $\hat{\phi}_{\text{R}}$ is, by construction, singular unless D is square.

rameter vector $\phi = D\varphi + d$ with condition (19) *in place fulfill s_φ linearly independent restrictions, i. e.,:*[11]

$$H_0 : \; R_\varphi \varphi = r_\varphi, \tag{26}$$

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:

$$\lim_{T\to\infty} A_\varphi^{-1} R_\varphi A_D = R_\varphi^* \tag{27}$$

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

$$T_{W,R} := (R_\varphi \hat{\varphi} - r_\varphi)' \left(R_\varphi \hat{A}^{-1} \hat{B} \hat{A}^{-1} R_\varphi' \right)^{-1} (R_\varphi \hat{\varphi} - r_\varphi) \Rightarrow \mathcal{O}_{s_\varphi}, \tag{28}$$

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

$$\hat{A} := D'(\hat{W} \otimes \tilde{S}^Z \tilde{S}^{Z\prime}) D, \tag{29}$$

$$\hat{B} := D'(\hat{W} \hat{\Omega}_{u \cdot v} \hat{W} \otimes CC') D. \tag{30}$$

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

Remark 1. 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 (19) 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 symmetric weighting matrix $\hat{W} \to W > 0$.

2.4 Fixed-b Inference

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

[11] Since Γ is not estimated consistently, the last nm columns of the restrictions matrix R_φ need to equal zero.

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. Full design means that the limit process $Z(r) = \Pi_Z Z_W(r)$, with Π_Z a regular matrix and $Z_W(r)$ a functional of standard Brownian motions.[12]

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 (12), 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. Whilst this is clearly restrictive, it, e. g., includes RESET-type specification testing with identical auxiliary regressors in all equations (see Vogelsang and Wagner, 2024, Section 2.4 for the $n = 1$ case).[13]

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:*[14]

$$H_0 : Rvec(\Phi') = (I_n \otimes \mathcal{R})\phi = r, \tag{31}$$

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:

$$\lim_{T \to \infty} A_{\mathcal{R}}^{-1} \mathcal{R} A_{IM} = \mathcal{R}^*, \tag{32}$$

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

[13] Since no additional restrictions are required for r, fixed-b inference is effectively available for testing the same set of hypotheses for every equation allowing for equation- (and restriction-) specific intercept terms.

[14] As in Proposition 2, we assume that none of the hypotheses tested involves elements of Γ, compare Footnote 6. This requires that the last m columns of $\mathcal{R}$ are zero.

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:

$$T_{W,b} := (R\hat{\phi} - r)' \left(R\hat{V}_{IM,M}R'\right)^{-1} (R\hat{\phi} - r) \Rightarrow \mathcal{Z}'_{s_\mathcal{R}n}(Q(P)^{-1} \otimes I_{s_\mathcal{R}})\mathcal{Z}_{s_\mathcal{R}n}, \tag{33}$$

with $\hat{V}_{IM,M}$ defined similarly as $\hat{V}_{IM}$ in (13), but with $\hat{\Omega}_{u\cdot v}$ replaced by $\hat{\Omega}_{u\cdot v,M}$, defined in (34) below, and $\mathcal{Z}_{s_\mathcal{R}n}$ an $s_\mathcal{R}n$-dimensional standard normally distributed random vector 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 $0 < b \leq 1$.[15]

It is key for asymptotically pivotal fixed-b inference, that $\mathcal{Z}_{s_\mathcal{R}n}$ and $Q(P)$ in (33) are independent random variables. Achieving independence requires (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 $\hat{S}^u_t := S^y_t - \hat{\Phi}\tilde{S}^Z_t$ and $\hat{S}^u := (\hat{S}^u_1, \ldots, \hat{S}^u_T)$. Instead, *orthogonalized* modified residuals, $\hat{S}^u_{t,\mathrm{M}}$, have to be used to annihilate (nuisance parameter-dependent) correlation. These are given by $\hat{S}^u_\mathrm{M} := \hat{S}_u \left(I_T - M^{\perp\prime}(M^\perp M^{\perp\prime})^{-1}M^\perp\right)$, 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 $M_t := t\sum_{j=1}^T \tilde{S}^Z_j - \sum_{j=1}^{t-1}\sum_{s=1}^j \tilde{S}^Z_s$ for $t = 1, \ldots, T$. The required modified estimator of $\Omega_{u\cdot v}$ is now defined as:

$$\begin{aligned} \hat{\Omega}_{u\cdot v,\mathrm{M}} &:= T^{-1}\sum_{i=2}^T\sum_{j=2}^T k\left(\frac{|i-j|}{B}\right)\Delta S^u_{i,\mathrm{M}}\Delta S^{u\prime}_{j,\mathrm{M}}, \\ &\Rightarrow \Omega^{1/2}_{u\cdot v}Q(P)\Omega^{1/2\prime}_{u\cdot v}, \end{aligned} \tag{34}$$

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

Remark 2. Note that in the case the restrictions considered in this subsection are not rejected, the discussion in Footnote 9 clarifies that the corresponding restricted estimation problem is, unsurprisingly, under the type of restrictions discussed in Remark 1. 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.

[15]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 Proposition 3 in Vogelsang and Wagner (2024), conveniently defined there already for the multivariate case.

Code for IM-OLS estimation and inference, including fixed-b inference – which necessitates (the generation of) fixed-b critical values that, as discussed, depend upon the specification of the SCMPR, the kernel function $k(\cdot)$ and the value of b – is available upon request.

References

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

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

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 (2005). 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.*

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., 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 (15) on A_R, it holds that:

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

with $\mathcal{Y}$ denoting the random variable (limiting distribution) given in (11). $R^*\mathcal{Y}$ is under the null hypothesis – conditional upon $W_v(r)$ – normally distributed with zero mean and covariance matrix $R^*V_{\text{IM}}R^{*\prime}$. Under condition (15), it furthermore holds that $A_R^{-1}R\hat{V}_{\text{IM}}R'A_R^{-1\prime} \Rightarrow R^*V_{\text{IM}}R^{*\prime}$. Combining the two results now immediately leads to the asymptotic chi-squared distribution for T_{W} as defined in (16) 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:

$$\hat{\varphi} - \varphi^* = \left(D'(\hat{W} \otimes \tilde{S}^Z \tilde{S}^{Z\prime})D\right)^{-1} \left((D'\text{vec}\left(\tilde{S}^Z (S^y - \Phi^* \tilde{S}^Z)'\hat{W}\right)\right) \tag{35}$$
$$= \left(D'(\hat{W} \otimes \tilde{S}^Z \tilde{S}^{Z\prime})D\right)^{-1} \left(D'\text{vec}\left(\tilde{S}^Z (S^u - \Omega_{uv}\Omega_{vv}^{-1} X)'\hat{W}\right)\right).$$

With condition (19) and $\hat{W} \to W > 0$ in place, it follows from straightforward calculations that:

$$A_D^{-1}(\hat{\varphi} - \varphi^*) \Rightarrow \left(D^{*\prime}\left(W \otimes \int_0^1 f(s)f(s)'ds\right)D^*\right)^{-1} \tag{36}$$
$$\times \left(D^{*\prime}\text{vec}\left(\int_0^1 f(s)B_{u\cdot v}(s)'dsW\right)\right),$$

with the result as given in the main text in (22) following by partial integration. □

Proof of Proposition 4. The result follows analogously to the result for the Wald-type statistic for linear hypothesis 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 (19), and one related to the restrictions matrix R_φ, given in (27), have to hold. 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.[16] The second important ingredient for asymptotically pivotal fixed-b inference is independence of $P(r)$ – as input in $Q(P)$ – and $\mathcal{Z}_{s_{\mathcal{R}}n} = (R^* V_{\mathrm{IM}} R^{*\prime})^{-1/2}(R^*\mathcal{Y})$. This can be shown analogously as in the $n = 1$ case in the proof of Vogelsang and Wagner (2024, Proposition 3), in particular (57)–(59).[17] Write V_{IM} as defined in (12) for brevity as $V_{\mathrm{IM}} = \Omega_{u\cdot v} \otimes \mathcal{M}$ and consider – to conclude the proof – the asymptotic behavior of

[16]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.

[17]Since $\mathcal{Y}$, as given in (11), can – in 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 $\text{vec}\left(\int_0^1 [G(1) - G(s)]dW_{u\cdot v}(s)'\right)$.

the modified covariance estimator which is the central term in the fixed-b Wald-type statistic $T_{\mathrm{W},b}$ defined in (33):

$$\begin{aligned} A_R^{-1} R \hat{V}_{\mathrm{IM,M}} R' A_R^{-1\prime} &\Rightarrow (I_n \otimes \mathcal{R}^*) \left(\Omega_{u \cdot v}^{1/2} Q(P) \Omega_{u \cdot v}^{1/2\prime} \otimes \mathcal{M} \right) (I_n \otimes \mathcal{R}^{*\prime}) \\ &= \Omega_{u \cdot v}^{1/2} Q(P) \Omega_{u \cdot v}^{1/2\prime} \otimes (\mathcal{R}^* \mathcal{M} \mathcal{R}^{*\prime}) \\ &= (\Omega_{u \cdot v} \otimes \mathcal{R}^* \mathcal{M} \mathcal{R}^{*\prime})^{1/2} (Q(P) \otimes I_{s_{\mathcal{R}}}) (\Omega_{u \cdot v} \otimes \mathcal{R}^* \mathcal{M} \mathcal{R}^{*\prime})^{1/2\prime} \\ &= (R^* V_{\mathrm{IM}} R^{*\prime})^{1/2} (Q(P) \otimes I_{s_{\mathcal{R}}}) (R^* V_{\mathrm{IM}} R^{*\prime})^{1/2\prime}. \end{aligned} \tag{37}$$

Combining the parts defining $T_{\mathrm{W},b}$ establishes the result. □