IHS Working Paper 53
April 2024

Integrated Modified OLS Estimation and Fixed-b Inference for Cointegrating Multivariate Polynomial Regressions

Timothy J. Vogelsang
Martin Wagner

Institut für Höhere Studien
Institute for Advanced Studies
Vienna

Integrated Modified OLS Estimation and Fixed-b Inference for Cointegrating Multivariate Polynomial Regressions

Timothy J. Vogelsang[1] and Martin Wagner[2,3,4]

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

Abstract

This paper shows that the integrated modified OLS (IM-OLS) estimator developed for cointegrating linear regressions in Vogelsang and Wagner (2014a) can be straightforwardly extended to cointegrating multivariate polynomial regressions. These are regression models that include as explanatory variables 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 regressors are allowed to be endogenous. The IM-OLS estimator is tuning-parameter free and does not require the estimation of any long-run variances. A scalar long-run variance, however, has to be estimated and scaled out when using IM-OLS for inference. In this respect, we consider both standard asymptotic inference as well as fixed-b inference. Fixed-b inference requires that the regression model is of full design. The results may be particularly interesting for specification testing of cointegrating relationships, with RESET-type specification tests following immediately. The simulation section also zooms in on RESET specification testing and illustrates that the performance of IM-OLS is qualitatively comparable to its performance in cointegrating linear regressions.

JEL Classification: C12, C13, C32

Keywords: Cointegration, fixed-b asymptotics, IM-OLS, multivariate polynomials, nonlinearity, RESET

1 Introduction

The integrated modified OLS (IM-OLS) estimator introduced in Vogelsang and Wagner (2014a) for cointegrating linear regressions is, algebraically and computationally, the simplest modified least squares estimator for cointegrating regressions in that, for estimation, no kernel, bandwidth or lead and lag choices need to be made: One OLS regression is all that is required. When using the IM-OLS estimator to test hypotheses, however, a scalar (conditional) long-run variance parameter needs to be estimated and scaled out, thus IM-OLS also involves some user choices on kernel and bandwidth when used for inference. The simplicity of IM-OLS is driven by the underlying idea of performing OLS estimation in the *partial sum* regression augmented by the integrated regressors rather than in the original cointegrating regression, where OLS would be consistent but result in a limiting distribution contaminated by so-called second-order bias terms. Partial summation has several advantages: First, it eliminates an additive bias term present in the limiting distribution of, e. g., the OLS estimator of the original cointegrating regression. Second, it renders asymptotic theory almost trivial, since stochastic integrals with respect to Brownian motion – resulting from the cross-product of the integrated regressors and the errors – are replaced by Riemann integrals of products of (integrals of) Brownian motions. Third, the inclusion of the integrated regressors suffices to allow for asymptotic standard inference, as these regressors "soak up" all second-order endogeneity biases. This stems directly from the fact that in the partial sum regression augmented by the integrated regressors, both the integrated regressors and the partially summed errors are integrated processes, see Section 2.2.

This paper demonstrates that the two building blocks of IM-OLS allow to extend the estimator straightforwardly to, what we call, cointegrating multivariate polynomial regressions (CMPR), see (3) below for the general formulation. These are regression models that include, potentially, the constant and polynomial time trends (with more general deterministic functions of time possible, see Remark 1), m integrated – but not cointegrated – processes of order one, as well as products of (non-negative) integer powers of these variables as regressors. The stationary errors are allowed to be serially correlated and the regressors are allowed to be endogenous.[1] As an illustration, consider (see Christensen *et al.*, 1971) a two-factor Translog (cost, demand or) production function (ignoring total factor productivity and deterministic components for brevity):

$$\ln Y_t = \ln K_t \beta_1 + \ln L_t \beta_2 + (\ln K_t)^2 \beta_3 + (\ln L_t)^2 \beta_4 + (\ln K_t \ln L_t)\beta_5 + u_t, \quad (1)$$

with Y_t denoting output, K_t capital and L_t labor. Under the assumption that both $\ln K_t$ and $\ln L_t$ are integrated (but not cointegrated), the above equation (1) is, in the case of stationary errors, an example of a cointegrating multivariate polynomial regression as considered in this paper.[2] For the example considered, the IM-OLS estimator is defined as the OLS estimator in the following regression model:

$$S_t^{\ln Y} = S_t^{\ln K}\beta_1 + S_t^{lnL}\beta_2 + S_t^{(\ln K)^2}\beta_3 + S_t^{(\ln L)^2}\beta_4 + S_t^{(\ln K \ln L)}\beta_5 + \ln K_t \gamma_1 + \ln L_t \gamma_2 + S_t^u, \quad (2)$$

[1] The term cointegrating multivariate polynomial regressions (CMPRs) is based on the term cointegrating polynomial regressions (CPRs) coined by Wagner and Hong (2016), who consider a regression in which deterministic regressors, integrated regressors and their powers, *but no* cross-products are included. Wagner and Hong (2016, Remark 1) discuss the potential of generalizing their analysis to the setting considered here and correctly observe that this will be easier for IM-OLS than for fully modified OLS (FM-OLS), the estimator considered in Wagner and Hong (2016), originally developed for cointegrating linear regressions in Phillips and Hansen (1990). The present paper is a substantially extended version of our earlier paper Vogelsang and Wagner (2014b) that only considered IM-OLS to perform RESET-type specification testing of the null hypothesis of a cointegrating linear relationship.

[2] Note for completeness that Hansen (1992) contains the necessary functional central limit results to extend the FM-OLS estimation principle to the two-factor Translog setting, i. e., to polynomials of degree two. To exemplify the difference between the CMPR and the CPR settings, the results of Wagner and Hong (2016) do not allow the inclusion of the regressor $(\ln K_t \ln L_t)$.

with $S_t^{\ln Y} := \sum_{j=1}^{t} \ln Y_j$ and the other partial sum variables defined analogously. To state the obvious, considering polynomials means that the regression problem is linear in parameters – also after partial summation, which implies, to continue stating the obvious, that the partial sum regression problem remains linear in parameters and thus allows for closed-form solutions of least squares estimators. We stress this fact because linearity-in-parameters is the key property of the regression function that interacts conveniently with the IM-OLS partial summation step. In other words, for cointegrating regression problems that are nonlinear in parameters (and feature additive stationary errors), partial summation may still be performed, but will result in nonlinear least squares problems with integrated errors, rather than as in the linear-in-parameters context, where partial summation plus augmentations leads to ordinary least squares estimation with a limiting distribution that allows for asymptotic standard inference. The restriction to polynomials in several variables is, of course, a restriction to a narrow set of functions, however, one that allows to interact different integrated regressors freely in this function class. A large part of the (nonlinear-in-parameters or nonparametric) nonlinear cointegration literature considers only one integrated regressor or restricts the integrated regressors to enter additively separably, for an early contribution see, e. g., Chang *et al.* (2001).[3] One exception to additive separability is, e. g., Dong *et al.* (2016), who allow for multiple integrated regressors that, however, enter the regression model in the form of a single index, i. e., as one linear combination with a parameter vector to be estimated. Gao and Phillips (2013) consider a triangular systems formulation in a semiparametric framework with, unavoidably, a finely calibrated set of assumptions on the system.

Compared to more general nonlinear settings, the CMPR setting is, as discussed, quite restrictive but easy to deal with in terms of statistical theory. On an imagined tradeoff line between generality of functional form and simplicity of estimation and inference, the CMPR setting is – in combination with the IM-OLS estimator – clearly on the simplicity end of the spectrum. Nevertheless, in addition to estimation and inference in, e. g., Translog-type relationships with an arbitrary number of factors, the CMPR setting is *useful in particular also* for developing simple specification tests for cointegrating relationships. In general, economic theory does not specify the precise form of a potential nonlinear cointegrating relationship in the case a linear (or CPR or Translog-type) cointegrating relationship is found to not adequately describe the relationship between the variables. Therefore, we believe that an *omnibus* specification test such as the RESET test (see Ramsey, 1969) is a useful and computationally cheap addition to the cointegration specification testing toolkit.[4] This type of test is convenient – and fits perfectly to the CMPR setting considered in this paper – as it is based on the replacement of an unknown nonlinear function by a finite sum approximation, in the RESET case a polynomial approximation. Such approaches have a long tradition in specification tests for stationary time series

[3]That paper is a generalization of Park and Phillips (2001), who consider nonlinear cointegrating regressions with only one integrated regressor and derive asymptotic theory for two classes of nonlinear functions, labelled as asymptotically homogenous (which include polynomials) and integrable, with quite different asymptotic behavior of parameter estimators in both cases. Since these early contributions, the literature on (effectively often bivariate) nonlinear cointegration has exploded, with both parametric and nonparametric approaches. See, from an even longer list (here in alphabetical order), Chan and Wang (2015), Chang and Park (2011), Dong and Linton (2018), Dong *et al.* (2022), Karlsen *et al.* (2007), Linton and Wang (2016), Saikkonen and Choi (2004), Shi and Phillips (2012), Wang and Phillips (2009, 2016) or Wang and Wang (2013). The listed papers differ *inter alia* with respect to precise assumptions, in particular with respect to whether error serial correlation and/or regressor endogeneity are allowed for. Some papers allow for time trends and some allow to also include stationary regressors, usually considered to be strictly exogenous. See Tjøstheim (2020) for a recent (partial) survey of this literature.

[4]In fact, a significant part of the applied nonlinear cointegration literature actually considers nonlinear adjustment mechanisms towards linear cointegrating relationships, see, e. g., Balke and Fomby (1997), Bec and Rahbek (2004) or Hansen and Seo (2002). There are also some contributions considering nonlinear cointegrating relationships with a specific functional form, e. g., Saikkonen and Choi (2004) consider cointegrating smooth transition regressions. Notwithstanding the fact that nonlinear relationships are usually not fully specified by economic theory, there is increasing interest in nonlinear cointegrating relationships spurred from different areas of application, ranging from empirical macroeconomics, e. g., deviations from purchasing power parity (Hong and Phillips, 2010), or linearity of money demand functions (Lütkepohl *et al.*, 1999; Choi and Saikkonen, 2010) to empirical finance, e. g., currency crises (Saikkonen and Choi, 2004) to environmental economics, e. g., the environmental Kuznets curve (Wagner, 2015).

models, see, e. g., Phillips (1983), Lee *et al.* (1993) or de Benedictis and Giles (1998). Given that we think that the RESET-type test may be of particular interest to many readers, we discuss RESET test options based on IM-OLS in detail in Section 2.4.[5]

As is well known, standard asymptotic theory does not capture the impact of kernel and bandwidth choices, required in our context for long-run variance estimation, on the sampling distributions of estimators and test statistics based upon them. Fixed-b asymptotic theory, put forward in the stationary context by Kiefer and Vogelsang (2005) and in the unit root and cointegration framework by Vogelsang and Wagner (2013, 2014a), captures the impact of kernel and bandwidth choices on the sampling distributions of HAC-type test statistics.[6] Fixed-b asymptotic theory for testing hypotheses on the parameters of a CMPR imposes some restrictions on the set of auxiliary regressors, in that a so-called *full design* (see the details in Section 2.3) has to be chosen, which essentially means that all powers and cross-products of powers of the integrated regressors up to the chosen maximum degree have to be included. Exactly as in the linear cointegration case treated in Vogelsang and Wagner (2014a), pivotal fixed-b inference (in the full design case) rests upon long-run variance estimation based on specifically modified residuals since the IM-OLS residuals cannot be used directly.

The theoretical analysis is complemented by a small simulation study to assess the finite sample performance of the proposed standard and fixed-b tests, where we only discuss the RESET-type specification test for brevity. The simulations show that the fixed-b limit theory describes the distribution of the test statistic well. Altogether, the findings of the simulation study are typical for the cointegration and fixed-b literatures. The performance of the tests is deteriorating if regressor endogeneity and error serial correlation are increasing for a given sample size, with this fact being true for both classical and to a lesser extent fixed-b testing. Fixed-b tests often – and also in the present situation – incur smaller size distortions at the expense of only minor losses in (size-adjusted) power than standard tests. It is worth noting that the RESET-type tests exhibit power also against smoothly varying logistic alternatives and not only against polynomial alternatives.[7]

The paper is organized as follows: Section 2 describes the setup and assumptions, IM-OLS estimation, fixed-b inference and the RESET-type test in its four subsections. Section 3 provides a finite sample performance analysis of the RESET-type test and Section 4 briefly concludes. All proofs are relegated to the appendix. Supplementary material available upon request provides fixed-b critical values as well as code to use IM-OLS for estimation and inference in CMPR settings, including, of course, material to perform RESET-type specification testing.

We use the following notation: Definitional equality is signified by $:=$, weak convergence by $\Rightarrow$, convergence in distribution by $\xrightarrow{d}$ and convergence in probability by $\xrightarrow{\mathbb{P}}$. The integer part of $x \in \mathbb{R}$ is given by $\lfloor x \rfloor$ and a diagonal matrix with entries specified throughout by $\mathrm{diag}(\cdot)$. For a matrix $A \in \mathbb{R}^{m\times n}$ we denote its Frobenius norm by $\|A\| := \sqrt{\mathrm{tr}(A'A)}$, where $\mathrm{tr}(\cdot)$ denotes the trace. Identity matrices with dimensions specified are denoted by I. We write $\mathbb{N}_0$ to denote the natural numbers extended by zero, $\mathbb{Z}$ for the integers, $\mathbb{R}$ for the real numbers and $|\mathcal{I}|$ for the number of elements in a (multi-index) set $\mathcal{I}$. L is the backward-shift operator, i. e., $L\{x_t\}_{t\in\mathbb{Z}} = \{x_{t-1}\}_{t\in\mathbb{Z}}$. Brownian motions, with covariance matrices specified in the context, are denoted by $B(r)$. Standard Brownian motions are denoted by $W(r)$.

[5] In close relation to the fast growing nonlinear cointegration literature referred to in Footnote 3, a number of tests for the correct specification of (nonlinear) cointegrating relationships has also been developed, see (again in alphabetical order), e. g., Choi and Saikkonen (2010), Dong and Gao (2018), Hong and Phillips (2010), Kasparis (2008), Wang and Phillips (2012), Wang and Zhu (2020) or Wang *et al.* (2018).

[6] Additional (partial) fixed-b treatments of cointegration models are contained in Bunzel (2006) and Jin *et al.* (2006). Hwang and Vogelsang (2022) extend fixed-b theory to regressions with high-frequency data.

[7] Grabarczyk and Wagner (2024) present a small comparative simulation study of IM-OLS in the context of an analysis of the environmental Kuznets curve for carbon dioxide emissions using three modified least squares estimators. The findings are qualitatively very similar to the findings in the cointegrating linear regression case considered in Vogelsang and Wagner (2014a).

2 Theory

2.1 Setup and Assumptions

We consider a *cointegrating multivariate polynomial regression* (CMPR) model including, potentially, the constant and (polynomial) time trends, m integrated – but not cointegrated – processes of order one, $x_{1t}, \dots, x_{mt}$, as well as products of (non-negative) integer powers of these variables as regressors, and a stationary error term u_t. Using multi-index notation, the regression model considered can be written compactly as:

$$\begin{aligned} y_t &= \sum_{\mathbf{i}:=(i_0,i_1,\dots,i_m)\in\mathcal{I}} t^{i_0} x_{1t}^{i_1} \cdots x_{mt}^{i_m} \theta_{i_0,i_1,\dots,i_m} + u_t \\ &= \sum_{i=1}^{|\mathcal{I}|} z_{it}\theta_i + u_t, = Z_t'\theta + u_t \\ x_t &= x_{t-1} + v_t, \end{aligned} \tag{3}$$

with the second sum considered for $i = 1, \dots, |\mathcal{I}|$ by attaching to every multi-index $\mathbf{i}$ an integer i, e. g., by lexicographic ordering. This defines $z_{it} := t^{i_0} x_{1t}^{i_1} \cdots x_{mt}^{i_m}$ for i corresponding to $\mathbf{i} = (i_0, i_1, \dots, i_m) \in \mathbb{N}_0^{m+1}$, $Z_t := [z_{1t}, \dots, z_{|\mathcal{I}|t}]'$, $\theta := [\theta_1, \dots, \theta_{|\mathcal{I}|}]'$ and $x_t := [x_{1t}, \dots, x_{mt}]'$. To avoid perfect multicollinearity of the regressors by construction, we assume that no multi-index $\mathbf{i}$ is contained more than once in $\mathcal{I}$ and, for finite sample calculations, that $|\mathcal{I}| \le T$. To not overload notational complexity we abstain from carrying these basic requirements through the paper.

Given that the unit root and cointegration literature offers a variety of primitive assumptions leading to the required convergence results, we do not posit a set of sufficient assumptions that lead to the required functional central limit theorems for $\{\eta_t\}_{t\in\mathbb{Z}} := \{[u_t, v_t']'\}_{t\in\mathbb{Z}}$ in the main text:[8]

$$T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} \eta_t \Rightarrow B(r) = \Omega^{1/2} W(r), \quad 0 \le r \le 1, \tag{4}$$

where $W(r)$ is an $(m+1)$-dimensional vector of independent standard Brownian motions and:

$$\Omega = \begin{bmatrix} \Omega_{uu} & \Omega_{uv} \\ \Omega_{vu} & \Omega_{vv} \end{bmatrix} := \sum_{j=-\infty}^{\infty} \mathbb{E}(\eta_t \eta_{t-j}'), \tag{5}$$

positive definite, where clearly $\Omega_{vu} = \Omega_{uv}'$. In the case $\Omega_{uv} \neq 0$, the regressors are endogenous and, in addition to regressor endogeneity, the setting also allows for relatively unrestricted forms of serial correlation of the errors $\{\eta_t\}_{t\in\mathbb{Z}}$. These two aspects, in general, necessitate some form of modified least squares estimation to allow for asymptotic standard inference. Partitioning $B(r) = [B_u(r), B_v(r)']'$ and $W(r) = [w_{u\cdot v}(r), W_v(r)']'$ we have, using, e. g., the Cholesky decomposition of $\Omega_{vv} = \Omega_{vv}^{1/2}(\Omega_{vv}^{1/2})'$:

$$\begin{bmatrix} B_u(r) \\ B_v(r) \end{bmatrix} = \begin{bmatrix} \omega_{u\cdot v}^{1/2} & \Omega_{uv}(\Omega_{vv}^{-1/2})' \\ 0 & \Omega_{vv}^{1/2} \end{bmatrix} \begin{bmatrix} w_{u\cdot v}(r) \\ W_v(r) \end{bmatrix}, \tag{6}$$

with $\omega_{u\cdot v} := \Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}$ the (innovation) variance of $B_{u\cdot v}(r) := B_u(r) - B_v(r)'\Omega_{vv}^{-1}\Omega_{vu}$.

[8] One set of sufficient assumptions for the results in this paper, related to Wagner and Hong (2016, Assumption 1), is to assume that $\eta_t = C(L)\eta_t^0 = \sum_{j=0}^{\infty} C_j \eta_{t-j}^0$, with $\sum_{j=0}^{\infty} j||C_j|| < \infty$ and $\det(C(1)) \neq 0$. Here, the process $\{\eta_t^0\}_{t\in\mathbb{Z}}$ is a strictly stationary and ergodic martingale difference sequence with natural filtration $\mathcal{F}_t = \sigma(\{\eta_s^0\}_{-\infty}^t)$, positive definite covariance matrix $\Sigma_{\eta^0\eta^0}$ and $\sup_{t\in\mathbb{Z}} \mathbb{E}[\|\eta_t^0\|^r | \mathcal{F}_{t-1}] < \infty$ a.s. for some $r > 4$.

Remark 1. In (3) 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:

$$y_t = D_t'\theta_D + \sum_{\mathbf{i}:=(i_0,i_1,\ldots,i_m)\in\mathcal{I}^*} t^{i_0} x_{1t}^{i_1} \cdots x_{mt}^{i_m} \theta_{i_0,i_1,\ldots,i_m} + u_t, \tag{7}$$

with $\mathcal{I}^*$ denoting a set of multi-indices $\mathbf{i}$ with $\min_{j=1,\ldots,m} i_j > 0$. 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 G_D and a p-dimensional vector of functions $D(z)$ such that for $0 \leq r \leq 1$ it holds that:

$$\lim_{T\to\infty} \sqrt{T} G_D^{-1} D_{\lfloor rT \rfloor} = D(r) \text{ with } 0 < \int_0^r D(z)D(z)'dz < \infty. \tag{8}$$

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

To exemplify the setting and the multi-index notation, consider again the two-factor Translog function considered already in the introduction in (1), i. e.,:

$$y_t = \theta_{0,0,0} + x_{1t}\theta_{0,1,0} + x_{2t}\theta_{0,0,1} + x_{1t}^2\theta_{0,2,0} + x_{2t}^2\theta_{0,0,2} + x_{1t}x_{2t}\theta_{0,1,1} + u_t, \tag{9}$$

with corresponding index set $\mathcal{I} = \{(0,0,0), (0,1,0), (0,0,1), (0,2,0), (0,0,2), (0,1,1)\}$ and $|\mathcal{I}| = 6$.

2.2 IM-OLS Estimation

Parameter estimation is based on a (straightforward) extension of the integrated modified OLS (IM-OLS) estimator introduced in Vogelsang and Wagner (2014a) from cointegrating linear regressions to the cointegrating multivariate polynomial regression setting considered. IM-OLS corresponds to OLS estimation in the partial sum version of (3) augmented by the integrated regressors x_t:

$$\begin{aligned} S_t^y &= S_t^{Z\prime}\theta + x_t'\gamma + S_t^u \\ &= \tilde{S}_t^{Z\prime}\vartheta + S_t^u, \end{aligned} \tag{10}$$

with $S_t^y := \sum_{j=1}^t y_j$ and S_t^Z and S_t^u defined analogously, $\tilde{S}_t^Z := [S_t^{Z\prime}, x_t']'$ and $\vartheta := [\theta', \gamma']'$. As mentioned in the introduction, the inclusion of the original integrated regressors x_t corrects for endogeneity and thus allows for a zero-mean Gaussian mixture limiting distribution, see also the discussion below Proposition 1. The above equation (10) can be written, as usual, in matrix form by combining all observations $t = 1, \ldots, T$, i. e.,:

$$\begin{aligned} S^y &= S^Z\theta + X\gamma + S^u \\ &= \tilde{S}^Z\vartheta + S^u, \end{aligned} \tag{11}$$

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 $\tilde{S}^Z := [\tilde{S}_1^Z, \ldots, \tilde{S}_T^Z]' = [S^Z, X]$. The IM-OLS estimator of ϑ is now defined as the OLS estimator of ϑ in (11), i. e.,:

$$\hat{\vartheta} := \left(\tilde{S}^{Z\prime}\tilde{S}^Z\right)^{-1} \tilde{S}^{Z\prime} S^u. \tag{12}$$

For a discussion of the asymptotic properties of the IM-OLS estimator, it remains to define two quantities: First, the scaling matrix $A_{\text{IM}} := \text{diag}(A_{\text{IM},\theta}, I_m)$, with $A_{\text{IM},\theta} \in \mathbb{R}^{|\mathcal{I}|\times|\mathcal{I}|}$ a diagonal matrix with the entry corresponding to regressor $t^{i_0}x_{1t}^{i_1}\cdots x_{mt}^{i_m}$ given by $T^{-\left(i_0+\frac{\sum_{j=1}^m i_j}{2}+\frac{1}{2}\right)}$. Second, the limiting process corresponding to the properly scaled regressors, 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 $Z(r) := [z_1(r),\ldots,z_{|\mathcal{I}|}(r)]'$, $z_i(r) := r^{i_o}B_{v_1}(r)^{i_1}\cdots B_{v_m}(r)^{i_m}$ for $i = 1,\ldots,|\mathcal{I}|$ and $B_{v_j}(r)$ denoting the j-th component of $B_v(r)$.

Proposition 1. *Let the data be generated by* (3) *with the assumptions discussed in the previous subsection in place. Define* $\vartheta^* := [\theta', (\Omega_{vv}^{-1}\Omega_{vu})']'$*, then as* $T\to\infty$ *it holds that:*

$$A_{IM}^{-1}\left(\hat{\vartheta}-\vartheta^*\right) \Rightarrow \omega_{u\cdot v}^{1/2}\left(\int_0^1 f(s)f(s)'ds\right)^{-1}\int_0^1 f(s)w_{u\cdot v}(s)ds \qquad (13)$$
$$= \omega_{u\cdot v}^{1/2}\left(\int_0^1 f(s)f(s)'ds\right)^{-1}\int_0^1 [F(1)-F(s)]dw_{u\cdot v}(s),$$

where:

$$f(r) := \begin{bmatrix} \int_0^r Z(s)ds \\ B_v(r)\end{bmatrix}, \quad F(r) := \int_0^r f(s)ds. \qquad (14)$$

Note that the parameter sub-vector γ is not centered at zero, but at $\Omega_{vv}^{-1}\Omega_{vu}$, reflecting that the inclusion of x_t in the partial sum regression "soaks" up all (long-run) correlation between $\{u_t\}_{t\in\mathbb{Z}}$ and $\{v_t\}_{t\in\mathbb{Z}}$. This is immediately seen from $S_t^y = S_t^Z\theta + S_t^u = S_t^Z\gamma + x_t'\Omega_{vv}^{-1}\Omega_{vu} + (S_t^u - x_t'\Omega_{vv}^{-1}\Omega_{vu})$, with $T^{-1/2}(S_{\lfloor rT\rfloor}^u - x_{\lfloor rT\rfloor}'\Omega_{vv}^{-1}\Omega_{vu}) \Rightarrow B_{u\cdot v}(r)$, which is independent of $B_v(r)$ by construction. Thus, the limiting process corresponding to the errors in the augmented partial sum regression (11) is asymptotically independent of the (limiting process corresponding to the stochastic) regressors, which is the key requirement for asymptotic standard inference. Also, note that the coefficient estimator $\hat{\gamma}$ is not consistent.

Conditional upon $W_v(r)$, the limiting distribution given in (13) is normal with zero mean and the following conditional covariance matrix:

$$V_{\text{IM}} = \omega_{u\cdot v}\left(\int_0^1 f(s)f(s)'ds\right)^{-1}\left(\int_0^1 [F(1)-F(s)][F(1)-F(s)]'ds\right)\left(\int_0^1 f(s)f(s)'ds\right)^{-1}. \qquad (15)$$

In conjunction with consistent estimation of $\omega_{u\cdot v}$, the limiting distribution given in (13) allows for asymptotic standard normal or chi-squared inference on θ, using, e. g., Wald-type tests. In this respect, three comments are in order: First, for the same reason as discussed in detail in Vogelsang and Wagner (2014a, Section 5), the (first) differences of the IM-OLS residuals, $\Delta\hat{S}_t^u$ say, with $\hat{S}_t^u := S_t^y - \tilde{S}_t^{Z\prime}\hat{\vartheta}$, cannot be used to consistently estimate $\omega_{u\cdot v}$ by standard long-run variance estimation procedures. It can be shown that their usage is bound to result in conservative test statistics even asymptotically. A consistent estimator of $\omega_{u\cdot v}$ is most easily obtained using the OLS residuals of (3), $\hat{u}_t$ say, and by using $\hat{\eta}_t := [\hat{u}_t, v_t']'$ to estimate Ω and thereby $\omega_{u\cdot v}$. Given a consistent estimator, $\hat{\omega}_{u\cdot v}$ say, of $\omega_{u\cdot v}$, an – up to scaling – estimator of V_{IM} immediately follows by simply using the sample counterparts of the expressions appearing in the limiting (conditional) covariance matrix, i. e.,:

$$\hat{V}_{\text{IM}} := \hat{\omega}_{u\cdot v}\left(\tilde{S}^{Z\prime}\tilde{S}^Z\right)^{-1}C'C\left(\tilde{S}^{Z\prime}\tilde{S}^Z\right)^{-1}, \qquad (16)$$

with $C := [c_1,\ldots,c_T]'$, $c_t := S_T^{\tilde{S}^Z} - S_{t-1}^{\tilde{S}^Z}$, $S_t^{\tilde{S}^Z} := \sum_{j=1}^t \tilde{S}_j^Z$ for $t = 1,\ldots,T$ and $S_0^{\tilde{S}^Z} = 0$. By construction, $A_{\text{IM}}^{-1}\hat{V}_{\text{IM}}A_{\text{IM}}^{-1} \Rightarrow V_{\text{IM}}$.

Second, since $\hat{\theta}$ (and thus $\hat{\vartheta}$) in general contains elements converging at different rates, obtaining a formal (standard) result for Wald-type tests requires a condition on the restriction matrix R (in case of linear hypotheses) that is unnecessary when all estimated coefficients converge at the same rate (see, e. g., Park and Phillips, 1988, 1989, Sims *et al.* 1990, Vogelsang and Wagner, 2014a or Wagner and Hong, 2016). As described in Vogelsang and Wagner (2014a): "*For a given constraint (a row of the restrictions matrix R) the coefficient estimator with the slowest convergence rate dominates the asymptotic distribution of the linear combination implied by this constraint. When there are several restrictions being tested, it is not necessarily the case that the slowest converging coefficient estimator dominates a given restriction. Should another restriction also involve that slowest converging coefficient estimator, it is possible that the restrictions can be rotated so that (i) the slowest rate estimator only appears in one restriction and (ii) the Wald-type statistics have asymptotic chi-squared null distributions. If none of the null hypotheses mixes coefficients with different convergence rates no additional complications arise compared to a standard situation with all estimated coefficients converging at the same rate.*" Third, one can only test restrictions involving elements of θ and not of γ, which is, as mentioned, not consistently estimated. This means that the last m columns of the restrictions matrix R have to be equal to zero, i. e., $R = [R_\theta, 0_{s\times m}]$.

Proposition 2. *Let the data be generated by* (3) *with the assumptions discussed in the previous subsection in place and assume that long-run covariance estimation is performed consistently.*[9] *Consider s linearly independent restrictions on θ collected in $H_0 : R\vartheta = R_\theta\theta = r$, with $R \in \mathbb{R}^{s\times(|\mathcal{I}|+m)}$ of full row rank s, $r \in \mathbb{R}^s$ and suppose that there exists a (matrix) sequence $A_R = A_R(T) \in \mathbb{R}^{s\times s}$ such that $\lim_{T\to\infty} A_R R A_{IM} = R^* \in \mathbb{R}^{s\times(|\mathcal{I}|+m)}$ has full row rank s. Then, it holds under the null hypothesis for $T\to\infty$ that the Wald-type statistic:*

$$T_W := (R\hat{\vartheta} - r)' \left[R\hat{V}_{IM}R'\right]^{-1} (R\hat{\vartheta} - r) \xrightarrow{d} \mathcal{O}_s, \tag{17}$$

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

$$T_t := \frac{R\hat{\vartheta} - r}{\sqrt{R\hat{V}_{IM}R'}} \xrightarrow{d} \mathcal{Z}, \tag{18}$$

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

2.3 Fixed-b Inference

An advantage of the IM-OLS estimator compared to, e. g., the FM-OLS estimator of Phillips and Hansen (1990) or the D-OLS estimator of Saikkonen (1991), is the possibility to perform asymptotically pivotal fixed-b inference.[10] One key necessary ingredient for fixed-b test statistics is that the limiting distribution of the Wald-type statistic can be written as a functional of standard Brownian motions, i. e., of $W_v(r)$ and $w_{u\cdot v}(r)$. Whilst such a bijection between a functional of $B(r)$ and a functional of $W(r)$ exists by construction in cointegrating linear regressions, this is not necessarily so for cointegrating (multivariate) polynomial regressions. To illustrate the matter, consider the stochastic regressors, or their limiting processes, for the example of the two-factor Translog function given

[9] Long-run covariance estimation based on the OLS residuals – based on consistency of the OLS estimator of θ in (3) – can be shown to be consistent under the usual assumptions on kernel and bandwidth, e. g., those given in Jansson (2002), and, e. g., $\{\eta_t\}_{t\in\mathbb{Z}}$ fulfilling the exemplary assumptions given in Footnote 8.

[10] The dynamic least squares estimation principle, developed for cointegrating linear regressions in Phillips and Loretan (1991), Saikkonen (1991) and Stock and Watson (1993), has been extended to even more general settings than polynomial functions in, e. g., Saikkonen and Choi (2004).

in (9), ignoring, as in (1), the deterministic regressors for brevity since they are irrelevant for the argument:

$$\left[\begin{array}{c} B_{v_1}(r) \\ B_{v_2}(r) \\ \hline B_{v_1}(r)^2 \\ B_{v_2}(r)^2 \\ B_{v_1}(r)B_{v_2}(r) \end{array}\right] = \left[\begin{array}{cc|ccc} \tau_{11} & \tau_{12} & 0 & 0 & 0 \\ 0 & \tau_{22} & 0 & 0 & 0 \\ \hline 0 & 0 & \tau_{11}^2 & \tau_{12}^2 & 2\tau_{11}\tau_{12} \\ 0 & 0 & 0 & \tau_{22}^2 & 0 \\ 0 & 0 & 0 & \tau_{12}\tau_{22} & \tau_{11}\tau_{22} \end{array}\right] \left[\begin{array}{c} W_{v_1}(r) \\ W_{v_2}(r) \\ \hline W_{v_1}(r)^2 \\ W_{v_2}(r)^2 \\ W_{v_1}(r)W_{v_2}(r) \end{array}\right], \tag{19}$$

with:

$$\Omega_{vv}^{1/2} = \left[\begin{array}{cc} \tau_{11} & \tau_{12} \\ 0 & \tau_{22} \end{array}\right].$$

In this case, a bijection between a vector involving powers of $B_{v_j}(r)$, i. e., $Z(r)$, and powers of $W_{v_j}(r)$ prevails, because – and this is an obviously necessary condition – the corresponding vectors are of the same dimension and because of the assumption that Ω is positive definite which implies that the lower right 3-by-3 block of the matrix after the equality sign in (19) has full rank. In the case $x_{1t}x_{2t}$ were not included as a regressor, the vector $Z(r)$ would consist of only four elements, but – unless $B_{v_1}(r)$ and $B_{v_2}(r)$ were independent, and hence $\tau_{12} = 0$, this four-dimensional vector would be a transformation of the same five-dimensional vector as the rightmost term in (19). In this case, obviously, a bijection would not prevail. Situations in which a bijection between $Z(r)$ and a vector involving only functions of standard Brownian motions prevails are henceforth referred to as *full design.*[11] In the case of full design, the limiting distribution of the IM-OLS estimator can, therefore, be rewritten as a functional of $W_v(r)$ and $w_{u\cdot v}(r)$:

Corollary 1. *Suppose that the setting of Proposition 1 and full design prevail. Then the limiting distribution for $T \to \infty$ of the OLS estimator $\hat{\vartheta}$ is given by:*

$$\begin{aligned} A_{IM}^{-1}\left(\hat{\vartheta} - \vartheta^*\right) &\Rightarrow \omega_{u\cdot v}^{1/2}(\Pi')^{-1}\left(\int_0^1 g(s)g(s)'ds\right)^{-1}\int_0^1 g(s)w_{u\cdot v}(s)ds \\ &= \omega_{u\cdot v}^{1/2}(\Pi')^{-1}\left(\int_0^1 g(s)g(s)'ds\right)^{-1}\int_0^1 [G(1) - G(s)]dw_{u\cdot v}(s), \end{aligned} \tag{20}$$

where $Z_W(r) := \Pi_Z^{-1}Z(r)$, with $\Pi := diag(\Pi_Z, \Omega_{vv}^{1/2})$, and:

$$g(r) := \left[\begin{array}{c} \int_0^r Z_W(s)ds \\ W_v(r) \end{array}\right], \quad G(r) := \int_0^r g(s)ds, \tag{21}$$

with $Z_W(r)$ containing powers of products of r and (the components of) $W_v(r)$.

Remark 2. Clearly, any CMPR relationship can be *extended* to full design by including the regressors missing for full design, e. g., in the two-variable example given above, the mixed regressor $x_{1t}x_{2t}$ is the key regressor for full design. When the true CMPR relationship does not have full design, the possibility to perform fixed-b inference, therefore, comes at the cost of including superfluous regressors,

[11] Another example of full design, in addition to Translog-type functions with an arbitrary number of variables and not just two as in the example, are cointegrating polynomial regressions, without cross-product terms, where only one of the integrated regressors appears with powers larger than one, compare, e. g., Wagner and Hong (2016, Proposition 5).
The prevalence of full design in cointegrating polynomial regressions is a necessary condition also for FM-OLS-based (non-)cointegration tests with limiting distributions that allow for critical values to be tabulated, see Wagner (2023), and for IM-OLS-based cointegration testing see Grabarczyk and Wagner (2024) and the discussion in Section 2.4 below.

i. e., at the cost of estimating a too big regression model. Whilst this is, of course, asymptotically innocuous, it may impact finite sample performance.

For exactly the same reason as discussed in Vogelsang and Wagner (2014a, Lemma 2) for cointegrating linear regressions, it can be shown that it is not possible to perform asymptotically pivotal fixed-b inference using the IM-OLS residuals $\hat{S}_t^u$. This is due to non-vanishing and nuisance parameter-dependent correlation between the limiting distribution of $A_{\mathrm{IM}}^{-1}(\hat{\vartheta} - \vartheta^*)$ and the limit process of $T^{-1/2}\sum_{t=2}^{\lfloor rT \rfloor} \Delta\hat{S}_t^u$. Also exactly as in Vogelsang and Wagner (2014a), valid fixed-b inference can be based upon estimating $\omega_{u\cdot v}$ using a modification of the IM-OLS residuals, $\hat{S}_{t,\mathrm{M}}^u$ say, that are constructed as described next. First, define:

$$M_t := t\sum_{j=1}^{t} \tilde{S}_j^Z - \sum_{j=1}^{t-1}\sum_{s=1}^{j} \tilde{S}_s^Z, \tag{22}$$

stacked into the matrix $M := [M_1, \ldots, M_T]'$. In the second step, the matrix M is orthogonalized to $\tilde{S}^Z$, which yields $M^{\perp} := \left(I_T - \tilde{S}^Z(\tilde{S}^{Z\prime}\tilde{S}^z)^{-1}\tilde{S}^{Z\prime}\right)M$. The modified residuals are finally given by orthogonalizing the IM-OLS residuals $\hat{S}^u$ to $M^{\perp}$, i. e.,:

$$\hat{S}_{\mathrm{M}}^u := \left(I_T - M^{\perp}(M^{\perp\prime}M^{\perp})^{-1}M^{\perp\prime}\right)\hat{S}^u, \tag{23}$$

with $\hat{S}_{\mathrm{M}}^u = [\hat{S}_{1,\mathrm{M}}^u, \ldots, \hat{S}_{T,\mathrm{M}}^u]'$. Exactly as in Vogelsang and Wagner (2014a, Lemma 2) it can be shown that $T^{-1/2}\sum_{t=2}^{\lfloor rT \rfloor} \Delta S_{t,\mathrm{M}}^u$ is, conditional upon $W_v(r)$, asymptotically independent of the limiting distribution of $A_{\mathrm{IM}}^{-1}\left(\hat{\vartheta} - \vartheta^*\right)$. Consequently, using a long-run variance estimator:

$$\hat{\omega}_{u\cdot v,\mathrm{M}} := T^{-1}\sum_{i=2}^{T}\sum_{j=2}^{T} k\left(\frac{|i-j|}{B}\right)\Delta\hat{S}_{i,\mathrm{M}}^u\Delta\hat{S}_{j,\mathrm{M}}^u, \tag{24}$$

with kernel function $k(\cdot)$ and bandwidth B, with $B = bT$ and $0 < b \leq 1$, allows for asymptotically pivotal fixed-b inference for Wald-type hypothesis testing as discussed for standard asymptotic inference in Proposition 2, but with $\hat{\omega}_{u\cdot v,\mathrm{M}}$ in place of (a consistent long-run variance estimator) $\hat{\omega}_{u\cdot v}$.

In order to efficiently describe the fixed-b limiting distributions of the Wald-type test statistic define for a stochastic process $P(r)$ the random variable $Q(P)$ as follows:[12] In the case $k(\cdot)$ is such that $k(0) = 1$ and $k(\cdot)$ is twice continuously differentiable with first and second derivatives given by $k'(\cdot)$ and $k''(\cdot)$ define:

$$\begin{aligned} Q(P) \;:=\; & -\frac{1}{b^2}\int_0^1\int_0^1 k''\left(\frac{|r-s|}{b}\right)P(s)P(r)'dsdr + \frac{1}{b}\int_0^1 k'\left(\frac{|1-s|}{b}\right)(P(1)P(s)' + P(s)P(1)')ds \\ & + P(1)P(1)'. \end{aligned} \tag{25}$$

The above case covers, e. g., the Quadratic Spectral (QS) kernel. The second case considered covers the Bartlett kernel (with $k(x) = 1 - |x|$ for $|x| \leq 1$ and 0 otherwise), where we define $Q(P)$ as:

$$\begin{aligned} Q(P) \;:=\; & \frac{2}{b}\int_0^1 P(s)P(s)'ds - \frac{1}{b}\int_0^{1-b}(P(s)P(s+b)' + P(s+b)P(s)')ds \\ & -\frac{1}{b}\int_{1-b}^1 (P(1)P(s)' + P(s)P(1)')ds + P(1)P(1)'. \end{aligned} \tag{26}$$

[12] In our case, with y_t and u_t univariate, also the process $P(r)$, see (28) below, is univariate. We nevertheless use the general formulation here, since the result extends analogously to situations where long-run covariance matrices are considered and $P(r)$ is indeed a vector.

Proposition 3. *Let the data be generated by* (3) *with the assumptions discussed in Subsection 2.1 in place. Assume that the augmented regression* (10) *has full design. Consider s linearly independent linear restrictions on θ collected in $H_0 : R\vartheta = R_\theta\theta = r$, with $R \in \mathbb{R}^{s\times(|\mathcal{I}|+m)}$ of full row rank s, $r \in \mathbb{R}^s$ and suppose that there exists a (matrix) sequence $A_R \in \mathbb{R}^{s\times s}$ such that $\lim_{T\to\infty} A_R R A_{IM} = R^* \in \mathbb{R}^{s\times(|\mathcal{I}|+m)}$ has full row rank s. Furthermore, $\hat{\omega}_{u\cdot v,M}$ is as given in* (24) *with $B = bT$ and $0 < b \leq 1$ held fixed. Then, it holds under the null hypothesis for $T \to \infty$ that the fixed-b Wald-type test statistic:*

$$T_{W,b} := (R\hat{\vartheta} - r)' \left[R\hat{V}_{IM,M}R'\right]^{-1} (R\hat{\vartheta} - r) \Rightarrow \frac{\mathcal{O}_s}{Q(P)}, \tag{27}$$

with $\hat{V}_{IM,M}$ defined similarly as $\hat{V}_{IM}$ in (16)*, but with $\hat{\omega}_{u\cdot v}$ replaced by $\hat{\omega}_{u\cdot v,M}$ and:*

$$P(r) := \int_0^r dw_{u\cdot v}(s) - h(r)' \left(\int_0^1 h(s)h(s)'ds\right)^{-1} \int_0^1 [H(1) - H(s)]dw_{u\cdot v}(s), \tag{28}$$

where:

$$h(r) := \begin{bmatrix} g(r) \\ \int_0^1 [G(1) - G(s)]ds \end{bmatrix}, \quad H(r) := \int_0^r h(s)ds, \tag{29}$$

with $\mathcal{O}_s$ denoting a chi-squared distributed random variable with s degrees of freedom that is independent of $Q(P)$.

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

$$T_{t,b} := \frac{R\hat{\vartheta} - r}{\sqrt{R\hat{V}_{IM,M}R'}} \Rightarrow \frac{\mathcal{Z}}{\sqrt{Q(P)}}, \tag{30}$$

with $\mathcal{Z}$ denoting a standard normally distributed random variable that is independent of $Q(P)$.

The precise form of $Q(P)$ depends upon the kernel and bandwidth chosen and is given by (25) *or* (26) *if the kernel satisfies the respective assumptions.*

Remark 3. As an immediate consequence of the Frisch-Waugh theorem or partitioned regression, it is possible to combine IM-OLS parameter estimation and the calculation of the modified residuals $\hat{S}^u_{t,\mathrm{M}}$ into one regression by performing OLS estimation of the following regression:

$$S^y_t = \tilde{S}^{Z\prime}_t \vartheta + M^{\perp\prime}_t \kappa + S^u_t, \tag{31}$$

with $M^{\perp\prime}$ the t-th row of the matrix $M^\perp$ as defined above (23). This leads on the one hand to an identical estimator of $\hat{\vartheta}$ as defined in (12) and to residuals identical to $\hat{S}^u_{t,\mathrm{M}}$ as defined in (23) on the other hand.

2.4 Specification and Cointegration Testing

As already mentioned in the introduction, one important application of the results in this paper is – in addition to their allowing one to estimate, e. g., Translog-type cointegrating relationships as a special case of the CMPR model – that they allow one to perform RESET-type specification tests (see Ramsey, 1969) in the variant of Thursby and Schmidt (1977), i. e., by performing RESET-type specification testing based on including powers and, in general, cross-products of powers of the regressors of the original regression in an augmented test regression.

For the CMPR setting considered, in fact for any cointegrating regression setting (considered in original and not in partial sum format), the augmented regression version of RESET is more convenient to deal with than the original formulation of Ramsey (1969), see also Keenan (1985) or Tsay (1986), where the residuals of the original model are regressed on the above-mentioned additional, auxiliary regressors in an auxiliary regression. This stems from the fact that – in cointegrating regression settings in general – the OLS residuals (more importantly, the corresponding partial sum limit process) feature nuisance parameter dependencies emanating from error serial correlation and regressor endogeneity; the very reasons for the development of modified least squares estimators that allow for asymptotic standard inference in cointegrating regressions.

These complications are clearly seen in, e. g., Hong and Phillips (2010), who develop a modified RESET test for the null hypothesis of a cointegrating linear regression with only one integrated regressor based on an auxiliary regression of the OLS residuals on powers of the integrated regressor. To achieve asymptotic chi-squared inference, the corresponding test statistic has to be adjusted to remove the nuisance parameter dependencies introduced by serial correlation and endogeneity. Wagner and Hong (2016, Section 2.3, Proposition 4) also consider a RESET-type specification test based on an auxiliary regression of the FM-OLS residuals of a cointegrating polynomial regression (without cross-products of integrated regressors) on additional deterministic regressors, higher-order powers of the included integrated regressors as well as additional (non-cointegrated) integrated regressors and their powers.[13] Using FM-OLS residuals rather than OLS residuals, however, only changes the precise form of the adjustments necessary to achieve asymptotic chi-squared inference and does not directly lead to a nuisance parameter-free test statistic. Similar in spirit, Wagner and Hong (2016, Section 2.3, Proposition 3) consider a Wald-type specification test based on an augmented regression, i. e., by extending the original model with the just mentioned auxiliary regressors. This latter test can *almost* be seen as an extension of the Thursby and Schmidt (1977) version of a RESET test to the CPR setting – *almost* because cross-products of the original regressors cannot be handled in the augmented regression.

The IM-OLS estimator, as discussed in this paper, allows one to *most straightforwardly* include also cross-products in augmented regressions in the CMPR setting. This, if nothing else, means that the resultant test options capture the original RESET spirit. To fix notation consider:

$$\begin{aligned} y_t &= \sum_{i=1}^{|\mathcal{I}|} z_{it}\theta_i + \sum_{j=1}^{|\mathcal{J}|} z_{jt}\theta_j + u_t \\ &= Z_t'\theta + Z_t^{*\prime}\theta^* = Z_t^{\mathrm{A}\prime}\theta^{\mathrm{A}} + u_t, \end{aligned} \tag{32}$$

with z_{it}, θ_i and $\mathcal{I}$ as in (3) and auxiliary regressors $z_{jt} = t^{j_0}x_{1t}^{j_1}\cdots x_{mt}^{j_m}x_{m+1,t}^{j_{m+1}}\cdots x_{m+m^*,t}^{j_{m+m*}}$ a vector of products of powers of t, the original integrated regressors $x_{1t},\dots,x_{mt}$ and m^* additional integrated regressors $x_{m+1,t},\dots,x_{m+m^*,t}$, with $j_r \in \mathbb{N}_0$ for $r = 0,\dots,m+m^*$, ordered again, e. g., lexicographically over the multi-index set $\mathcal{J}$. This defines $Z_t^{\mathrm{A}} := [Z_t', Z_t^{*\prime}]'$, with Z_t as before, $Z_t^* := [z_{1t}^*,\dots,z_{|\mathcal{J}|t}]'$ and $\theta^{\mathrm{A}} := [\theta', \theta^{*\prime}]'$. The joint vector of integrated regressors $x_t^{\mathrm{A}} := [x_t', x_t^{*\prime}]'$ is assumed to be non-cointegrated. In brief, the same set of assumptions assumed to hold for (3) have to hold for the augmented CMPR model (32). It is, e. g., obviously necessary to ensure that no element in Z_t^* is al-

[13]The idea of including superfluous higher-order deterministic trends in unit root or cointegration tests has been explored originally in Park (1990) and Park and Choi (1988). It is, of course, based on the fact that in the case of alternatives that involve some elements of nonstationary errors in a "spurious" regression model, the deterministic trends will pick up some of the asymptotic dependence between deterministic trends and functions (e. g., polynomials) of Brownian motions. A similar argument, directly in the heart of spurious regression arguments, holds for the inclusion of additional integrated regressors and their powers in augmented regressions. Note that it is an option to consider simulated integrated variables as additional integrated (test) regressors. Simulated regressors are very convenient as they do, by construction, not add any problems with respect to cointegration amongst the integrated regressors.

ready an element in Z_t. The corresponding augmented partial sum regression is, entirely analogously to (11), given by:

$$\begin{aligned} S_t^y &= S_t^{Z^{\mathrm{A}}\prime}\theta^{\mathrm{A}} + x_t^{\mathrm{A}\prime}\gamma^{\mathrm{A}} + S_t^u \\ &= \tilde{S}_t^{Z^{\mathrm{A}}\prime}\vartheta^{\mathrm{A}} + S_t^u. \end{aligned} \tag{33}$$

The RESET test now corresponds to the block-zero restriction $\theta^* = 0$, i. e., $R\vartheta^{\mathrm{A}} = 0$, with $R = [0,\ I_{|\mathcal{J}|},\ 0]$ and $r = 0$, tested based on OLS estimation of (33). It can be seen immediately that this matrix satisfies the constraints on restriction matrices discussed above Proposition 2 without further ado, as each of the $|\mathcal{J}|$ restrictions involves only a single coefficient. Also, of course, given the simple block-zero nature of the restrictions, the Wald-type RESET test statistic is of the simple form:

$$T_{W,\mathrm{RESET}} := \hat{\theta}^{*\prime}\hat{V}_{\mathrm{IM}}(\hat{\theta}^*)^{-1}\hat{\theta}^*, \tag{34}$$

with $\hat{V}_{\mathrm{IM}}(\hat{\theta}^*) = R\hat{V}_{\mathrm{IM}}(\hat{\theta}^*)R'$, the block of $\hat{V}_{\mathrm{IM}}$ corresponding to $\hat{\theta}^*$.

It is also straightforwardly possible to consider an LM-type auxiliary regression version of the RESET test, which differs from the Wald-type version only in terms of the used (conditional long-run) variance estimator. This stems from the fact that IM-OLS estimation is, as mentioned several times, an OLS regression (see, e. g., Engle, 1984), i. e., it follows immediately from the Frisch-Waugh theorem or partitioned regression. More precisely, an LM-type, i. e., residual-based, version of the RESET test can be based on the coefficient estimator of regressing the residuals of (11), $\hat{S}^u_{t,(11)}$ say, on the auxiliary regressors, i. e., on the partial sum version of the additional regressors z_{jt}, $j = 1, \ldots, |\mathcal{J}|$, $S_t^{Z^*}$ say, and on $x_t^* := [x_{m+1,t}, \ldots, x_{m+m^*,t}]'$:

$$\hat{S}^u_{t,(11)} = S_t^{\perp,Z^*\prime}\theta^* + x_t^{\perp,*\prime}\gamma^* + S_t^u, \tag{35}$$

and testing the null hypothesis $H_0 : \theta^* = 0$ based on OLS estimation of (35). Note that (35) does not contain $S_t^{Z^*}$ and x_t^* as regressors, but the orthogonalized, with respect to the regressors in (11), versions of these regressors, constructed exactly as $M_t^{\perp}$ above. This orthogonalization step is common in LM-type specification testing in cointegrating regressions (see, e. g., Propositions 3 and 4 in Wagner and Hong, 2016) and this orthogonalization implies that the IM-OLS estimator of θ^* in (35) is identical to the IM-OLS estimator of θ^* from the augmented regression (32). If one were to use $S_t^{Z^*}$ and x_t^* in (35), the resultant estimator of θ^* would, in general, not be identical. Furthermore, equivalence also holds true – *up to* the scalar long-run variance $\omega_{u\cdot v}$ – for the corresponding covariance matrix estimators $\hat{V}_{\mathrm{IM}}(\hat{\theta}^*)$. Here, the only difference is the estimator $\hat{\omega}_{u\cdot v}$ used: For the Wald-type RESET test the estimator is based on the OLS residuals of (32), whereas for the LM-type RESET test the estimator is based on the OLS residuals of (3).

The fact that we have established that the only difference between Wald- and LM-type RESET test statistics is the estimator of the scalar conditional long-run variance $\omega_{u\cdot v}$ implies that for fixed-b inference there are *no differences* between these two variants, i. e., there is only one fixed-b statistic: First, it is an immediate consequence of the above-discussed equivalence of the parameter estimators that the residuals also coincide exactly, i. e., both regressions, the augmented regression (33) and the auxiliary regression (35), lead to exactly the same OLS residuals $\hat{S}_t^u$. Second, to achieve asymptotic independence of numerator and denominator of the fixed-b test statistic, the construction of the regressors M_t, compare (22), has to be based on the full set of all regressors, i. e., it has to include $S_t^{Z^*}$ and x_t^*. Therefore, there is only one set of modified residuals $\hat{S}^u_{t,\mathrm{M}}$ and consequently only one estimator $\hat{\omega}_{u\cdot v,\mathrm{M}}$ and only one fixed-b RESET test statistics.

Note that one can use, as a side-product, the construction of the set of augmented regressors Z_t^* to ensure full design of the augmented regression model to allow for fixed-b inference. The fixed-b critical values depend upon the specification of both the null model and the augmented regression, in addition to the kernel and the bandwidth sample size ratio b. Code and critical values for a variety of specifications are available as supplementary material. Of course, the code also provides an immediate starting point to adapt to any CMPR model not covered by the available critical values a researcher faces, given that full design prevails.

Remark 4. In the spirit of Remark 1 it may be convenient to "organize" the RESET test with a focus on the null hypothesis being a cointegrating linear regression, leading to the auxiliary regression model being, similar in spirit to (7), *(re-)ordered* with the deterministic regressors D_t and the integrated regressors x_t singled out:

$$y_t = D_t'\theta_D + x_t'\theta_x + \sum_{j\in\mathcal{J}^*} z_{jt}\theta_j + u_t. \tag{36}$$

Of course, this is just a re-ordering of the above equation (32) for a specific – but likely very important – null hypothesis of a cointegrating linear relationship between y_t and x_t. The finite sample performance analysis considered in the following Section 3 also considers cointegrating linear regressions under the null hypothesis, see (38) below. Tailor-made code corresponding to the null hypothesis of cointegrating linear regression is available as supplementary material, for both standard and fixed-b inference. In particular, we provide tables (and code) for cointegrating linear regressions with up to four integrated regressors, no deterministic components, intercept or intercept and linear trend, for the quadratic ($q = 2$) and cubic ($q = 3$) specification of the augmented regression for five kernels (Bartlett, Bohman, Daniell, Parzen and Quadratic Spectral) for the grid of 50 values of $b \in \{0.02, 0.04, \ldots, 0.98, 1\}$.

With respect to cointegration testing, let us only briefly mention that Grabarczyk and Wagner (2024, Proposition 3) present the extension of a Shin (1994)-type cointegration test to the CMPR setting considered here. The test statistic is closely related to the extension of Shin (1994) considered in Wagner and Hong (2016, Propositions 5 and 6) to CPR models without cross-products. The main conceptual difference between the two variants of the test is that Wagner and Hong (2016) consider FM-OLS estimation and, therefore, construct their test using the FM-OLS residuals. The test statistic based on the IM-OLS estimator has to be based – because of partial summation – on the first differences of the IM-OLS residuals and is consequently defined as:

$$CT_{\mathrm{IM}} := \frac{1}{T^2\hat{\omega}_{u\cdot v}} \sum_{t=2}^{T} \left(\sum_{i=2}^{t} \Delta\hat{S}_i^u \right)^2. \tag{37}$$

When using a consistent estimator $\hat{\omega}_{u\cdot v}$ of $\omega_{u\cdot v}$, the CT_{IM} test statistic converges under the null hypothesis *in the case of full design* towards a limiting distribution that can be tabulated, with the critical values depending upon the specification of the CMPR model. Grabarczyk and Wagner (2024) provide critical values for the CT_{IM} statistic for specifications relevant for the EKC literature, a small simulation study, as well as code to use the CT_{IM} test (or to generate critical values for other full-design specifications a researcher might be interested in).[14]

[14] Grabarczyk and Wagner (2024) also provide critical values for fixed-b hypothesis tests for the same set of specifications. Furthermore, Veldhuis and Wagner (2024) consider a fixed-b version of the CT_{IM} statistic, defined similarly as CT_{IM} above, but with $\hat{\omega}_{u\cdot v,\mathrm{M}}$ in place of a consistent estimator $\hat{\omega}_{u\cdot v}$. This test extends the fixed-b Kwiatkowski *et al.* (1992)-type test of Amsler *et al.* (2009) from a stationarity test to a cointegration test for the CMPR setting also considered here.

3 RESET: Finite Sample Performance

To keep the paper short, we consider in the simulation section only the performance of the RESET specification test. Grabarczyk and Wagner (2024, Section 3) present a simulation study that compares IM-OLS with both, fully modified OLS (FM-OLS) and dynamic OLS (D-OLS) for a (simple) quadratic cointegrating polynomial regression model, which is the workhorse model in the environmental Kuznets curve (EKC) literature (see, e. g., Wagner, 2015). Whilst their simulation study does not cover the width of the possible CMPR specifications that can be handled with IM-OLS, it suffices to give an impression of the relative performance of the three estimators and tests based upon them in a setting for which all three estimators have been developed. The findings are qualitatively very similar to the finite sample performance findings reported in Vogelsang and Wagner (2014a) for cointegrating linear regressions.

To assess the performance of the RESET test we consider a data generating process (DGP) corresponding to (the most widely-used setting in the cointegration literature of) a cointegrating linear regression under the null hypothesis:

$$\begin{aligned} y_t &= c + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t, \\ x_{it} &= x_{i,t-1} + v_{it}, \quad x_{i0} = 0, \qquad i = 1, 2, \end{aligned} \tag{38}$$

where:

$$\begin{aligned} u_t &= \rho_1 u_{t-1} + \varepsilon_t + \rho_2(e_{1t} + e_{2t}), \quad u_0 = 0, \\ v_{it} &= e_{it} + 0.5 e_{i,t-1}, \quad i = 1, 2, \end{aligned} \tag{39}$$

with ε_t, e_{1t} and e_{2t} are iid standard normal random variables independent of each other. The parameter values chosen are $c = 3$, $\beta_1 = \beta_2 = 1$, where we note that the values of these parameters have no effect on the results because the IM-OLS estimator and a fortiori the RESET test are exactly invariant to the values of c, β_1 and β_2. The values for ρ_1 and ρ_2 are chosen from the set $\{0, 0.3, 0.6, 0.8\}$. Parameter ρ_1 controls serial correlation in the regression error process $\{u_t\}_{t\in\mathbb{Z}}$, whereas parameter ρ_2 controls whether the regressors are endogenous or not. The kernels chosen for long-run covariance estimation are the Bartlett and the Quadratic Spectral (QS) kernels and the results are reported for bandwidths using the grid $B = bT$ with $b \in \{0.02, 0.04, ..., 0.98, 1.0\}$ or for data-dependent bandwidths chosen according to Andrews (1991), labelled AND, and Newey and West (1994), labelled NW in the tables. We report a selection of representative results for sample sizes $T = 100, 200, 500$. The number of replications is 10,000 in all cases and all tests are carried out at the 5% nominal significance level.

For the discussion of the simulation results we label T_W as IM(O) and $T_{W,b}$ as IM(fb). We implement the IM(O) and IM(fb) RESET tests via estimation of the quadratic and cubic specifications given by the regression models (which are partially summed and augmented by the original x_{1t} and x_{2t} for IM-OLS estimation):

$$y_t = c + x_{1t}\beta_1 + x_{2t}\beta_2 + x_{1t}^2\theta_{0,2,0} + x_{2t}^2\theta_{0,0,2} + x_{1t}x_{2t}\theta_{0,1,1} + u_t \quad (q = 2), \tag{40}$$

$$\begin{aligned} y_t = {} & c + x_{1t}\beta_1 + x_{2t}\beta_2 + x_{1t}^2\theta_{0,2,0} + x_{2t}^2\theta_{0,0,2} + x_{1t}x_{2t}\theta_{0,1,1} + x_{1t}^3\theta_{0,3,0} + x_{2t}^3\theta_{0,0,3} \\ & + x_{1t}^2 x_{2t}\theta_{0,2,1} + x_{1t}x_{2t}^2\theta_{0,1,2} + u_t \quad (q = 3), \end{aligned} \tag{41}$$

where we test the null hypothesis that the θ parameters are jointly equal to zero in the two specifications using the IM-OLS estimates. These augmented regressions do not correspond to the DGP, neither under the null of linearity nor – with one exception – for the alternatives considered. The IM(O) tests are based on using the OLS residuals of the above augmented regressions (not partially summed) to obtain a consistent estimator $\hat{\omega}_{u\cdot v}$ of $\omega_{u\cdot v}$. The IM(fb) tests are based on the fixed-b

asymptotic results given in Proposition 3 and use the modified residuals $\Delta\hat{S}^u_{t,\mathrm{M}}$ for estimating $\omega_{u\cdot v}$ by $\hat{\omega}_{u\cdot v,\mathrm{M}}$ as given in (24).

We first present null-rejection probabilities for IM(fb) for a grid of 50 bandwidth values. These null-rejection probabilities are computed using the relevant fixed-b critical values depending on the kernel, the 50 equidistant values of the bandwidth sample size ratio b and q. Figures 1 and 2 plot null-rejection probabilities for the Bartlett and QS kernels, respectively, for $q = 2$ and $T = 100$. Each figure depicts results for our range of ρ_1 and ρ_2 values with $\rho_1 = \rho_2$ in all cases. In Figures 1 and 2 we see that null rejections are close to 0.05 for all bandwidths when $\rho_1 = \rho_2 = 0$. This shows that the fixed-b critical values are successful in capturing the impact of the kernel and bandwidth choices on the finite sample behavior of IM(fb) when there is no serial correlation. As ρ_1, ρ_2 increase, we see that over-rejections appear. Over-rejections are more substantial for the Bartlett kernel than for the QS kernel. Increasing the bandwidth helps to reduce over-rejection problems for the QS kernel but has relatively little effect in the Bartlett kernel case.

Figures 3 and 4 depict null rejections for all three sample sizes and both values of q for $\rho_1 = \rho_2 = 0.3$. Increasing the sample size reduces the over-rejection problem. For a given value of T, increasing q from 2 to 3 tends to, as expected, inflate the over-rejection problem especially when $T = 100$.[15] Overall, the QS kernel delivers a test with null rejections closer to 0.05 than the Bartlett kernel.

Figures 5 and 6 have the same configuration as Figures 3 and 4, but with $\rho_1 = \rho_2 = 0.8$. The Bartlett kernel leads to a test that has severe over-rejection problems when $T = 100$, but the tendency to over-reject quickly falls as T increases. The QS kernel leads to a test with substantially less over-rejection problems especially if T is not small and/or the bandwidth is not too small. When $T = 500$, the QS test is much less sensitive to q than the Bartlett kernel test. Clearly, the QS kernel has some advantages over the Bartlett kernel regarding size control when using IM(fb).

Table 1 reports empirical null rejections for IM(O) and IM(fb) when the data-dependent bandwidths of Andrews (1991) and Newey and West (1994) are used. The IM(O) statistic uses chi-squared critical values whereas IM(fb) uses fixed-b critical values. When there is no serial correlation, IM(fb) has rejections close to 0.05 in nearly all cases. In contrast, IM(O) has over-rejections that tend to be higher for the QS kernel and $q = 3$. These over-rejections fall as T increases. As ρ_1, ρ_2 increase, both tests tend to have larger over-rejections for a given sample size. When $\rho_1, \rho_2 = 0.8$, IM(fb) can have substantial over-rejections that are larger than the IM(O) over-rejections. This is because the data dependent bandwidths tend to be relatively small, thereby not fully exploiting the potential of fixed-b inference to reduce the extent of over-rejections by using larger bandwidths. That fixed-b inference, in particular in conjunction with the QS kernel, reduces over-rejections for larger bandwidths is, e.g., indicated in Figure 6.[16]

Overall, both IM(O) and IM(fb) can have over-rejection problems when the serial correlation is strong enough relative to the sample size. The QS kernel leads to less over-rejection problems than the Bartlett kernel and increasing q tends to inflate over-rejection problems. When the sample size is not too small and the serial correlation is not too strong, both IM(O) and especially IM(fb) have null rejections relatively close to 0.05.

[15] Sun (2014) shows, using higher-order asymptotic theory, that finite sample size distortions of tests based on kernel HAC estimators increase as the dimension of the null being tested increases when chi-squared critical values are used for the test. Whilst the results in Sun (2014) do not apply to models with cointegration, our simulations suggest that Sun's results might extend to the cointegration case, although this would most likely be very challenging to establish formally.

[16] As argued by Sun *et al.* (2008), bandwidth rules designed to balance size distortions and power of the tests would be preferable to using bandwidth rules that minimize an approximate MSE of the HAC estimator as in Andrews (1991) and Newey and West (1994). Such testing-oriented bandwidths tend to be larger than MSE-based bandwidths. Unfortunately, the theory of Sun *et al.* (2008) does not apply in our setting because of the integrated regressors. It would be very challenging but worthwhile to extend the Sun *et al.* (2008) approach to cointegration settings.

Next, we present some simulations that illustrate size-adjusted power of the tests.[17] The DGP under the alternative is specified as:

$$y_t = c + \beta_1 x_{1t} + \beta_2 x_{2t} + \phi G(X_t) + u_t, \tag{42}$$

where $G(X_t)$ takes on six possible functional forms: i) x_{1t}^2, ii) $x_{1t}^2 + x_{1t}x_{2t}$, iii) $x_{1t}^2 + x_{1t}x_{2t} + x_{2t}^2$, iv) $x_{1t}x_{2t}$, v) x_{1t}^3 and vi) $x_{1t}\left(1 + e^{-x_{1t}}\right)^{-1}$. The DGPs for x_{1t},x_{2t} and u_t are the same as in (38). Note that alternative iii) corresponds exactly to the augmented regression for $q = 2$. Alternative vi) corresponds to a cointegrating smooth transition regression, as, e.g., considered by Saikkonen and Choi (2004). Table 2 presents power of the IM(O) test for sample size $T = 200$. Null-rejection probabilities are also reported in the tables to put the power results in context. For each alternative, power is reported for the same values of ρ_1, ρ_2 as used for the size simulations and a given value of ϕ chosen so that power is nontrivial for the sample size considered. Table 3 reports power for the IM(fb) test and has the same format as Table 2. The two data-dependent bandwidths are used in both tables.

Several general patterns are evident in Tables 2 and 3. First, power is decreasing in ρ_1, ρ_2, which is expected. Second, power is sometimes higher with $q = 2$ compared to $q = 3$, but the opposite is also often true. Because we are not holding bandwidths constant across values of q, since they are data dependent to mimic applications, we cannot easily disentangle the effect of changing q on power. Third, power is similar between IM(O) and IM(fb), with IM(O) having slightly higher power in some cases. As expected, power is highest for alternative iii), with $G(X_t) = x_{1t}^2 + x_{2t}^2 + x_{1t}x_{2t}$ corresponding exactly to the augmented regression for $q = 2$. Changing $q = 2$ to $q = 3$ in this case entails relatively little power loss with $T = 200$, thus illustrating the minimal impact of the degrees-of-freedom loss incurred by including four extra regressors.

Whilst Tables 2 and 3 have useful information, they give no indication of the shape of the power curves. Figures 7 to 10 plot power for various configurations of the tests for the alternative with $G(X_t) = x_{1t}^2$ for an equidistant grid of 21 values of ϕ ranging from 0 (the null hypothesis) to 0.04. Figures 7 and 8 depict power of the Andrews (1991) data-dependent bandwidth versions of the tests for the Bartlett and QS kernels respectively for our range of values for ρ_1, ρ_2. We see that the entire power curves shift down as ρ_1, ρ_2 increase. Power is similar across kernels and is minimally higher for IM(O) than IM(fb).

Figures 9 and 10 plot power for IM(fb) for a selection of values for the bandwidth sample size ratio, b. Figure 9 gives power for the Bartlett kernel and shows that for this kernel power is not that sensitive to the choice of bandwidth. In contrast, Figure 10 shows that power for the QS kernel is very sensitive to b and power curves shift down as b increases. Recall that the over-rejection problem was smallest when using the QS kernel with large values of b. We see that there is a tradeoff between reduction in size distortions and power when choosing the bandwidth for IM(fb), especially for the QS kernel.

Tables 2 and 3 show that the tests also have power to detect departures from linearity given by the logistic function, $G(X_t) = x_{1t}\left(1 + e^{-x_{1t}}\right)^{-1}$. Whilst the logistic function departs slowly from linearity, the polynomial terms in the augmented regressions used to compute IM(O) and IM(fb) are able to detect that the cointegrating relationship is not linear. To give a better sense of the shape of the power functions in this case, Figures 11 and 12 plot power for the logistic alternative for an equidistant grid of 21 values of ϕ ranging from 0 (the null hypothesis) to 1. In the figures, results are reported for both $q = 2$ and $q = 3$, the Andrews (1991) data-dependent bandwidth and only the Bartlett kernel, given that power is similar for the QS kernel. As the figures show, power increases as ϕ increases, but power tends to flatten out for large values of ϕ. The fact that power is lower and increases not as fast for the logistic alternative compared to the polynomial alternatives stems from the fact that

[17] Throughout the paper, all power considerations refer to size-adjusted power, for notational brevity simply referred to as power.

larger values of ϕ do not generate as large and fluctuating regressors in the logistic case as in the polynomial case. Comparing Figures 11 and 12, we see that increasing q increases power, especially when ρ_1, ρ_2 are not too large. This reflects the fact that the logistic function is better approximated by a higher-order polynomial and the effect of the better approximation outweighs the loss of degrees of freedom throughout our simulations in the logistic case.

Our simulation results suggest the following about the practical performance of the IM-OLS RESET test for detecting departures from linear cointegration. First, if the sample size is large enough relative to the strength of endogeneity and serial correlation, then both tests, IM(O) and IM(fb), have null rejections not too far from the nominal level when data-dependent bandwidths are used. In addition, the IM(fb) statistic has stable null rejections for the full range of bandwidths especially when the QS kernel is used. Second, if the sample size is small relative to the strength of serial correlation and endogeneity, the tests can have over-rejection problems under the null that are sometimes severe. If the QS kernel is used, then increasing the bandwidth of IM(fb) can substantially reduce over-rejection problems. This is less true for the Bartlett kernel. Third, the tests have respectable power in detecting nonlinearities in the cointegrating relationship, especially, by construction, when the nonlinearity is a polynomial in the regressors. In addition, the tests do have power to also detect smooth, gradual departures from linearity as in the case of a logistic alternative and, in such a case, increasing the order of the polynomial in the augmented regression can increase power. Finally, increasing the bandwidth has relatively little effect on power of IM(fb) when the Bartlett kernel is used but causes power to drop substantially when the QS kernel is used. That increasing the bandwidth when using the QS kernel reduces size distortions at the expense of power is a typical finding in the fixed-b literature, also found in cointegrating linear regression settings by Vogelsang and Wagner (2014a).

4 Summary and Conclusions

The paper has shown that the simple idea underlying the IM-OLS estimator of Vogelsang and Wagner (2014a), which uses partial summation cum augmentation with the original integrated regressors, allows to immediately extend the scope of IM-OLS estimation from cointegrating linear regression to cointegrating multivariate polynomial regressions (CMPRs). Compared to cointegrating polynomial regressions (CPRs) as considered in, e. g., Wagner and Hong (2016), CMPRs allow to include arbitrary products of (non-negative) integer powers of integrated regressors (and time). The CMPR setting, therefore, overcomes, admittedly only for polynomial functions, the almost omnipresent additively separable (between integrated regressors) setting used in the nonlinear cointegration literature, which in very large parts actually only considers one integrated regressor.

One important (simple) special case of CMPRs, not covered, e. g., in Wagner and Hong (2016), are Translog-type (cost, demand or production) functions. The second important illustration of the applicability of the IM-OLS estimator in the CMPR setting is that a RESET specification test for testing the null hypothesis of a CMPR – with important special cases being cointegrating linear regressions, cointegrating polynomial regressions and Translog-type cointegrating relationships – follows directly. For the special case of *full design* of the CMPR relationship, we have furthermore developed fixed-b hypothesis tests. Fixed-b inference allows to capture the impact of kernel and bandwidth choices, required for the estimation of a scalar conditional long-run variance parameter, on the sampling distributions of test statistics. As in the cointegrating linear case, the IM-OLS residuals themselves cannot be used to construct asymptotically pivotal fixed-b test statistics and, instead, a specifically modified version of the IM-OLS residuals has to be used.

The finite-sample RESET test simulations show that the fixed-b asymptotic distribution adequately captures the impact of kernel and bandwidth choices on the test statistics. Comparing the standard with the fixed-b test statistic shows that fixed-b asymptotic theory trades partly substantially smaller size distortions for marginally lower size-adjusted power. This is a typical finding in the fixed-b literature. As is also well known in the fixed-b literature the size/power tradeoff as a function of bandwidth sample size ratio b differs substantially between the Bartlett and the QS kernels, with the latter being more sensitive in this respect. The tests do not only have power against the polynomial alternatives simulated but also against smoothly varying logistic alternatives. It has to be noted that commonly used data-dependent bandwidth rules, like those of Andrews (1991) or Newey and West (1994), which typically lead to relatively small bandwidths are not necessarily optimal for (fixed-b) inference in terms of size/power tradeoffs. Developing in some sense optimal bandwidth rules for fixed-b inference in a cointegration framework thus appears to be an important but challenging problem.

Acknowledgements

The authors gratefully acknowledge financial support from the Jubiläumsfonds (Anniversary Fund) of the Oesterreichische Nationalbank (Austrian Central Bank) under grant no. 15334. The second author, furthermore, gratefully acknowledges financial support from Deutsche Forschungsgemeinschaft via the Collaborative Research Center 823: *Statistical Modelling of Nonlinear Dynamic Processes*. Both authors thank Oliver Stypka for superb research assistance. This paper was completed 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 those of the authors 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

Amsler, C., Schmidt, P. and T. J. Vogelsang (2009). The KPSS test using fixed-b critical values: Size and power in highly autocorrelated time series. *Journal of Time Series Econometrics* **1**, Article 5.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**, 817–854.

Balke, N. S. and T. B. Fomby (1997). Threshold cointegration. *International Economic Review* **38**, 627–645.

Bec, F. and A. Rahbek (2004). Vector equilibrium error correction models with non-linear discontinuous adjustments. *Econometrics Journal* **7**, 628–651.

Bunzel, H. (2006). Fixed-b asymptotics in single-equation cointegration models with endogenous regressors *Econometric Theory* **22**, 743–755.

Chan, N. and Q. Wang (2015). Nonlinear regressions with nonstationary time series. *Journal of Econometrics* **185**, 182–195.

Chang, Y. and J. Y. Park (2011). Endogeneity in nonlinear regressions with integrated time series. *Econometric Reviews* **30**, 51–87.

Chang, Y., Park, J. Y. and P. C. B. Phillips (2001). Nonlinear econometric models with cointegrated and deterministically trending regressors. *Econometrics Journal* **4**, 1–36.

Choi, I. and P. Saikkonen (2010). Tests for nonlinear cointegration. *Econometric Theory* **26**, 682–709.

Christensen, L. R., Jorgenson, D. W. and L. J. Lau (1971). Conjugate duality and the transcendental logarithmic production function. *Econometrica* **39**, 255–256.

de Benedictis, L. F. and D. E. A. Giles (1998). Diagnostic testing in econometrics: Variable addition, RESET and Fourier approximations. In Ullah, A. and D. E. A. Giles (eds.), *Handbook of Applied Economic Statistics*, 383–417. New York: Marcel Dekker.

Dong, C. and J. Gao (2018). Specification testing by orthogonal series for nonlinear cointegration with endogeneity. *Econometric Theory* **34**, 754–789.

Dong, C. and O. Linton (2018). Additive nonparametric models with time variable and both stationary and nonstationary regressors. *Journal of Econometrics* **207**, 212–236.

Dong, C., Gao, J. and D. Tjøstheim (2016). Estimation for single-index and partially linear single-index integrated models. *Annals of Statistics* **44**, 425–453.

Dong, C., Linton, O. and P. Peng (2022). A weighted sieve estimator for nonparametric time series models with nonstationary regressors. *Journal of Econometrics* **222**, 909–932.

Engle, R. F. (1984). Wald, likelihood ratio and Lagrange multiplier tests in econometrics. In Griliches, Z. and Intriligator, M. (eds.), *Handbook of Econometrics*, Volume II, 775–826. Amsterdam: Elsevier.

Gao, J. and P. C. B. Phillips (2013). Semiparametric estimation in triangular system equations with nonstationarity. *Journal of Econometrics* **176**, 59–79.

Grabarczyk, P. and M. Wagner (2024). An integrated modified OLS analysis of the environmental Kuznets curve for CO_2 emissions. *Mimeo.*

Hansen, B. E. (1992). Convergence to stochastic integrals for dependent heterogeneous processes. *Econometric Theory* **8**, 489–500.

Hansen, B. E. and B. Seo (2002). Testing for two-regime threshold cointegration in vector error correction models. *Journal of Econometrics* **110**, 293–318.

Hong, S. H. and P. C. B. Phillips (2010). Testing linearity in cointegrating relations with an application to purchasing power parity. *Journal of Business and Economic Statistics* **28**, 96–114.

Hwang, T. and T. J. Vogelsang (2022). Some Fixed-b Results for Regressions with High Frequency Data over Long Spans. *Mimeo.*

Jansson, M. (2002). Consistent covariance matrix estimation for linear processes, *Econometric Theory* **18**, 1449–1459.

Jin, S., Phillips, P. C. B. and Y. Sun (2006). A new approach to robust inference in cointegration. *Economics Letters* **91**, 300–306.

Karlsen, H. A., Myklebust, T. and D. Tjøstheim (2007). Nonparametric estimation in a nonlinear cointegration type model. *Annals of Statistics* **35**, 252–299.

Kasparis, I. (2008). Detection of functional form misspecification in cointegrating relations. *Econometric Theory* **24**, 1373–1403.

Keenan, D. M. (1985). A Tukey nonadditivity type test for time series nonlinearity. *Biometrika* **72**, 39–44.

Kiefer, N. M. and T. J. Vogelsang (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory* **21**, 1130–1164.

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics* **54**, 159–178.

Lee, T. H., White, H. and C. W. J. Granger (1993). Testing for neglected nonlinearity in time series models: A comparison of neural network methods and alternative tests. *Journal of Econometrics* **56**, 269–290.

Linton, O. and Q. Wang (2016). Nonparametric transformation regression with nonstationary data. *Econometric Theory* **32**, 1–29.

Lütkepohl, H., Teräsvirta, T. and J. Wolters (1999). Investigating stability and linearity of a German M1 demand function. *Journal of Applied Econometrics* **14**, 511–525.

Newey, W. and K. West (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* **61**, 631–654.

Park, J. Y. (1990). Testing for unit roots and cointegration by variable addition. In Fomby, T. B. and G. F. Rhodes (eds.), *Advances in Econometrics: Co-Integration, Spurious Regression, and Unit Roots*, Volume 8, 107–133. Greenwich: JAI Press.

Park, J. Y. and B. Choi (1988). A new approach to testing for a unit root. *CAE Working Paper* 88–23, Cornell University.

Park, J. Y. and P. C. B. Phillips (1988). Statistical inference in regressions with inregrated processes: Part I. *Econometric Theory* **4**, 468–497.

Park, J. Y. and P. C. B. Phillips (1989). Statistical inference in regressions with inregrated processes: Part II. *Econometric Theory* **5**, 95–131.

Park, J. Y. and P. C. B. Phillips (2001). Nonlinear regressions with integrated time series. *Econometrica* **69**, 117–161.

Phillips, P. C. B. (1983). Best uniform and modified Padé approximants to probability densities in econometrics. In Hildenbrand, W. (ed.), *Advances in Econometrics*, 123–167. Cambridge: Cambridge University Press.

Phillips, P. C. B. and B. E. Hansen (1990). Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies* **57**, 99–125.

Phillips, P. C. B. and M. Loretan (1991). Estimating long-run economic equilibria. *Review of Economic Studies* **58**, 407–436.

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.

Saikkonen, P. (1991). Asymptotically efficient estimation of cointegrating regressions. *Econometric Theory* **7**, 1–21.

Saikkonen, P. and I. Choi (2004). Cointegrating smooth transition regressions. *Econometric Theory* **20**, 301–340.

Shi, X. and P. C. B. Phillips (2012). Nonlinear cointegrating regression under weak identification. *Econometric Theory* **28**, 509–547.

Shin, Y. (1994). A residual-based test for the null of cointegration against the alternative of no cointegration. *Econometric Theory* **10**, 91–115.

Sims, C. A., Stock, J. H. and M. W. Watson (1990). Inference in linear time series models with some unit roots. *Econometrica* **58**, 113–144.

Stock, J. H. and M. W. Watson (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* **61**, 783–820.

Sun, Y. (2014). Let's fix it: Fixed-b asymptotics versus small-b asymptotics in heteroskedasticity and autocorrelation robust testing. *Journal of Econometrics* **178**, 659–677.

Sun, Y., Phillips, P. C. B. and S. Jin (2008). Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing. *Econometrica* **76**, 175–194.

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.

Tjøstheim, D. (2020). Some notes on nonlinear cointegration: A partial review with some novel perspectives. *Econometric Reviews* **39**, 655–673.

Tsay, R. S. (1986). Nonlinearity tests for time series. *Biometrika* **73**, 461–466.

Veldhuis, S. and M. Wagner (2024). IM-OLS residual-based cointegration tests for cointegrating multivariate polynomial regressions. *Mimeo.*

Vogelsang, T. J. and M. Wagner (2013). A fixed-b perspective on the Phillips-Perron unit root tests. *Econometric Theory* **29**, 609–628.

Vogelsang, T. J. and M. Wagner (2014a). Integrated modified OLS estimation and fixed-b inference for cointegrating regressions. *Journal of Econometrics* **178**, 741–760.

Vogelsang, T. J. and M. Wagner (2014b). An integrated modified OLS RESET test for cointegrating regressions. *SFB823 Discussion Paper* 37/14.

Wagner, M. (2015). The environmental Kuznets curve, cointegration and nonlinearity. *Journal of Applied Econometrics* **30**, 948–967.

Wagner, M. (2023). Residual-based cointegration and non-cointegration tests for cointegrating polynomial regressions. *Empirical Economics* **65**, 1–31.

Wagner, M. and S. H. Hong (2016). Cointegrating polynomial regressions: Fully modified OLS estimation and inference. *Econometric Theory* **32**, 1289–1315.

Wang, Q. and P. C. B. Phillips (2009). Structural nonparametric cointegrating regression. *Econometrica* **77**, 1901–1948.

Wang, Q. and P. C. B. Phillips (2012). A specification test for nonlinear nonstationary models. *Annals of Statistics* **40**, 727–758.

Wang, Q. and P. C. B. Phillips (2016). Nonparametric cointegrating regression with endogeneity and long memory. *Econometric Theory* **32**, 359–401.

Wang, Q. and Y. X. R. Wang (2013). Nonparametric cointegrating regression with NNH errors. *Econometric Theory* **29**, 1–27.

Wang, Q. and K. Zhu (2020). On a measure of lack of fit in nonlinear cointegrating regression with endogeneity. *Statistica Sinica* **30**, 371–396.

Wang, Q., Wu, D. and K. Zhu (2018). Model checks for nonlinear cointegrating regression. *Journal of Econometrics* **207**, 261–284.

Proofs

Proof of Proposition 1

Using standard algebra and the fact that the block of A_{IM} corresponding to γ is an identity matrix, allows to write:

$$\begin{aligned} A_{\text{IM}}^{-1}\left(\hat{\vartheta}-\vartheta^*\right) &= \left(T^{-2}A_{\text{IM}}\tilde{S}^{Z\prime}\tilde{S}^{Z}A_{\text{IM}}\right)^{-1}\left(T^{-2}A_{\text{IM}}\tilde{S}^{Z\prime}S^{u}\right)-[0,(\Omega_{vv}^{-1}\Omega_{vu})']' \\ &= \left(T^{-1}\left(T^{-1/2}A_{\text{IM}}\tilde{S}^{Z}\right)'\left(T^{-1/2}A_{\text{IM}}\tilde{S}^{Z}\right)\right)^{-1}\left(T^{-1}\left(T^{-1/2}A_{\text{IM}}\tilde{S}^{Z}\right)'\left(T^{-1/2}S^{u}\right)\right) \\ &\quad -[0,(\Omega_{vv}^{-1}\Omega_{vu})']'. \end{aligned} \tag{43}$$

The asymptotic distribution follows from the continuous mapping theorem upon combining the limits of the constituent components, which are given by $T^{-1/2}S^{u}_{\lfloor rT\rfloor} \Rightarrow B_u(r)$ for $0 \leq r \leq 1$ and $T^{-1/2}A_{\text{IM}}\tilde{S}^{Z}_{\lfloor rT\rfloor} \Rightarrow [\int_0^r Z(s)'ds, B_v(r)']' = f(r)$ for $0 \leq r \leq 1$, with $Z(r)$ as defined in the main text above Proposition 1. These results in turn lead to:

$$A_{\text{IM}}^{-1}\left(\hat{\vartheta}-\vartheta^*\right) \Rightarrow \left(\int_0^1 f(s)f(s)'ds\right)^{-1}\int_0^1 f(s)B_u(s)ds-[0,(\Omega_{vv}^{-1}\Omega_{vu})']'. \tag{44}$$

Next, using $B_u(s) = \omega_{u\cdot v}^{1/2}w_{u\cdot v}(s)+B_v(s)'\Omega_{vv}^{-1}\Omega_{vu}$ implies that:

$$\int_0^1 f(s)B_u(s)ds = \omega_{u\cdot v}^{1/2}\int_0^1 f(s)w_{u\cdot v}(s)ds+\int_0^1 f(s)B_v(s)'ds\Omega_{vv}^{-1}\Omega_{vu}. \tag{45}$$

Given that $B_v(s)$ corresponds to the last m elements of $f(s)$ it follows that:

$$\left(\int_0^1 f(s)f(s)'ds\right)^{-1}\int_0^1 f(s)B_v(s)'ds\Omega_{vv}^{-1}\Omega_{vu} = [0,(\Omega_{vv}^{-1}\Omega_{vu})']', \tag{46}$$

which implies that:

$$A_{\text{IM}}^{-1}\left(\hat{\vartheta}-\vartheta^*\right) \Rightarrow \omega_{u\cdot v}^{1/2}\left(\int_0^1 f(s)f(s)'ds\right)^{-1}\int_0^1 f(s)w_{u\cdot v}(s)ds, \tag{47}$$

which is the result given in (13). The representation in the second line of (13) in terms of $dw_{u\cdot v}(s)$ follows using integration by parts and the definition of $F(s)$.

Proof of Proposition 2

Under the null hypothesis $H_0: R\vartheta = r$, the test statistic can be written as:

$$\begin{aligned} T_W &= \left(R(\hat{\vartheta}-\vartheta^*)\right)'\left[R\hat{V}_{\text{IM}}R'\right]^{-1}\left(R(\hat{\vartheta}-\vartheta^*)\right) \\ &= \left(\left(A_R R A_{\text{IM}}\right)A_{\text{IM}}^{-1}(\hat{\vartheta}-\vartheta^*)\right)'\left[\left(A_R R A_{\text{IM}}\right)A_{\text{IM}}^{-1}\hat{V}_{\text{IM}}A_{\text{IM}}^{-1}\left(A_R R A_{\text{IM}}\right)\right]^{-1}\left((A_R R A_{\text{IM}})A_{\text{IM}}^{-1}(\hat{\vartheta}-\vartheta^*)\right). \end{aligned} \tag{48}$$

By assumption $A_R R A_{\text{IM}} \to R^* \in \mathbb{R}^{s\times s}$ of full rank and Proposition 1 has established that $A_{\text{IM}}^{-1}\left(\hat{\vartheta}-\vartheta^*\right)$ converges to the limiting distribution given in (13). It thus remains to consider $A_{\text{IM}}^{-1}\hat{V}_{\text{IM}}A_{\text{IM}}^{-1}$, with $\hat{V}_{\text{IM}}$ as defined in (16):

$$\hat{V}_{\text{IM}} = \hat{\omega}_{u\cdot v}\left(T^{-2}A_{\text{IM}}\tilde{S}^{Z\prime}\tilde{S}^{Z}A_{\text{IM}}\right)^{-1}\left(T^{-4}A_{\text{IM}}C'CA_{\text{IM}}\right)\left(T^{-2}A_{\text{IM}}\tilde{S}^{Z\prime}\tilde{S}^{Z}A_{\text{IM}}\right)^{-1}. \tag{49}$$

The asymptotic behavior of the outer term has been established in the proof of Proposition 1 and it remains to consider the term in the middle. Straightforward calculations show that $T^{-3/2}A_{\text{IM}}C_{\lfloor rT \rfloor} \Rightarrow F(1) - F(r)$ for $0 \leq r \leq 1$, which implies that the central term converges to $\int_0^1 [F(1) - F(s)][F(1) - F(s)]'ds$ for $0 \leq r \leq 1$. Consequently, when $\hat{\omega}_{u \cdot v} \xrightarrow{\mathbb{P}} \omega_{u \cdot v}$ it follows that $A_{\text{IM}}^{-1} \hat{V}_{\text{IM}} A_{\text{IM}}^{-1} \Rightarrow V_{\text{IM}}$. The, conditional upon $B_v(r)$, limiting distribution of $A_{\text{IM}}^{-1}\left(\hat{\vartheta} - \vartheta^*\right)$ is normal with covariance matrix V_{IM}. This immediately – and as usual in the cointegration literature – implies that T_W converges under the null hypothesis, conditional upon $B_v(r)$ and due to the independence of the limiting distribution from $B_v(r)$ also unconditionally, to a chi-squared distribution with s degrees of freedom. The t-type test statistic result for T_t as defined in (18) follows entirely analogously.

Proof of Corollary 1

The result follows immediately by rewriting using $f(s) = \Pi g(s)$.

Proof of Proposition 3

The key results to be established are (i) the asymptotic behavior of the partial sum process $T^{-1/2}\sum_{t=2}^{\lfloor rT \rfloor} \hat{S}_{t,\text{M}}^u$ and (ii) that this limiting process is independent of the chi-squared random variable appearing in the numerator of the null limiting distribution of the test statistic. Both steps are similar to the corresponding results for cointegrating linear regressions derived in Vogelsang and Wagner (2014a, Theorem 3).

In relation to Remark 3, note that $\hat{S}_{t,\text{M}}^u$ is equivalently given by the OLS residuals of estimating (31) with $M_t^{\perp}$ replaced by M_t. Defining $S_t^* := [\tilde{S}_t^{Z\prime}, M_t']'$, it thus follows that:

$$\hat{S}_{t,\text{M}}^u = S_t^u - S_t^{*\prime}\left(\sum_{t=1}^{T} S_t^* S_t^{*\prime}\right)^{-1} \sum_{t=1}^{T} S_t^* S_t^u. \tag{50}$$

With respect to the asymptotic behavior of a properly scaled version of S_t^*, we already know from the proof of Proposition 1 that $T^{-1/2}A_{\text{IM}}\hat{S}_{\lfloor rT \rfloor}^Z \Rightarrow f(r)$ for $0 \leq r \leq 1$. It thus remains to consider the second block-component:

$$\begin{aligned} T^{-5/2}A_{\text{IM}}M_{\lfloor rT \rfloor} &= \frac{\lfloor rT \rfloor}{T} T^{-1}\sum_{j=1}^{T} T^{-1/2}A_{\text{IM}}\tilde{S}_j^Z - T^{-1}\sum_{j=1}^{\lfloor rT \rfloor - 1} T^{-1}\sum_{i=1}^{j} T^{-1/2}A_{\text{IM}}\tilde{S}_j^Z \\ &\Rightarrow r\int_0^1 f(s)ds - \int_0^r \left(\int_0^s f(v)dv\right) ds = \int_0^r F(1) - \int_0^r F(s)ds \\ &= \int_0^r [F(1) - F(s)]ds. \end{aligned} \tag{51}$$

Defining $A_{\text{IM}}^* := \text{diag}(A_{\text{IM}}, T^{-2}A_{\text{IM}})$ leads to:

$$T^{-1/2}A_{\text{IM}}^* S_{\lfloor rT \rfloor}^* = \begin{bmatrix} T^{-1/2}A_{\text{IM}}\hat{S}_{\lfloor rT \rfloor}^Z \\ T^{-5/2}A_{\text{IM}}M_{\lfloor rT \rfloor} \end{bmatrix} \Rightarrow \begin{bmatrix} f(r) \\ \int_0^r [F(1) - F(s)]ds \end{bmatrix}. \tag{52}$$

Defining $\Pi^* := \text{diag}(\Pi, \Pi)$, it follows immediately that:

$$\begin{bmatrix} f(r) \\ \int_0^r [F(1) - F(s)]ds \end{bmatrix} = \begin{bmatrix} \Pi g(r) \\ \Pi \int_0^r [G(1) - G(s)]ds \end{bmatrix} = \Pi^* h(r). \tag{53}$$

The next step requires full design, i. e., invertibility of Π and thus of Π^*:

$$T^{-1/2}\hat{S}^u_{\lfloor rT\rfloor,\mathrm{M}} = T^{-1/2}S^u_{\lfloor rT\rfloor} - T^{-1/2}\hat{S}^{*\prime}_{\lfloor rT\rfloor}\left(\sum_{t=1}^{T} S^*_t S^{*\prime}_t\right)^{-1}\sum_{t=1}^{T} S^*_t S^u_t \tag{54}$$
$$\Rightarrow B_u(r) - (\Pi^* h(r))'\left(\Pi^*\int_0^1 h(s)h(s)'ds\Pi^{*\prime}\right)^{-1}\left(\Pi^*\int_0^1 h(s)B_u(s)ds\right)$$
$$= B_u(s) - h(s)'\left(\int_0^1 h(s)h(s)'ds\right)^{-1}\int_0^1 h(s)B_u(s)ds.$$

The argument now continues similarly to the argument used in the proof of Proposition 2, using that $B_u(s) = \omega_{u\cdot v}^{1/2}w_{u\cdot v}(r) + W_v(s)'\Omega_{vv}^{-1/2}\Omega_{vu}$. Noting next that $W_v(s)$ is a block of $h(r)$ implies, similarly to (46), that:

$$h(r)'\left(\int_0^1 h(s)h(s)'ds\right)^{-1}\int_0^1 h(s)W_v(s)'ds\Omega_{vu}^{-1/2}\Omega_{vu} = h(r)'\begin{bmatrix} 0 \\ I_m \\ 0 \end{bmatrix}\Omega_{vv}^{-1/2}\Omega_{vu} = W_v(r)'\Omega_{vv}^{-1/2}\Omega_{vu}. \tag{55}$$

This yields:

$$T^{-1/2}\hat{S}^u_{\lfloor rT\rfloor,\mathrm{M}} \Rightarrow \omega_{u\cdot v}^{1/2}\left(\int_0^r dw_{u\cdot v}(s) - h(r)'\left(\int_0^1 h(s)h(s)'ds\right)^{-1}\int_0^1 h(s)w_{u\cdot v}(s)ds\right) \tag{56}$$
$$= \omega_{u\cdot v}^{1/2}\left(\int_0^r dw_{u\cdot v}(s) - h(r)'\left(\int_0^1 h(s)h(s)'ds\right)^{-1}\int_0^1 [H(1)-H(s)]dw_{u\cdot v}(s)\right)$$
$$= \omega_{u\cdot v}^{1/2}P(r).$$

It remains to establish independence of $\mathcal{O}_s$, the chi-squared distributed random variable in the numerator of the fixed-b test statistic, and the process $P(r)$. This is shown, using again the same argument as in Vogelsang and Wagner (2014a, Theorem 3) by showing that the two quantities – $\mathcal{O}_s$ and $P(r)$ – are independent conditionally upon $W_v(r)$. With $\mathcal{O}_s$ independent of $W_v(r)$ this also establishes unconditional independence.

Conditional upon $W_v(r)$, both the numerator and the denominator are (deterministic) functionals of $w_{u\cdot v}(r)$, i. e., of a Gaussian process. The random variable $\mathcal{O}_s$ is effectively a (scaled) quadratic form of the (conditionally) Gaussian process $\int_0^1[G(1)-G(s)]dw_{u\cdot v}(s)$, compare (20). Therefore, to establish (conditional) independence it suffices to calculate:

$$\mathrm{Cov}\left(P(r).\int_0^1 [G(1)-G(s)]dw_{u\cdot v}(s)\right) = \int_0^1 [G(1)-G(s)]'ds - \tag{57}$$
$$- h(r)'\left(\int_0^1 h(s)h(s)'ds\right)^{-1}\int_0^1 [H(1)-H(s)][G(1)-G(s)]'ds.$$

The first term above in (57) equals, by definition, the transposed second lower block of $h(r) = [g(r)', h_2(r)']'$ as defined in (29), referred to as $h_2(r)$ from now on. Zero covariance is, therefore, established upon showing that the second term on the right hand side of (57) is also equal to $h_2(r)'$. In this respect, using integration by parts shows that:

$$\int_0^1 [H(1)-H(s)][G(1)-G(s)]'ds = \int_0^1 h(s)h_2(s)'ds, \tag{58}$$

which implies that:

$$h(r)'\left(\int_0^1 h(s)h(s)'ds\right)^{-1}\int_0^1 h(s)h_2(s)'ds = h(r)'\begin{bmatrix} 0 \\ I_{|\mathcal{I}|+m} \end{bmatrix} = h_2(r)'. \tag{59}$$

This establishes that $\mathrm{Cov}(P(r), \int_0^1 [G(1) - G(s)]dw_{u\cdot v}(s)) = 0$. The result for the fixed-b t-type test statistic $T_{t,b}$ as defined in (30) follows entirely analogously.

Table 1: Empirical Null-Rejection Probabilities, Data-Dependent Bandwidths

Panel A: IM(O)

		Bartlett				QS			
		And		NW		And		NW	
T	ρ_1,ρ_2	$q=2$	$q=3$	$q=2$	$q=3$	$q=2$	$q=3$	$q=2$	$q=3$
100	0	.1545	.3187	.0947	.1667	.1963	.4054	.1194	.2319
	.3	.2033	.4006	.2009	.3560	.2043	.4137	.1781	.3310
	.6	.2706	.4902	.3573	.5946	.2449	.4453	.3023	.5071
	.8	.4034	.6767	.5667	.8343	.3962	.6545	.4901	.7633
200	0	.1008	.1888	.0733	.1128	.1128	.2139	.0870	.1494
	.3	.1314	.2403	.1452	.2457	.2457	.2231	.1188	.2035
	.6	.1714	.3117	.2513	.4349	.4349	.2686	.2019	.3428
	.8	.2460	.4905	.4152	.7008	.7008	.4593	.3342	.6016
500	0	.0735	.1085	.0600	.0725	.0740	.1080	.0622	.0766
	.3	.0917	.1416	.1083	.1662	.0803	.1183	.0910	.1279
	.6	.1046	.1696	.1646	.2752	.0910	.1387	.1444	.2342
	.8	.1158	.2256	.2673	.4932	.1020	.2001	.2445	.4596

Panel B: IM(fb)

		Bartlett				QS			
		And		NW		And		NW	
T	ρ_1,ρ_2	$q=2$	$q=3$	$q=2$	$q=3$	$q=2$	$q=3$	$q=2$	$q=3$
100	0	.0566	.0743	.0580	.0825	.0529	.0659	.0548	.0828
	.3	.1716	.2850	.1327	.2508	.1300	.2037	.0858	.1325
	.6	.4062	.7442	.3283	.6290	.3349	.6180	.2041	.3296
	.8	.7045	.9707	.6926	.9432	.6148	.9433	.5418	.7778
200	0	.0520	.0521	.0520	.0521	.0524	.0549	.0524	.0549
	.3	.0682	.0833	.0682	.0833	.0547	.0612	.0547	.0612
	.6	.0996	.1586	.0996	.1586	.0692	.0877	.0692	.0877
	.8	.1829	.3873	.1836	.3884	.1240	.2314	.1240	.2314
500	0	.0520	.0499	.0520	.0499	.0524	.0521	.0524	.0521
	.3	.0681	.0816	.0681	.0816	.0547	.0590	.0547	.0590
	.6	.1026	.1609	.1026	.1609	.0702	.0853	.0702	.0853
	.8	.2048	.3921	.2075	.3942	.1326	.2342	.1326	.2343

Table 2: Empirical Size-Adjusted Power, Data-Dependent Bandwidths, $T = 200$

Panel A: IM(O)										
			Bartlett				QS			
			And		NW		And		NW	
$G(X_t)$	ϕ	ρ_1,ρ_2	$q=2$	$q=3$	$q=2$	$q=3$	$q=2$	$q=3$	$q=2$	$q=3$
NA, Null	0	0	.1008	.1888	.0733	.1128	.1128	.2139	.0870	.1494
		.3	.1314	.2403	.1452	.2457	.2457	.2231	.1188	.2035
		.6	.1714	.3117	.2513	.4349	.4349	.2686	.2019	.3428
		.8	.2460	.4905	.4152	.7008	.7008	.4593	.3342	.6016
x_{1t}^2	.01	0	.8600	.8487	.8631	.8547	.8549	.8388	.8615	.8499
		.3	.7473	.7260	.7520	.7325	.7453	.7193	.7488	.7289
		.6	.5243	.4830	.5162	.4714	.5209	.4789	.5184	.4668
		.8	.2495	.2157	.2611	.2263	.2441	.2097	.2589	.2256
$x_{1t}^2 + x_{1t}x_{2t}$	.01	0	.9202	.9136	.9220	.9193	.9176	.9048	.9208	.9151
		.3	.8270	.8164	.8312	.8177	.8239	.8099	.8276	.8181
		.6	.6191	.5798	.6136	.5693	.6177	.5748	.6125	.5675
		.8	.3066	.2660	.3199	.2871	.3010	.2617	.3177	.2822
$x_{1t}^2 + x_{2t}^2 + x_{1t}x_{2t}$	.01	0	.9756	.9758	.9774	.9784	.9745	.9728	.9765	.9772
		.3	.9283	.9214	.9286	.9243	.9270	.9179	.9288	.9229
		.6	.7755	.7458	.7715	.7368	.7733	.7415	.7721	.7350
		.8	.4551	.3967	.4691	.4195	.4490	.3894	.4668	.4163
$x_{1t}x_{2t}$	.01	0	.8209	.8028	.8267	.8099	.8176	.7899	.8244	.8024
		.3	.6786	.6404	.6790	.6411	.6744	.6314	.6765	.6435
		.6	.4082	.3548	.4015	.3402	.4049	.3508	.4017	.3390
		.8	.1572	.1324	.1647	.1409	.1543	.1287	.1628	.1396
x_{1t}^3	.001	0	.8380	.8690	.8470	.8780	.8310	.8670	.8480	.8750
		.3	.7810	.8200	.7920	.8290	.7780	.8180	.7850	.8240
		.6	.6850	.7180	.6910	.7090	.6830	.7160	.6880	.7020
		.8	.5300	.5480	.5410	.5550	.5170	.5370	.5390	.5620
$x_{1t}\left(1+e^{-x_{1t}}\right)^{-1}$	.5	0	.5320	.5856	.5456	.5901	.5250	.5773	.5390	.5850
		.3	.4660	.5115	.4753	.5157	.4645	.5066	.4724	.5143
		.6	.3261	.3446	.3260	.3377	.3241	.3414	.3237	.3334
		.8	.1466	.1336	.1535	.1429	.1445	.1299	.1511	.1395

Table 3: Empirical Size-Adjusted Power, Data-Dependent Bandwidths, $T = 200$

Panel B: IM(fb)

			Bartlett				QS			
			And		NW		And		NW	
$G(X_t)$	ϕ	ρ_1,ρ_2	$q=2$	$q=3$	$q=2$	$q=3$	$q=2$	$q=3$	$q=2$	$q=3$
NA, Null	0	0	.0520	.0521	.0520	.0521	.0524	.0549	.0524	.0549
		.3	.0682	.0833	.0682	.0833	.0547	.0612	.0547	.0612
		.6	.0996	.1586	.0996	.1586	.0692	.0877	.0692	.0877
		.8	.1829	.3873	.1836	.3884	.1240	.2314	.1240	.2314
x_{1t}^2	.01	0	.8544	.8395	.8544	.8395	.8455	.8205	.8455	.8205
		.3	.7362	.7000	.7362	.7000	.7258	.6853	.7258	.6853
		.6	.4997	.4404	.4956	.4381	.4910	.4246	.4899	.4246
		.8	.2274	.1982	.2215	.1951	.2221	.1928	.2190	.1891
$x_{1t}^2 + x_{1t}x_{2t}$	.01	0	.9160	.9031	.9160	.9031	.9087	.8918	.9087	.8918
		.3	.8185	.7912	.8185	.7912	.8091	.7763	.8091	.7763
		.6	.5925	.5363	.5898	.5353	.5858	.5174	.5842	.5172
		.8	.2774	.2458	.2771	.2388	.2782	.2377	.2733	.2331
$x_{1t}^2 + x_{2t}^2 + x_{1t}x_{2t}$	.01	0	.9741	.9720	.9741	.9720	.9711	.9656	.9711	.9656
		.3	.9226	.9084	.9226	.9084	.9157	.8987	.9157	.8987
		.6	.7549	.7052	.7522	.7035	.7460	.6886	.7452	.6885
		.8	.4145	.3622	.4138	.3584	.4094	.3522	.4093	.3472
$x_{1t}x_{2t}$	.01	0	.8163	.7865	.8163	.7865	.8050	.7648	.8050	.7648
		.3	.6625	.6059	.6625	.6059	.6510	.5838	.6510	.5838
		.6	.3827	.3176	.3782	.3164	.3731	.3024	.3717	.3023
		.8	.1430	.1221	.1404	.1209	.1427	.1212	.1391	.1192
x_{1t}^3	.001	0	.8400	.8700	.8400	.8700	.8360	.8620	.8360	.8620
		.3	.7940	.8140	.7940	.8140	.7890	.8040	.7890	.8040
		.6	.6820	.6930	.6790	.6920	.6780	.6830	.6770	.6830
		.8	.5150	.5220	.5160	.5190	.5220	.5110	.5130	.5080
$x_{1t}\left(1+e^{-x_{1t}}\right)^{-1}$	.5	0	.5440	.5816	.5440	.5816	.5344	.5702	.5344	.5702
		.3	.4711	.4970	.4710	.4970	.4629	.4862	.4629	.4862
		.6	.3197	.3147	.3161	.3135	.3119	.3008	.3105	.3007
		.8	.1302	.1221	.1291	.1197	.1282	.1162	.1275	.1153

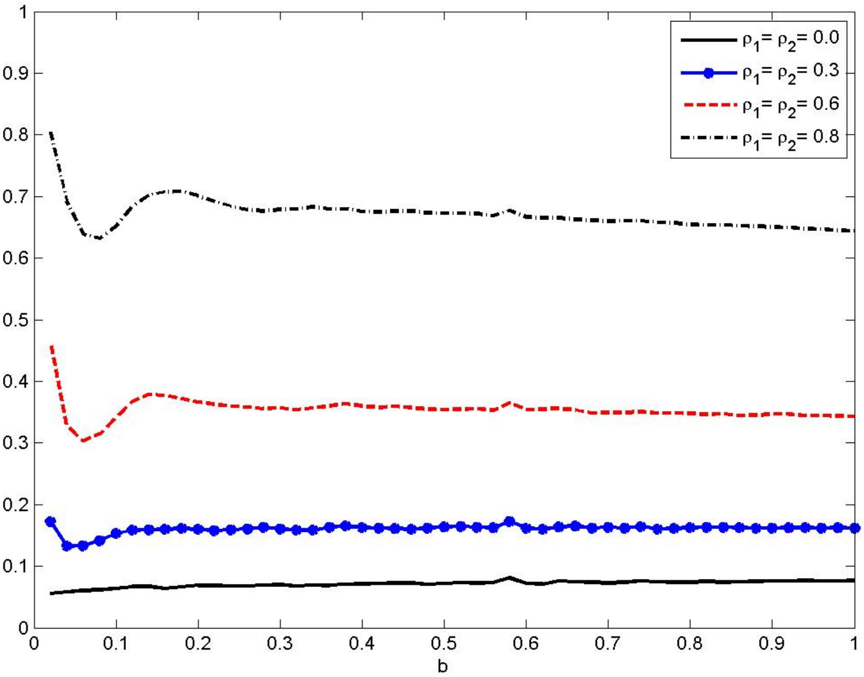

Figure 1: Empirical Null Rejections, IM(fb), $q = 2$, $T = 100$, Bartlett Kernel, 5% Nominal Level

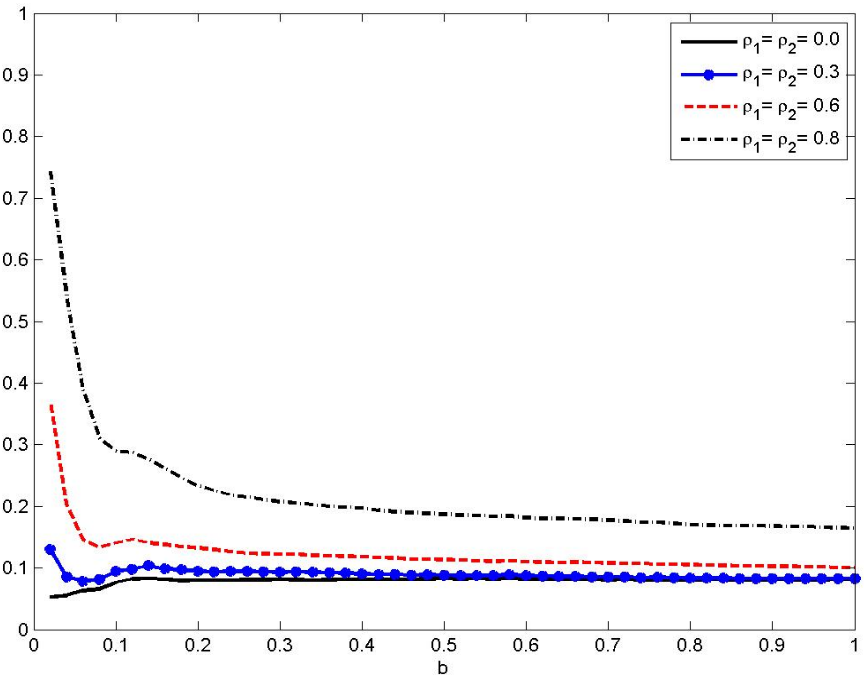

Figure 2: Empirical Null Rejections, IM(fb), $q = 2$, $T = 100$, QS Kernel, 5% Nominal Level

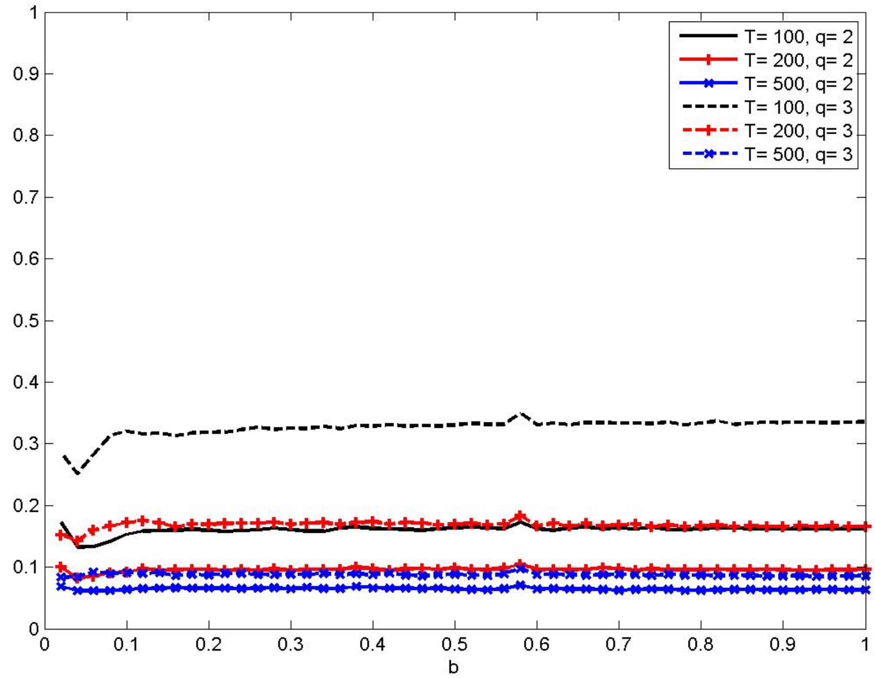

Figure 3: Empirical Null Rejections, IM(fb), $\rho_1 = \rho_2 = .3$, Bartlett Kernel, 5% Nominal Level

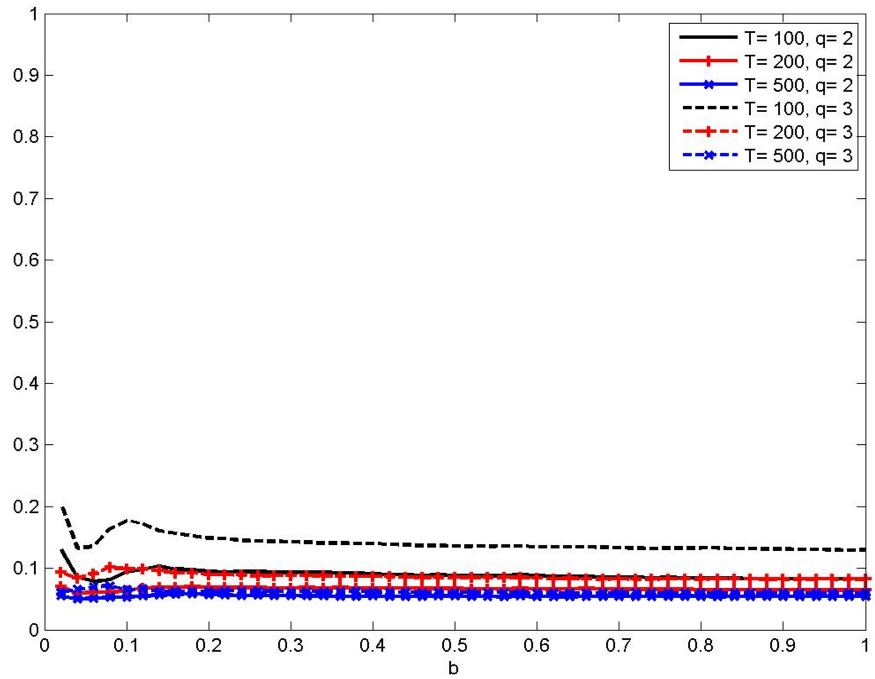

Figure 4: Empirical Null Rejections, IM(fb), $\rho_1 = \rho_2 = .3$, QS Kernel, 5% Nominal Level

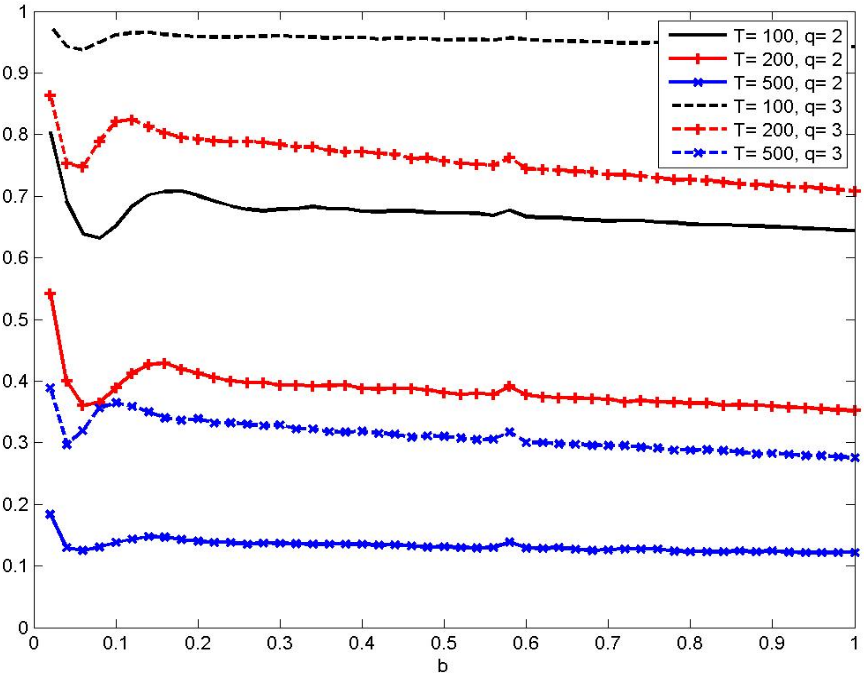

Figure 5: Empirical Null Rejections, IM(fb), $\rho_1 = \rho_2 = .8$, Bartlett Kernel, 5% Nominal Level

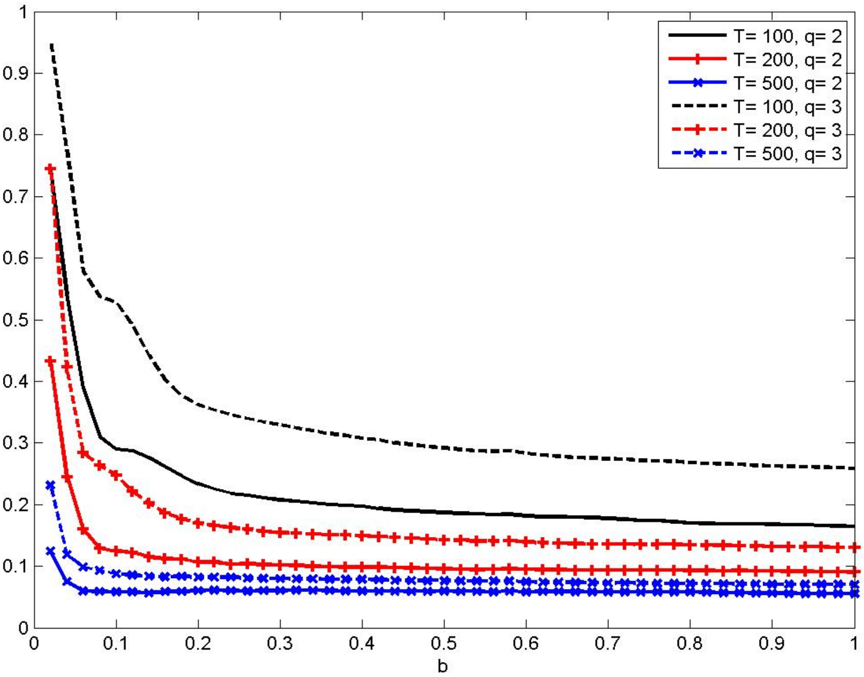

Figure 6: Empirical Null Rejections, IM(fb), $\rho_1 = \rho_2 = .8$, QS Kernel, 5% Nominal Level

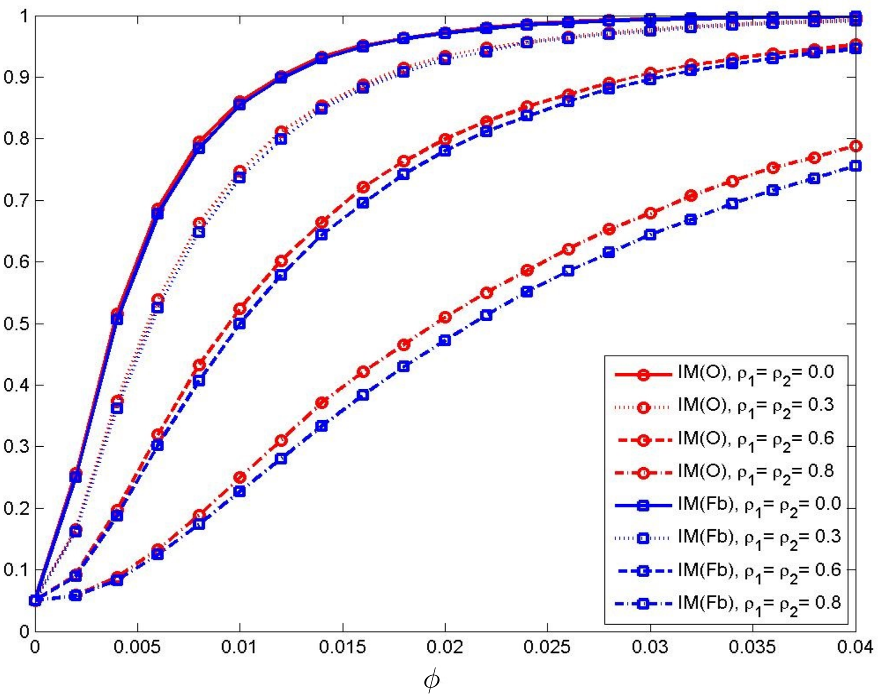

Figure 7: Empirical Size-Adjusted Power, $T = 200$, $q = 2$, $G(X_t) = x_{1t}^2$, Bartlett Kernel, Andrews Bandwidth, 5% Nominal Level

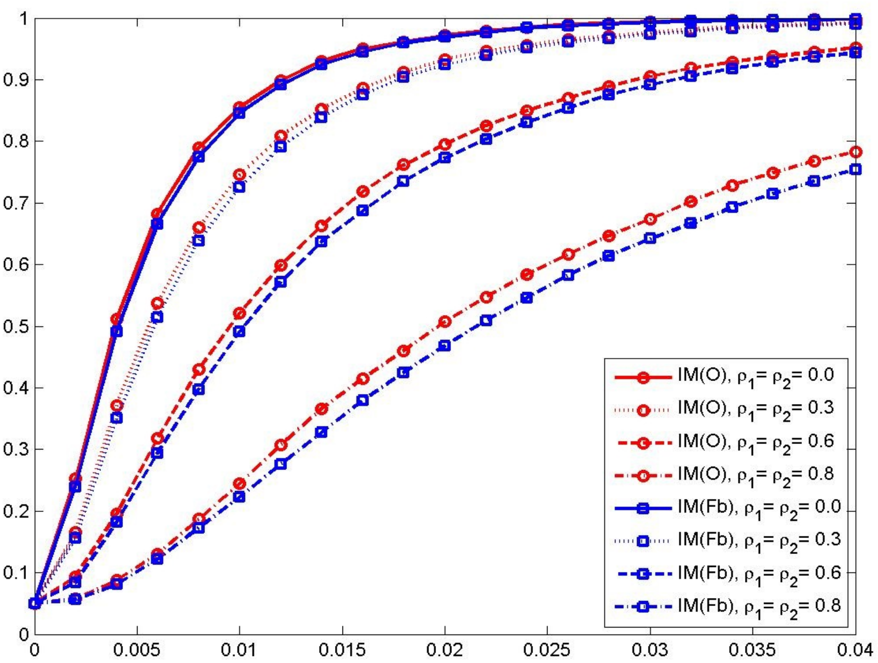

Figure 8: Empirical Size-Adjusted Power, $T = 200$, $q = 2$, $G(X_t) = x_{1t}^2$, QS Kernel, Andrews Bandwidth, 5% Nominal Level

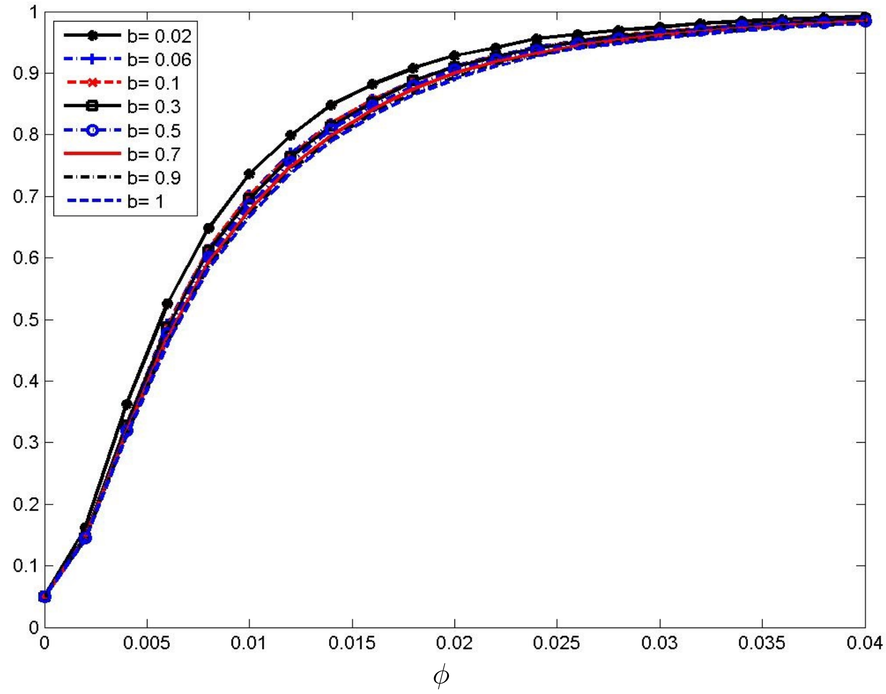

Figure 9: Empirical Size-Adjusted Power, IM(fb), $T = 200$, $q = 2$, $\rho_1 = \rho_2 = .3$, $G(X_t) = x_{1t}^2$, Bartlett Kernel, 5% Nominal Level

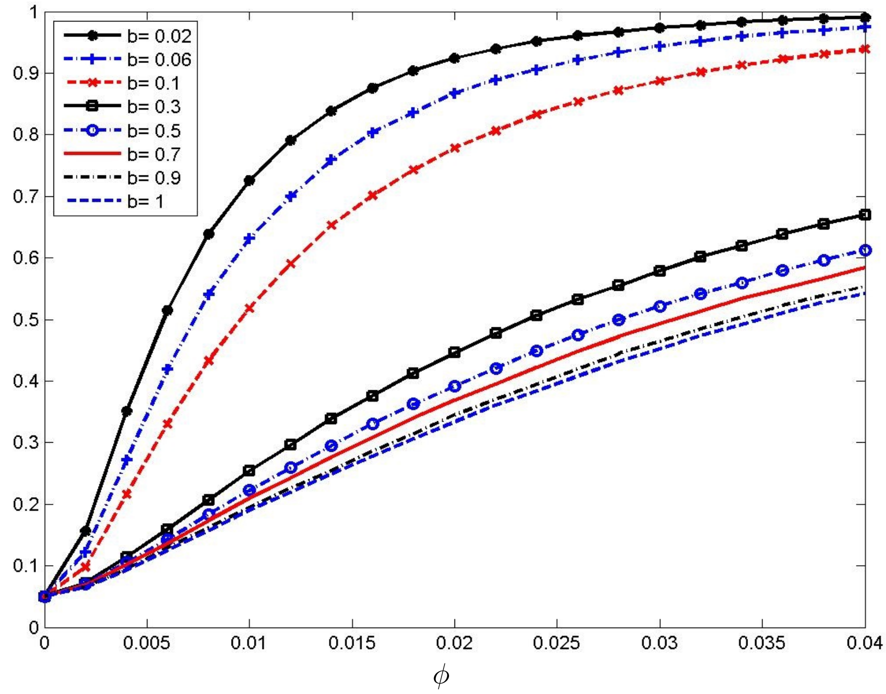

Figure 10: Empirical Size-Adjusted Power, IM(fb), $T = 200$, $q = 2$, $\rho_1 = \rho_2 = .3$, $G(X_t) = x_{1t}^2$, QS Kernel, 5% Nominal Level

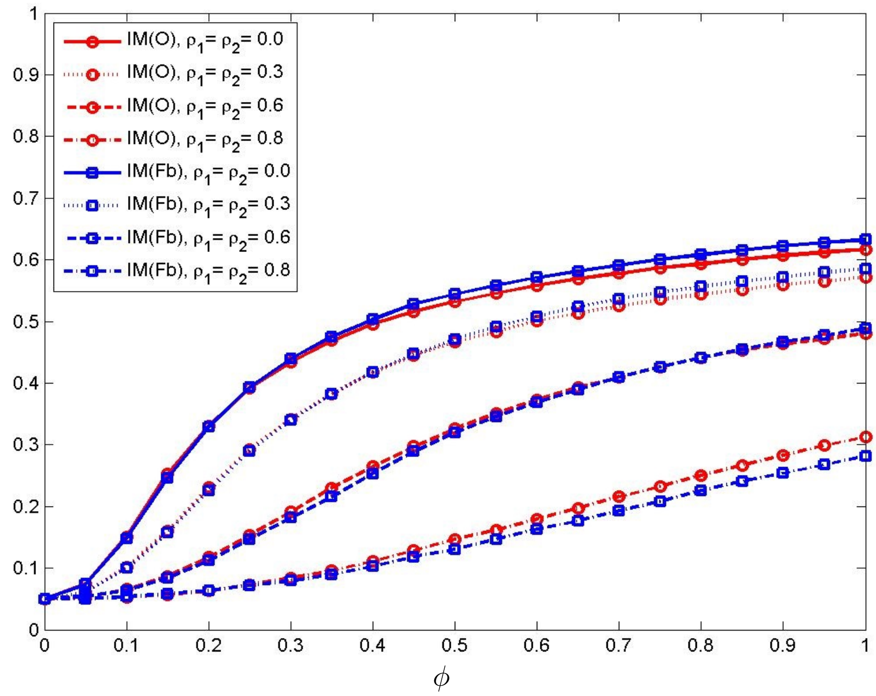

Figure 11: Empirical Size-Adjusted Power, $T = 200$, $q = 2$, $G(X_t) = x_{1t}\left(1 + e^{-x_{1t}}\right)^{-1}$, Bartlett Kernel, Andrews Bandwidth, 5% Nominal Level

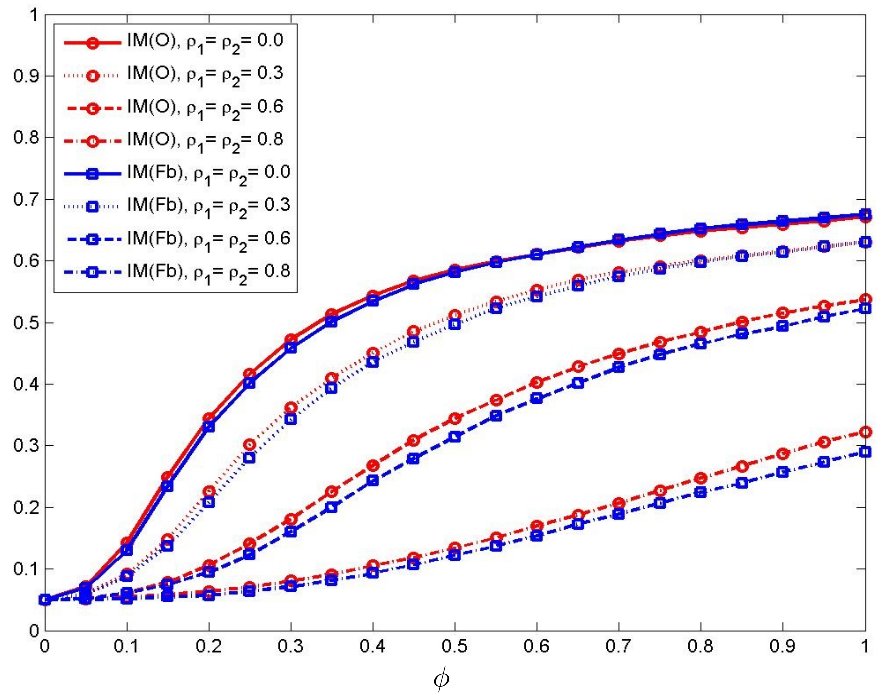

Figure 12: Empirical Size-Adjusted Power, $T = 200$, $q = 3$, $G(X_t) = x_{1t}\left(1 + e^{-x_{1t}}\right)^{-1}$, Bartlett Kernel, Andrews Bandwidth, 5% Nominal Level