SFB 823

An integrated modified OLS RESET test for cointegrating regressions

Timothy J. Vogelsang, Martin Wagner

Nr. 37/2014

Discussion Paper

SFB 823

An Integrated Modified OLS RESET Test for Cointegrating Regressions

Timothy J. Vogelsang
Department of Economics
Michigan State University

Martin Wagner
Faculty of Statistics
Technical University Dortmund
&
Institute for Advanced Studies, Vienna

October 2014

Abstract

We propose a RESET-type test for the null hypothesis of linearity of a cointegrating relationship with an asymptotic chi-squared null distribution. The test is based on an extension of the Integrated Modified OLS estimator of Vogelsang and Wagner (2014) from linear cointegrating relationships to multivariate cointegrating polynomial relationships. For the case of *full design* we furthermore provide fixed-b asymptotic theory for our RESET test. The theoretical results are complemented by a small simulation study.

JEL Classification: C12, C13, C32
Keywords: Bandwidth, cointegration, fixed-b asymptotics, IM-OLS, kernel, nonlinearity, specification test

1 Introduction

The vast majority of both theoretical as well as applied work dealing with cointegration is concerned with linear cointegrating relationships, be it in a regression setting or when considering dynamic linear models featuring cointegration, e.g., cointegrated VAR or state space models. For many applications, linearity may well be an apt description of or at least approximation to the problem. Nevertheless there is growing interest in nonlinear cointegrating relationships spurred from different areas of application ranging from empirical macroecoonomics, 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 hypothesis (Wagner, 2014).[1]

[1] A significant part of the 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. In this paper, however, we are concerned with *general or unspecified* nonlinearity of a cointegrating relationship. As such our tests may be useful also for the emerging nonparametric cointegration literature, see, e.g., Karlsen *et al.* (2007) or Wang and Phillips (2009).

More often than not economic theory does not specify the precise form of a potential nonlinear cointegrating relationship under the alternative and therefore, we believe, an *omnibus* specification test such as the RESET test (see Ramsey, 1969) is potentially useful. This type of test is convenient as it is based on replacement of an unknown nonlinear function by a finite sum approximation, in the RESET case a polynomial approximation. Such an approach has a long tradition in specification testing for stationary time series models, see, e.g., Phillips (1983), Lee *et al.* (1993) or de Benedictis and Giles (1998). It turns out that for our purposes it is convenient to resort to the Thursby and Schmidt (1977) formulation of the RESET test, in which higher order powers and/or cross-products of powers of the regressors are added to the regression rather than using the original formulation of Ramsey (1969); see also Keenan (1985) or Tsay (1986), in which the least squares residuals of the linear model are regressed on the auxiliary regressors.

Adding powers and/or cross-products of powers of the integrated regressors to the cointegrating regression has the advantage that the resulting estimation problem is linear in parameters. Therefore these parameters can be estimated under the null with a zero mean Gaussian mixture limiting distribution that allows for asymptotic standard inference by using a slight extension of the integrated modified OLS (IM-OLS) estimator of Vogelsang and Wagner (2014). For IM-OLS estimation no choices with respect to tuning parameters like kernel and bandwidths or lead and lag choices have to be made.[2] When using the IM-OLS estimator to test hypotheses, however, a scalar long run variance needs to be estimated, thus also this estimation procedure involves some user choices when used for inference. The null hypothesis of linearity of the cointegrating relationship simply corresponds to the hypothesis that all coefficients corresponding to powers of variables or cross-products of powers of variables are jointly equal to zero, which is tested with the corresponding Wald-type test.[3]

Hong and Phillips (2010) present a modified RESET test for cointegrating regressions with only one regressor. They follow the Ramsey (1969) version of the RESET test in that they use a regression of the OLS residuals on the auxiliary regressors, i.e. the powers of the integrated regressor. Endogeneity of the regressor and error serial correlation necessitate to perform corrections to the standard RESET test statistic to allow for asymptotic chi-squared inference. The calculation of the required correction factors essentially amounts to performing FM-OLS estimation of the regression augmented by the powers.[4] Hong and Wagner (2012) extend, with respect to specification testing, these results in that they allow for multiple integrated regressors and their powers but they do not allow for cross-products of (powers of) the regressors. In addition to an LM-type test based

[2]The IM-OLS estimator, as seen below, is simpler to implement than other modified least squares estimators used in the cointegration literature. For the fully modified OLS (FM-OLS) estimator of Phillips and Hansen (1990) additive correction factors involving long run and half long run variance matrices need to be calculated. For the dynamic OLS (D-OLS) estimator of Saikkonen (1991) leads and lag choices have to be made and, when used for inference, also a long run variance needs to be estimated.

[3]Although the focus of this paper is on the RESET specification test it should be noted, see Remark 1, that the results in this paper show that the IM-OLS estimator allows for computationally very simple estimation of multivariate cointegrating polynomial relationships as well for asymptotic standard inference using Wald-type tests in this setting. Extending either of the other two mentioned modified OLS estimators to this generality appears to be rather cumbersome, whereas it is straightforward with IM-OLS.

[4]Hong and Phillips (2010) rests upon results for (fully modified) least squares estimation of specific types of nonlinear cointegrating relationships, including polynomial relationships, developed *inter alia* by de Jong (2002), Ibragimov and Phillips (2008) or Park and Phillips (2001) in slightly varying settings.

on an auxiliary regression they also consider a Wald-type test based on an augmented regression.[5] Our test needs no correction factors, is thus simpler to implement and is *more generally applicable* than existing RESET type tests for cointegrating relationships.

As is well-known, standard asymptotic theory does not capture the impact of kernel and bandwidth choices, required 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 cointegration framework by Vogelsang and Wagner (2014), captures the impact of kernel and bandwidth choices on the sampling distributions of HAC-type test statistics.[6] Fixed-b asymptotic theory for the RESET test imposes some restrictions on the set of auxiliary regressors, in that a so-called *full design* (see the details in the following section) has to be chosen, which essentially means that all powers and cross-products of powers of the integrated regressors up to the chosen maximal degree have to be included. Exactly as in the linear cointegration case treated in Vogelsang and Wagner (2014), 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 directly be used.

The theoretical analysis is complemented by a small simulation study to assess the finite sample performance of the proposed standard and fixed-b tests. The simulations show that the fixed-b limit theory well describes the distribution of the test statistic. 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 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 tests exhibit power also against smoothly varying logistic alternatives and not only against polynomial alternatives.

The paper is organized as follows: In the following Section 2 we discuss the underlying assumptions as well as the RESET test and its standard and fixed-b asymptotic distributions. Section 3 presents selected results from a simulation evaluation of the developed IM-OLS RESET test and Section 4 briefly summarizes and concludes. All proofs are relegated to the appendix. Supplementary material available upon request provides tables with the fixed-b critical values.

[5] Hong and Wagner (2012) consider for their LM-type test the FM-OLS residuals from the cointegrating regression rather than the OLS residuals considered by Hong and Phillips (2010). This essentially affects the precise form of the required correction factors. Another difference is that Hong and Wagner (2012) derive their specification test statistics for the null hypothesis of a cointegrating polynomial relationship of arbitrary form (without cross-terms), of which the linearity null hypothesis is only one special case. Clearly, the present paper extends the scope of such test procedures, compare Footnote 3, to general multivariate polynomial relationships.
Let us note in addition that there exist also some tests for the null hypothesis of nonlinear cointegration of a specific form, see, e.g., Choi and Saikkonen (2010) who present a (subsampling) KPSS-type test and Kasparis (2008) who presents a CUSUM type test for additively separable nonlinear cointegrating relationships, i.e. functions where in each of the nonlinear components only one integrated regressor appears.

[6] Additional fixed-b treatments of cointegration models are contained in Bunzel (2006) and Jin *et al.* (2006).

2 Model, Assumptions and the RESET Test

We consider a similar setup as Vogelsang and Wagner (2014), i.e. a cointegrating regression of the form

$$y_t = D_t'\delta + X_t'\beta + u_t \tag{1}$$

$$X_t = X_{t-1} + v_t, \tag{2}$$

with $\eta_t = [u_t, v_t']'$ fulfilling a functional central limit theorem (FCLT) of the form

$$T^{-1/2}\sum_{t=1}^{[rT]} \eta_t \Rightarrow B(r) = \Omega^{1/2}W(r), \quad r \in [0,1], \tag{3}$$

where $[rT]$ denotes the integer part of rT, $W(r)$ is a $(k+1)$-dimensional vector of independent standard Brownian motions and

$$\Omega = \sum_{j=-\infty}^{\infty} \mathbb{E}(\eta_t\eta_{t-j}') = \begin{bmatrix} \Omega_{uu} & \Omega_{uv} \\ \Omega_{vu} & \Omega_{vv} \end{bmatrix} > 0, \tag{4}$$

where clearly $\Omega_{vu} = \Omega_{uv}'$. In case that $\Omega_{uv} \neq 0$ the regressors are endogenous and in addition to regressor endogeneity the setting also allows for relatively unrestricted forms of serial correlation of the errors η_t. These two aspects in general necessitate some form of modified least squares estimation in conjunction with HAC arguments 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 the Cholesky decomposition of Ω, that

$$\begin{bmatrix} B_u(r) \\ B_v(r) \end{bmatrix} = \begin{bmatrix} \omega_{u\cdot v}^{1/2} & \Omega_{uv}(\Omega_{vv}^{-1/2})' \\ 0 & \Omega_{vv}^{1/2} \end{bmatrix} \begin{bmatrix} w_{u\cdot v}(r) \\ W_v(r) \end{bmatrix}, \tag{5}$$

with $\omega_{u\cdot v} = \Omega_{uu} - \Omega_{uv}\Omega_{vv}^{-1}\Omega_{vu}$.

With respect to the deterministic component D_t we merely assume 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

$$\lim_{T\to\infty} \sqrt{T}G_D^{-1}D_{[rT]} = D(r) \ \text{ with } \ 0 < \int_0^r D(z)D(z)'dz < \infty. \tag{6}$$

If, e.g., $D_t = (1, t, t^2, ..., t^{p-1})'$, then $G_D = \text{diag}(T^{1/2}, T^{3/2}, T^{5/2}, \ldots, T^{p-1/2})$ and $D(z) = (1, z, z^2, ..., z^{p-1})'$.

As discussed in the introduction, we want to test for correct specification of the above linear cointegrating relationship (1). The proposed test statistic follows the Thursby and Schmidt (1977) version of the RESET test originally proposed by Ramsey (1969), i.e., it is based on an augmented regression including powers and cross-products of powers of the regressors. The specific choice of auxiliary regressors to be included is to be made by the user. The general form of the augmented regression that underlies the test statistic can be written as:

$$\begin{aligned} y_t &= D_t'\delta + X_t'\beta + \sum_{(p_1,\ldots,p_k)\in\mathcal{I}} \vartheta_{p_1,\ldots,p_k} x_{1t}^{p_1}\cdots x_{kt}^{p_k} + u_t \\ &= D_t'\delta + X_t'\beta + M_t'\vartheta + u_t, \end{aligned} \tag{7}$$

where the RESET null hypothesis of linear cointegration is given by $H_0 : \vartheta = 0$.[7] The set $\mathcal{I}$ denotes all multi-indices, i.e. combinations of powers p_i of the different regressors $x_{1t}, \ldots, x_{kt}$, such that the corresponding term is included in the chosen set of auxiliary regressors.

Consider as a simple illustration the case of two integrated regressors and as auxiliary regressors all terms with maximal degree two. In this case $\mathcal{I} = \{(2,0),(0,2),(1,1)\}$, $M_t = [x_{1t}^2, x_{2t}^2, x_{1t}x_{2t}]'$ and $\vartheta \in \mathbb{R}^3$.

Developing a test statistic for $H_0 : \vartheta = 0$ requires an estimator of the parameters in (7) with a limiting distribution that allows for asymptotically pivotal inference. The estimator we use is (a straightforward extension of) the IM-OLS estimator developed in Vogelsang and Wagner (2014) for linear cointegrating relationships of the form (1). The extended version of this estimator is based on the partial summed version of (7) augmented by the original X_t:

$$\begin{aligned} S_t^y &= S_t^{D\prime}\delta + S_t^{X\prime}\beta + S_t^{M\prime}\vartheta + X_t'\gamma + S_t^u \\ &= S_t^{\widetilde{X}\prime}\theta + S_t^u, \end{aligned} \tag{8}$$

with $S_t^y = \sum_{j=1}^t y_j$ and S_t^D, S_t^X and S_t^M defined analogously. The OLS estimator of the parameters of this regression is referred to as IM-OLS estimator, i.e. when using obvious matrix notation $S^y = S^{\widetilde{X}}\theta + S^u$ we have

$$\widetilde{\theta} = \left(S^{\widetilde{X}\prime}S^{\widetilde{X}}\right)^{-1} S^{\widetilde{X}\prime}S^y. \tag{9}$$

The basis of the test is the asymptotic distribution of the IM-OLS estimator given in the following proposition, for which we need to define some further quantities first. We define the scaling matrix

$$A_{IM} = \begin{bmatrix} G_D^{-1} & 0 & 0 & 0 \\ 0 & T^{-1}I_k & 0 & 0 \\ 0 & 0 & G_\vartheta^{-1} & 0 \\ 0 & 0 & 0 & I_k \end{bmatrix}, \tag{10}$$

with G_ϑ a diagonal matrix corresponding to the components of ϑ, i.e. the entry corresponding to the regressor $x_{1t}^{p_1}x_{2t}^{p_2}\cdots x_{kt}^{p_k}$ is given by $T^{\frac{p_1+\cdots+p_k+1}{2}}$.

Proposition 1 *Assume that the data are generated by* (1) *and (2) with the FCLT* (3) *in place with* $\Omega > 0$ *and the deterministic components fulfilling* (6)*. Define* $\theta^0 = [\delta', \beta', 0, (\Omega_{vv}^{-1}\Omega_{vu})']'$*, then as* $T \to \infty$ *it holds that*

$$\begin{aligned} A_{IM}^{-1}\left(\widetilde{\theta} - \theta^0\right) &\Rightarrow \omega_{u\cdot v}^{1/2}\left(\int f(s)f(s)'ds\right)^{-1}\int f(s)w_{u\cdot v}(s)ds \\ &= \omega_{u\cdot v}^{1/2}\left(\int f(s)f(s)'ds\right)^{-1}\int [F(1)-F(s)]dw_{u\cdot v}(s), \end{aligned} \tag{11}$$

[7] If interpreted as data generating process rather than as augmented regression, with arbitrary values of the coefficient vector ϑ and with X_t and u_t as given in the assumptions, then (7) is referred to as *multivariate* cointegrating polynomial regression, using the nomenclature of Hong and Wagner (2012). A linear cointegrating regression model is, by construction, a special case thereof.

where

$$f(r) = \begin{bmatrix} \int_0^r D(s)ds \\ \int_0^r B_v(s)ds \\ \int_0^r \mathbf{B}_M(s)ds \\ B_v(r) \end{bmatrix}, \quad F(r) = \int_0^r f(s)ds,$$

with $\mathbf{B}_M(r)$ *denoting the limiting process corresponding to the auxiliary regressors, whose precise form depends upon* $\mathcal{I}$.

For our illustrative example it is obvious that $\mathbf{B}_M(r) = [B_{v_1}(r)^2, B_{v_2}(r)^2, B_{v_1}(r)B_{v_2}(r)]'$.

Remark 1 *Clearly, the above result holds true if* (7) *is considered not as augmented regression but as data generating process with* X_t *and* u_t *fulfilling the stated assumptions for any value of* ϑ *and not just for* $\vartheta = 0$ *and thus establishes consistency, with a zero mean Gaussian mixture limiting distribution, of the IM-OLS estimator for the case of multivariate cointegrating polynomial regressions. The IM-OLS limiting distribution allows for hypothesis testing on the parameters, with the null hypothesis of linearity being just one specific hypothesis. The result thereby substantially extends the realm of polynomial relationship type cointegration analysis since the existing literature dealing with this type of relationship imposes additive separability, i.e. does not allow for terms involving cross-products (of powers) of the integrated regressors.*

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

$$V_{IM} = \omega_{u \cdot v} \left(\int_0^1 f(s)f(s)'ds \right)^{-1} \left(\int_0^1 [F(1) - F(s)][F(1) - F(s)]'ds \right) \left(\int_0^1 f(s)f(s)'ds \right)^{-1}. \quad (12)$$

Therefore, the limiting distribution given in (11) allows for asymptotic chi-squared inference to test, e.g., the null hypothesis $H_0 : \vartheta = 0$, given a consistent estimator of $\omega_{u \cdot v}$ to scale out this conditional long run variance from V_{IM}.[8]

For the same reason as in Vogelsang and Wagner (2014), the (first differences) of the IM-OLS residuals $\widetilde{S}_t^u$ cannot be used to consistently estimate $\omega_{u \cdot v}$ by standard long run variance estimation procedures, and in particular it can be shown that their usage is bound to result in a conservative test statistic even asymptotically. A consistent estimator of $\omega_{u \cdot v}$ is most easily obtained using the OLS residuals of (1), $\widehat{u}_t$ say, and by using $\widehat{\eta}_t = [\widehat{u}_t, v_t']'$ to estimate Ω and thereby $\omega_{u \cdot v}$.[9] Given a consistent estimator, $\widehat{\omega}_{u \cdot v}$ say, of $\omega_{u \cdot v}$ an estimator of V_{IM} suggests itself by simply using the sample counterparts of the expressions appearing in the limiting covariance matrix, i.e.

$$\widehat{V}_{IM} = \widehat{\omega}_{u \cdot v} A_{IM}^{-1} \left(S^{\widetilde{X}\prime} S^{\widetilde{X}} \right)^{-1} C'C \left(S^{\widetilde{X}\prime} S^{\widetilde{X}} \right)^{-1} A_{IM}^{-1}, \quad (13)$$

[8] Clearly, compare Remark 1, also other hypotheses can be estimated using exactly the same approach and, e.g., the Wald-type statistic.

[9] Under usual assumptions on kernel and bandwidth this approach allows for consistent estimation of Ω in cointegrating regressions. For detailed discussions and sets of conditions see, e.g., Jansson (2002) or Phillips (1995). We could, for instance, formulate primitive assumptions in terms of linear processes as specified in Jansson (2002), rather than just stating the FCLT (3), and then rely upon Corollary 3 of that paper, since the OLS estimator converges sufficiently fast, in conjunction with the formulated assumptions on kernel (A3) and bandwidth choice (A4). Any set of assumptions, however, under which convergence prevails will of course do for our purposes.

with $C = [c_1, \ldots, c_T]'$, $c_t = S_T^{S^{\widetilde{X}}} - S_{t-1}^{S^{\widetilde{X}}}$ and $S_t^{S^{\widetilde{X}}} = \sum_{j=1}^{t} S_j^{\widetilde{X}}$. Having collected the required quantities we can test the null hypothesis of linearity of the cointegrating relationship by, e.g., the corresponding Wald-type test statistic, i.e.

$$W_{\mathrm{R}} = \widetilde{\vartheta}' \left[R A_{IM} \widehat{V}_{IM} A_{IM} R' \right]^{-1} \widetilde{\vartheta}, \tag{14}$$

where $R = [0, 0, I_{|\mathcal{I}|}, 0]$ is the selection matrix corresponding to $\vartheta \in \theta$, $|\mathcal{I}|$ denotes the number of elements in $\mathcal{I}$ and $I_{|\mathcal{I}|}$ the identity matrix of that size.

In our illustrative example, assuming for concreteness that $D_t = 1$, we have $|\mathcal{I}| = 3$ and $R = [0_{3\times 1}\ 0_{3\times 2}\ I_3\ 0_{3\times 2}]$.

The asymptotic behavior of W_{R} is given next:

Proposition 2 *Suppose the assumptions of Proposition 1 hold and a consistent estimator $\widehat{\omega}_{u\cdot v}$ of $\omega_{u\cdot v}$ is used in $\widetilde{V}_{IM}$, then under the null hypothesis $H_0 : \vartheta = 0$, as $T \to \infty$ it holds that*

$$W_R \xrightarrow{d} \chi^2_{|\mathcal{I}|}, \tag{15}$$

where $\chi^2_{|\mathcal{I}|}$ is a chi-squared distributed random variable with $|\mathcal{I}|$ degrees of freedom.

In order to construct also a fixed-b test statistic it is first necessary to express the limiting distribution of the estimator $\widetilde{\theta}$ as a function involving only powers and products of powers of standard Wiener processes, $\mathbf{W}_M(r)$ say.[10] This imposes some restrictions on the design of the augmented regression. Our usual example can be used to illustrate the issues, where for notational simplicity we denote

$$\Omega_{vv}^{1/2} = \begin{bmatrix} \tau_{11} & \tau_{12} \\ 0 & \tau_{22} \end{bmatrix}.$$

In the example, the limiting process $\mathbf{B}_M(r) = [B_{v_1}(r)^2, B_{v_2}(r)^2, B_{v_1}(r)B_{v_2}(r)]'$ corresponding to the auxiliary regressors is related to a vector comprising powers and cross-products of powers of the elements of $W_v(r)$, the vector $\mathbf{W}_M(r) = [W_{v_1}(r)^2, W_{v_2}(r)^2, W_{v_1}(r)W_{v_2}(r)]'$, as follows:

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

With the assumptions on Ω and because we include *all terms up to the maximal power* in $\mathcal{I}$ in the example, the lower 3–3 block, $\mathbf{\Omega}_M$, of the matrix in the middle has full rank. A bijection between $\mathbf{B}_M(r)$ and $\mathbf{W}_M(r)$ would not prevail, if, e.g., the mixed term $x_{1t}x_{2t}$ were not included in the auxiliary regressor set. We refer to situations in which a bijection between $\mathbf{B}_M(r)$ and $\mathbf{W}_M(r)$

[10] It is clear that in this respect only the "nonlinear regressors" matter, as of course for the other stochastic regressors in the IM-OLS regression, S_t^X and X_t, a bijection of the corresponding limiting processes expressed in terms of $B_v(r)$ and in terms of $W_v(r)$ prevails by construction.

prevails as *full design*.[11] In case of full design we can thus express the IM-OLS limiting distribution as a function of $W(r)$.

Corollary 1 *Suppose the assumptions of Proposition 1 hold and full design prevails, then as $T \to \infty$ it holds that*

$$A_{IM}^{-1}\left(\widetilde{\theta}-\theta^0\right) \Rightarrow \omega_{u\cdot v}^{1/2}\left(\Pi\int g(s)g(s)'ds\Pi'\right)^{-1}\Pi\int g(s)w_{u\cdot v}(s)ds \tag{16}$$

$$= \omega_{u\cdot v}^{1/2}(\Pi')^{-1}\left(\int g(s)g(s)'ds\right)^{-1}\int[G(1)-G(s)]dw_{u\cdot v}(s), \tag{17}$$

where

$$g(r) = \begin{bmatrix} \int_0^r D(s)ds \\ \int_0^r W_v(s)ds \\ \int_0^r \mathbf{W}_M(s)ds \\ W_v(r) \end{bmatrix}, \quad \Pi = \begin{bmatrix} I_p & 0 & 0 & 0 \\ 0 & \Omega_{vv}^{1/2} & 0 & 0 \\ 0 & 0 & \mathbf{\Omega}_M & 0 \\ 0 & 0 & 0 & \Omega_{vv}^{1/2} \end{bmatrix}, \tag{18}$$

with $\mathbf{W}_M(r)$ the vector containing the terms corresponding to the auxiliary regressors expressed as functions of $W_v(r)$.

Clearly, $f(r) = \Pi g(r)$, with this relation being bijective in case of full design. Again similarly to Vogelsang and Wagner (2014) it can be shown that it is not possible to perform asymptotically pivotal fixed-b inference using the IM-OLS residuals, $\widetilde{S}_t^u$, because of non-vanishing and nuisance parameter dependent correlation between the limiting distribution of $A_{IM}^{-1}\left(\widetilde{\theta}-\theta_0\right)$ and the limit process of $T^{-1/2}\sum_{t=2}^{[rT]}\Delta\widetilde{S}_t^u$.

Valid fixed-b inference can, however, be performed by basing the estimator of $\omega_{u\cdot v}$ on a modification of the IM-OLS residuals $\widetilde{S}_t^u$. For doing so it is useful to define

$$z_t = t\sum_{j=1}^{T} S_j^{\widetilde{X}} - \sum_{j=1}^{t-1}\sum_{s=1}^{j} S_s^{\widetilde{X}}, \tag{19}$$

and $z_t^{\perp}$ as the vector of residuals from individually regressing each element of z_t on $S_t^{\widetilde{X}}$. The adjusted residuals $\widetilde{S}_t^{u*}$ are then given by the residuals of the OLS regression of $\widetilde{S}_t^u$ on $z_t^{\perp}$, i.e. by

$$\widetilde{S}_t^{u*} = \widetilde{S}_t^u - z_t^{\perp\prime}\widehat{\pi}, \tag{20}$$

with $\widehat{\pi} = \left(\sum_{t=1}^T z_t^{\perp}z_t^{\perp\prime}\right)^{-1}\sum_{t=1}^T z_t^{\perp}\widetilde{S}_t^u$.[12] For $\widetilde{S}_t^{u*}$ it can be shown that $T^{-1/2}\sum_{t=2}^{[rT]}\Delta\widetilde{S}_t^{u*}$ is, conditional upon $W_v(r)$, asymptotically independent of $A_{IM}^{-1}\left(\widetilde{\theta}-\theta^0\right)$. Consequently, using a long run

[11]For most situations this is equivalent to full rank of $\mathbf{\Omega}_M$ and a bijection between $\mathbf{B}_M(r)$ and $\mathbf{W}_M(r)$ that contain similar powers and cross-products of powers of the elements of $B_v(r)$ and $W_v(r)$, respectively. Of course, it is, if so desired, possible to achieve full design in the augmented regression by simply including all necessary regressors in the augmented regression. By construction this restricts the flexibility of specifying the augmented regression but allows for fixed-b inference.

[12]Clearly, the required residuals $\widetilde{S}_t^{u*}$ can be obtained together with the (unchanged) IM-OLS parameter estimates $\widetilde{\theta}$ in one step by augmenting the IM-OLS regression (7) by $z_t^{\perp}$.

variance estimator

$$\widetilde{\omega}_{u\cdot v}^* = T^{-1} \sum_{i=2}^{T} \sum_{j=2}^{T} k\left(\frac{|i-j|}{B}\right) \Delta \widetilde{S}_i^{u*} \Delta \widetilde{S}_j^{u*} \tag{21}$$

with kernel function $k(\cdot)$ and bandwidth B allows for asymptotically pivotal fixed-b inference for the RESET null hypothesis in case of full design.

In order to efficiently describe the fixed-b limiting distributions of the RESET statistic define for a stochastic process $P(r)$ the random variable $Q(P)$ as follows. In case that $k(\cdot)$ 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

$$\begin{aligned} Q(P) &= -\frac{1}{b^2} \int_0^1 \int_0^1 k''\left(\frac{|r-s|}{b}\right) P(s)P(r)' ds dr + \frac{1}{b} \int_0^1 k'\left(\frac{|1-s|}{b}\right) (P(1)P(s)' + P(s)P(1)')ds \\ &\quad + P(1)P(1)'. \end{aligned} \tag{22}$$

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

$$\begin{aligned} Q(P) &= \frac{2}{b} \int_0^1 P(s)P(s)' ds - \frac{1}{b} \int_0^{1-b} (P(s)P(s+b)' + P(s+b)P(s)')ds \\ &\quad - \frac{1}{b} \int_{1-b}^1 (P(1)P(s)' + P(s)P(1)')ds + P(1)P(1)'. \end{aligned} \tag{23}$$

With the necessary notation collected we can now give the result for the fixed-b RESET test statistic.

Proposition 3 *Suppose the assumptions of Proposition 1 hold, the augmented regression has full design and the null hypothesis $H_0 : \vartheta = 0$ is correct. Furthermore, $\widetilde{\omega}_{u\cdot v}^*$ is given as in (21) with $B = bT$, where $b \in (0, 1]$ is held fixed as $T \to \infty$, then as $T \to \infty$ it holds that*

$$W_R^* = \widetilde{\vartheta}' [R A_{IM} \widetilde{V}_{IM}^* A_{IM} R']^{-1} \widetilde{\vartheta} \Rightarrow \frac{\chi_{|\mathcal{I}|}^2}{Q(\widetilde{P}^*)}, \tag{24}$$

with $\widetilde{V}_{IM}^*$ as given in (13) with $\widehat{\omega}_{u\cdot v}$ replaced by $\widetilde{\omega}_{u\cdot v}^*$ and

$$\widetilde{P}^*(r) = \int_0^r dw_{u\cdot v}(s) - h(r)' \left(\int_0^1 h(s)h(s)' ds\right)^{-1} \int_0^1 [H(1) - H(s)] dw_{u\cdot v}(s), \tag{25}$$

where

$$h(r) = \left[g(r)', \int_0^1 [G(1) - G(s)]' ds\right]', \quad H(r) = \int_0^r h(s) ds.$$

Here $\chi_{|\mathcal{I}|}^2$ is a chi-squared distributed random variable with $|\mathcal{I}|$ degrees of freedom independent of $Q(\widetilde{P}^*)$. The precise form of $Q(\widetilde{P}^*)$ depends upon the kernel chosen and is given by (22) or (23) if the kernel satisfies the respective assumptions.

Critical values based on simulating the limiting distribution given in the above proposition are available in tabulated form for $k = 2, 3, 4$ regressors for the full design quadratic and cubic specifications of the RESET test for the cases with intercept only and with intercept and linear trend. These critical values are available for five different kernels (Bartlett, Bohman, Daniell, Parzen and Quadratic Spectral) for a grid of 50 values of $b = 0.02, 0.04, \ldots, 0.98, 1$. Supplementary material available upon request contains the tables with the corresponding critical values.[13]

Let us close this section with some reflections on the power properties of the IM-OLS RESET test, which are by construction conceptually similar to, e.g., the power properties discussed for their univariate RESET test in Hong and Phillips (2010). First, the tests do not only have power against nonlinear cointegration but also have power against the alternative of no cointegration between y_t and X_t. In that case spurious regression asymptotics leads to non-zero limits of the estimated coefficients. Second, by construction power will be the higher, the better the augmented regression captures the true unknown function and will be particularly high in case the augmented regression coincides with a true unknown multivariate cointegrating polynomial relationship. Thus, the choice of auxiliary regressors is bound to have impacts on test performance. Third, power can be expected to be lower against smoothly varying alternatives, like logistic functions, which dampen the fluctuations of the integrated regressors rather than exacerbating them like polynomial transformations, and which are furthermore not necessarily well approximated by low order polynomials. The discussion of power properties in the following simulation section illustrates these properties.[14]

3 Finite Sample Performance

We generate data under the null hypothesis of linear cointegration according to the data generating process (DGP)

$$\begin{aligned} y_t &= \delta + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t, \\ x_{it} &= x_{i,t-1} + v_{it}, \quad x_{i0} = 0, \quad i = 1, 2, \end{aligned} \tag{26}$$

where

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

where ε_t, e_{1t} and e_{2t} are i.i.d. standard normal random variables independent of each other. The parameter values chosen are $\delta = 3$, $\beta_1 = \beta_2 = 1$, where we note that the values of these parameters have no effect on the results because the IM-OLS estimator corresponding to the RESET test is exactly invariant to the values of δ, β_1 and β_2. The values for ρ_1 and ρ_2 are chosen from the set $\{0, 0.3, 0.6, 0.8\}$. The parameter ρ_1 controls serial correlation in the regression error u_t, whereas the parameter ρ_2 controls whether the regressors are endogenous or not. The kernels chosen for long run variance estimation are the Bartlett and the Quadratic Spectral (QS) kernels and the results are reported for bandwidths using the grid $M = bT$ with $b \in \{0.02, 0.04, ..., 0.98, 1.0\}$ or for

[13] Furthermore, also `MATLAB` code implementing the test procedure, with both classical and fixed-b inference, is available.

[14] Because power in the no cointegration case is well documented in many settings and behaves similarly for our test, we do not include the no cointegration case in our set of alternatives simulated.

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 simulations we label W_{R} as IM(O) and W_{R}^* 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 partial summed and augmented by the original x_{1t} and x_{2t} for IM-OLS estimation):

$$\begin{aligned} y_t &= \delta + x_{1t}\beta_1 + x_{2t}\beta_2 + x_{1t}^2\vartheta_1 + x_{2t}^2\vartheta_2 + x_{1t}x_{2t}\vartheta_3 + u_t \quad (q=2), \\ y_t &= \delta + x_{1t}\beta_1 + x_{2t}\beta_2 + x_{1t}^2\vartheta_1 + x_{2t}^2\vartheta_2 + x_{1t}x_{2t}\vartheta_3 + x_{1t}^3\vartheta_4 + x_{2t}^3\vartheta_5 \\ &\quad + x_{1t}^2x_{2t}\vartheta_6 + x_{1t}x_{2t}^2\vartheta_7 + u_t \quad (q=3), \end{aligned}$$

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 considered alternatives. The IM(O) tests are based on using the OLS residuals of the above augmented regressions to obtain a consistent estimator $\widehat{\omega}_{u\cdot v}$ 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 $\Delta\widetilde{S}_t^{u*}$ for estimating $\omega_{u\cdot v}$ as given in (21).

We first present null rejection probabilities for IM(fb) for the grid of 50 bandwidth values. These null rejection probabilities are computed using the relevant fixed-b critical values depending on the kernel, the bandwidth sample size ratio (b) and q. Figures 1 and 2 plot null rejection probabilities for the Bartlett and QS kernels respectively for the case of $q = 2$ 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 doing a good job of 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 the case of $\rho_1 = \rho_2 = 0.3$. 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 gives 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 gives a test with substantially less over-rejection problems especially if T is not small and/or the bandwidth is not too small.

[15] A recent paper by Sun (2014) has shown, 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 testing increases when chi-squared critical values are used for the test. While 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.

When $T = 500$, the QS test is much less sensitive to q than is 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.

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

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

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 (26). 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 considered sample size. 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

[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 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 I(1) regressors. It would be very challenging but worthwhile to extend the Sun *et al.* (2008) approach to the cointegration setting.

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

slightly higher power in some cases. As expected power is highest in case of 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.

While Tables 2 and 3 have useful information, they give no indication of the shape of the power curves. Figures 7-10 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 very slightly 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 trade-off in reduction of 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, $G(X_t) = x_{1t}\left(1 + e^{-x_{1t}}\right)^{-1}$. While 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 in the case of the logistic alternative compared to the case of polynomial alternatives stems from the fact that larger values of ϕ do not generate as big 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 linear cointegration settings by Vogelsang and Wagner (2014).

4 Summary and Conclusions

We have presented an easy-to-use RESET test for the null hypothesis of linear cointegration. Given that economic theory typically does not specify the exact form of a potential nonlinear cointegrating relationship under the alternative, an omnibus specification test like the RESET test may be a useful tool for applied work, just like the widely-used original RESET test in standard regression settings. The test is based on the IM-OLS estimator of Vogelsang and Wagner (2014), which, as is shown in the paper, straightforwardly extends to the multivariate cointegrating polynomial regression case. The fact that the extension is so easy is a unique feature of the IM-OLS approach and similar extensions appear much more cumbersome with competing modified least squares estimators.

The RESET test is performed as a Wald-type test in the regression augmented by powers and cross-products of powers of the integrated regressors that is estimated by IM-OLS. By construction such a test has power not only against nonlinear alternatives but also power against the spurious regression alternative. The null of linearity is easily tested by testing whether all coefficients to the nonlinear terms are jointly equal to zero. Standard asymptotic theory delivers an asymptotic chi-squared null limiting distribution of the RESET test. In order to capture the impact of kernel and bandwidth choices on the sampling distribution of the test statistics we also develop fixed-b asymptotic theory for the full design case. Fixed-b asymptotic theory rests upon long run variance estimation using specifically modified IM-OLS residuals since pivotal fixed-b limiting distributions cannot be based on the IM-OLS residuals directly.

The finite sample 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 in against 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 Austrian Central Bank (grant 15334). The second author furthermore gratefully acknowledges financial support from Deutsche Forschungsgemeinschaft via the Collaborative Research Center 823: *Statistical Modelling of Nonlinear Dynamic Processes* (Projects A3 and A4). The authors are particularly indebted to Oliver Stypka for superb research assistance. The usual disclaimer applies.

References

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

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

Bec, F. and Rahbek, A.: (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.

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

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

de Jong, R.: (2002), Nonlinear estimators with integrated regressors but without exogeneity, Mimeo.

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

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

Hong, S. H. and Wagner, M.: (2012), Cointegrating polynomial regressions: fully modified OLS estimation and inference, Mimeo.

Ibragimov, R. and Phillips, P. C. B.: (2008), Regression asymptotics using martingale convergence methods, *Econometric Theory* **24**, 888–947.

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

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

Karlsen, H. A., Myklebust, T. and Tjøstheim, D.: (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 Vogelsang, T. J.: (2005), A new asymptotic theory for heteroskedasticity-autocorrelation robust tests, *Econometric Theory* **21**, 1130–1164.

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

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

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

Park, J.Y. and Phillips, P. C. B.: (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*, Cambridge University Press, Cambridge, 123–167.

Phillips, P. C. B.: (1995), Fully modified least squares and vector autoregression, *Econometrica* **59**, 1023–1078.

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

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

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 Jin, S.: (2008), Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing, *Econometrica* **76**, 175–194.

Thursby, J. G. and Schmidt, P.: (1977), Some properties of tests for specification error in a linear regression model, *Journal of the American Statistical Association* **72**, 635–641.

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

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

Wagner, M.: (2014), The environmental Kuznets curve, cointegration and nonlinearity, *Journal of Applied Econometrics*, forthcoming (DOI: 10.1002/jae.2421).

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

Appendix: Proofs

Proof of Proposition 1

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

$$\begin{aligned} A_{IM}^{-1}\left(\widetilde{\theta}-\theta^0\right) &= \left(T^{-2}\left(A_{IM}S^{\widetilde{X}\prime}\right)S^{\widetilde{X}}A_{IM}\right)^{-1}T^{-2}\left(A_{IM}S^{\widetilde{X}\prime}\right)S^u-[0,0,0,(\Omega_{vv}^{-1}\Omega_{vu})']' \\ &= \left(T^{-1}\left(T^{-1/2}A_{IM}S^{\widetilde{X}\prime}\right)\left(T^{-1/2}S^{\widetilde{X}}A_{IM}\right)\right)^{-1}T^{-1}\left(T^{-1/2}A_{IM}S^{\widetilde{X}\prime}\right)\left(T^{-1/2}S^u\right) \\ &\quad -[0,0,0,(\Omega_{vv}^{-1}\Omega_{vu})']'. \end{aligned} \tag{27}$$

Because $T^{-1/2}S^u_{[rT]} \Rightarrow B_u(r)$, the limit of $A_{IM}^{-1}\left(\widetilde{\theta}-\theta^0\right)$ follows by the continuous mapping theorem (CMT) once we obtain the limit of $T^{-1/2}A_{IM}S^{\widetilde{X}}_{[rT]}$. Breaking $T^{-1/2}A_{IM}S^{\widetilde{X}}_{[rT]}$ into its components we have

$$T^{-1/2}A_{IM}S^{\widetilde{X}}_{[rT]} = \left[\begin{array}{c} T^{-1}\sum_{t=1}^{[rT]}\sqrt{T}G_D^{-1}D_t \\ T^{-3/2}\sum_{t=1}^{[rT]}X_t \\ T^{-1/2}G_\vartheta^{-1}\sum_{t=1}^{[rT]}M_t \\ T^{-1/2}\sum_{t=1}^{[rT]}v_t \end{array}\right].$$

The limits of the first, second and fourth sub-vectors follow from (6) and (3). To establish the limit of $T^{-1/2}G_\vartheta^{-1}\sum_{t=1}^{[rT]}M_t$, consider a typical element given by

$$\begin{aligned} T^{-1/2}T^{-\frac{p_1+\cdots+p_k+1}{2}}\sum_{t=1}^{[rT]}x_{1t}^{p_1}x_{2t}^{p_2}\cdots x_{kt}^{p_k} &= T^{-1}\sum_{t=1}^{[rT]}\left(T^{-p_1/2}x_{1t}^{p_1}\right)\left(T^{-p_2/2}x_{2t}^{p_2}\right)\cdots\left(T^{-p_k/2}x_{kt}^{p_k}\right) \\ &= T^{-1}\sum_{t=1}^{[rT]}\left(T^{-1/2}x_{1t}\right)^{p_1}\left(T^{-1/2}x_{2t}\right)^{p_2}\cdots\left(T^{-1/2}x_{kt}\right)^{p_k} \\ &\Rightarrow \int_0^r B_{v_1}(r)^{p_1}B_{v_2}(r)^{p_2}\cdots B_{v_k}(r)^{p_k}dr, \end{aligned}$$

where $B_{v_1}(r)^{p_1} B_{v_2}(r)^{p_2} \cdots B_{v_k}(r)^{p_k}$ is the corresponding element of $\mathbf{B}_M(s)$. Therefore, it follows that $T^{-1/2} G_\vartheta^{-1} \sum_{t=1}^{[rT]} M_t \Rightarrow \int_0^r \mathbf{B}_M(s) ds$ and we have

$$T^{-1/2} A_{IM} S_{[rT]}^{\widetilde{X}} \Rightarrow f(r), \tag{28}$$

which in turn gives

$$\begin{aligned}
&\left(T^{-2}\left(A_{IM} S^{\widetilde{X}\prime}\right)\left(S^{\widetilde{X}} A_{IM}\right)\right)^{-1}\left(T^{-2}\left(A_{IM} S^{\widetilde{X}\prime}\right) S^{u}\right) \Rightarrow \left(\int f(s) f(s)' ds\right)^{-1} \int f(s) B_u(s) ds \\
&\quad = \left(\int f(s) f(s)' ds\right)^{-1}\left(\omega_{u \cdot v}^{1/2} \int f(s) w_{u \cdot v}(s) ds + \int f(s) B_v(s)' \Omega_{vv}^{-1} \Omega_{vu} ds\right) \\
&\quad = \left(\int f(s) f(s)' ds\right)^{-1} \omega_{u \cdot v}^{1/2} \int f(s) w_{u \cdot v}(s) ds + [0,0,0,(\Omega_{vv}^{-1}\Omega_{vu})']', \qquad (29)
\end{aligned}$$

where the second line follows from $B_u(s) = \omega_{u \cdot v}^{1/2} w_{u \cdot v}(s) + B_v(s)' \Omega_{vv}^{-1} \Omega_{vu}$ and the third line follows because $B_v(s)$ the vector of the last k elements of $f(s)$, which implies $\left(\int f(s) f(s)' ds\right)^{-1} \int f(s) B_v(s)' ds = [0,0,0,I_k]'$. The proposition follows from (27) and (29) with the representation in terms of $dw_{u \cdot v}(s)$ easily calculated using integration by parts.

Proof of Proposition 2

Manipulating the expression for W_{R} gives

$$\begin{aligned}
W_{\mathrm{R}} &= \widetilde{\vartheta}' [R A_{IM} \widetilde{V}_{IM} A_{IM} R']^{-1} \widetilde{\vartheta} = \left(G_\vartheta \widetilde{\vartheta}\right)' [R \widetilde{V}_{IM} R']^{-1} \left(G_\vartheta^{-1} \widetilde{\vartheta}\right) \\
&= \left[R A_{IM}^{-1}\left(\widetilde{\theta} - \theta^0\right)\right]' [R \widetilde{V}_{IM} R']^{-1} \left[R A_{IM}^{-1}\left(\widetilde{\theta} - \theta^0\right)\right].
\end{aligned}$$

Using consistency of $\widehat{\omega}_{u \cdot v}$, the limit in (28), and arguments used by Vogelsang and Wagner (2014) in the proof of their Theorem 3, it follows that

$$\widetilde{V}_{IM} \Rightarrow V_{IM},$$

and the chi-squared limit for W_{R} follows by the mixture normal limit of $A_{IM}^{-1}\left(\widetilde{\theta} - \theta^0\right)$ in Proposition 1, upon conditioning on $B_v(s)$, i.e. upon conditioning on $f(s)$.

Proof of Corollary 1

Rewrite $f(r)$ as

$$f(r) = \begin{bmatrix} \int_0^r D(s) ds \\ \int_0^r B_v(s) ds \\ \int_0^r \mathbf{B}_M(s) ds \\ B_v(r) \end{bmatrix} = \begin{bmatrix} \int_0^r D(s) ds \\ \Omega_{vv}^{1/2} \int_0^r W_v(s) ds \\ \Omega_M \int_0^r \mathbf{W}_M(s) ds \\ \Omega_{vv}^{1/2} B_v(r) \end{bmatrix} = \Pi g(r),$$

and the corollary follows immediately.

Proof of Proposition 3

The key calculation is to derive the limit process of $T^{-1/2}\widetilde{S}^{u*}_{[rT]} = T^{-1/2}\sum_{t=2}^{[rT]}\Delta\widetilde{S}^{u*}_t$. Once this limit has been obtained, the limits for $\widetilde{\omega}_{u\cdot v}$ and W^*_{R} follow using similar arguments as in Vogelsang and Wagner (2014) and details are omitted. It is clear that $\widetilde{S}^{u*}_t$ is equivalently given by the OLS residuals from the regression of S^y_t on $S^{\widetilde{X}}_t$ and z_t. Letting $S^*_t = \left[S^{\widetilde{X}\prime}_t, z'_t\right]'$ we can therefore write

$$\widetilde{S}^{u*}_t = S^u_t - S^{*\prime}_t\left(\sum_{t=1}^T S^*_t S^{*\prime}_t\right)^{-1}\sum_{t=1}^T S^*_t S^u_t.$$

The limit of $T^{-1/2}A_{IM}S^{\widetilde{X}}_{[rT]}$ is given in the proof of Proposition 1 and thus we next show that $T^{-5/2}A_{IM}z_{[rT]}$ has a well defined limit:

$$\begin{aligned}
T^{-5/2}A_{IM}z_{[rT]} &= \frac{[rT]}{T}T^{-1}\sum_{j=1}^T T^{-1/2}A_{IM}S^{\widetilde{X}}_j - T^{-1}\sum_{j=1}^{[rT]-1}T^{-1}\sum_{i=1}^j T^{-1/2}A_{IM}S^{\widetilde{X}}_i\\
&\Rightarrow r\int_0^1 f(s)ds - \int_0^r\left[\int_0^s f(v)dv\right]ds = \int_0^r F(1)ds - \int_0^r F(s)ds\\
&= \int_0^r\left[F(1) - F(s)\right]ds.
\end{aligned}$$

Defining

$$A^*_{IM} = \begin{bmatrix} A_{IM} & 0\\ 0 & T^{-2}A_{IM}\end{bmatrix},$$

it follows that

$$\begin{aligned}
T^{-1/2}A^*_{IM}S^*_{[rT]} &= \begin{bmatrix} T^{-1/2}A_{IM}S^{\widetilde{X}}_{[rT]}\\ T^{-5/2}A_{IM}z_{[rT]}\end{bmatrix} \Rightarrow \begin{bmatrix} f(r)\\ \int_0^r\left[F(1)-F(s)\right]ds\end{bmatrix}\\
&= \begin{bmatrix} \Pi g(r)\\ \Pi\int_0^r\left[G(1)-G(s)\right]ds\end{bmatrix} = \Pi^* h(r),
\end{aligned}\tag{30}$$

where $\Pi^* = \begin{bmatrix}\Pi & 0\\ 0 & \Pi\end{bmatrix}$. Using (30) we easily obtain

$$\begin{aligned}
T^{-1/2}\widetilde{S}^{u*}_{[rT]} &= T^{-1/2}S^u_{[rT]}\\
&\quad - T^{-1/2}A^*_{IM}S^*_{[rT]}\left(T^{-1}\sum_{t=1}^T\left(T^{-1/2}A^*_{IM}S^*_t\right)\left(T^{-1/2}A^*_{IM}S^*_t\right)'\right)^{-1}T^{-1}\sum_{t=1}^T\left(T^{-1/2}A^*_{IM}S^*_t\right)\left(T^{-1/2}S^u_t\right)\\
&\quad\Rightarrow B_u(r) - \Pi^*h(r)\left(\Pi^*\int_0^1 h(s)h(s)'ds\Pi^*\right)^{-1}\Pi^*\int_0^1 h(s)B_u(s)ds\\
&\quad= B_u(r) - h(r)\left(\int_0^1 h(s)h(s)'ds\right)^{-1}\int_0^1 h(s)B_u(s)ds.
\end{aligned}\tag{31}$$

Now recall that $B_u(s) = \omega_{u\cdot v}^{1/2} w_{u\cdot v}(s) + W_v(s)'\Omega_{vv}^{-1/2}\Omega_{vu}$ and therefore, because $W_v(s)$ is contained in $h(r)$, the $W_v(s)$ component of $B_u(s)$ is projected out of (31) giving

$$\begin{aligned} T^{-1/2}\widetilde{S}_{[rT]}^{u*} &\Rightarrow \omega_{u\cdot v}^{1/2}\left(\int_0^r dw_{u\cdot v}(s) - h(r)\left(\int_0^1 h(s)h(s)'ds\right)^{-1}\int_0^1 h(s)w_{u\cdot v}(s)ds\right) \qquad (32) \\ &= \omega_{u\cdot v}^{1/2}\left(\int_0^r dw_{u\cdot v}(s) - h(r)\left(\int_0^1 h(s)h(s)'ds\right)^{-1}\int_0^1 \left[H(1)-H(s)\right]dw_{u\cdot v}(s)\right) \\ &= \omega_{u\cdot v}^{1/2}\widetilde{P}^*(r). \end{aligned}$$

The independence of the $\chi^2_{|\mathcal{I}|}$ random variable and $\widetilde{P}^*(r)$ is established as follows: We show that, conditional on $W_v(r)$, $\chi^2_{|\mathcal{I}|}$ and $\widetilde{P}^*(r)$ are independent. If that is established, because $\chi^2_{|\mathcal{I}|}$ is independent of $W_v(r)$, this implies that $\chi^2_{|\mathcal{I}|}$ and $\widetilde{P}^*(r)$ are also independent unconditionally.

Thus it remains to establish conditional independence of $\chi^2_{|\mathcal{I}|}$ and $\widetilde{P}^*(r)$. Conditional on $W_v(s)$, $\chi^2_{|\mathcal{I}|}$ and $\widetilde{P}^*(r)$ are Gaussian processes defined in terms of the Gaussian process $w_{u\cdot v}(r)$. Therefore conditional independence is established by showing that the covariance between the two conditional Gaussian processes is zero. For $\chi^2_{|\mathcal{I}|}$, the relevant conditional Gaussian process is $\int[G(1)-G(s)]dw_{u\cdot v}(s)$ and thus it suffices to calculate

$$\begin{aligned} \mathrm{Cov}\left(\int[G(1)-G(s)]dw_{u\cdot v}(s), \widetilde{P}^*(r)\right) = &\int_0^r [G(1)-G(s)]'ds - \qquad (33) \\ &- h(r)'\left(\int h(s)h(s)'\right)^{-1}\int [H(1)-H(s)][G(1)-G(s)]'ds \end{aligned}$$

Let $h_2(r) = \int_0^r [G(1)-G(s)]'ds$. Using integration by parts

$$\int [H(1)-H(s)][G(1)-G(s)]'ds = \int h(s)h_2(s)'ds,$$

and it follows that

$$\mathrm{Cov}\left(\int[G(1)-G(s)]dw_{u\cdot v}(s), \widetilde{P}^*(r)\right) = h_2(r) - h(r)'\left(\int h(s)h(s)'\right)^{-1}\int h(s)h_2(s)'ds = 0,$$

where we obtain zero because $h_2(r)$ is contained in $h(r)$.

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 2: 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

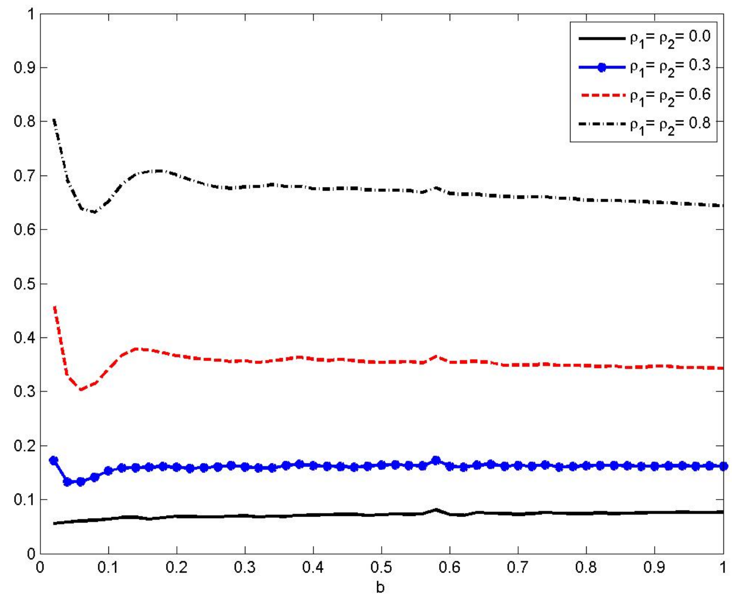


Figure 1: Empirical Null Rejections, IM(fb), $q = 2$, $T = 100$
Bartlett Kernel, 5% Nominal Level

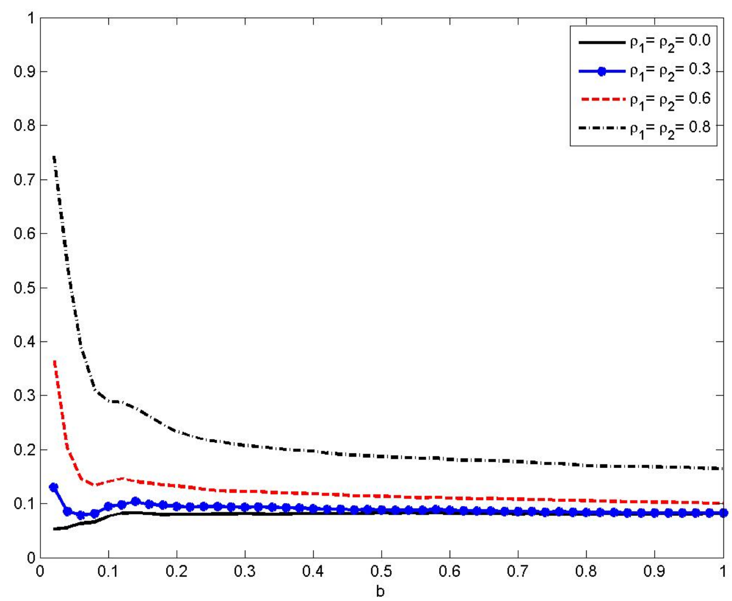


Figure 2: Empirical Null Rejections, IM(fb), $q = 2$, $T = 100$
QS Kernel, 5% Nominal Level

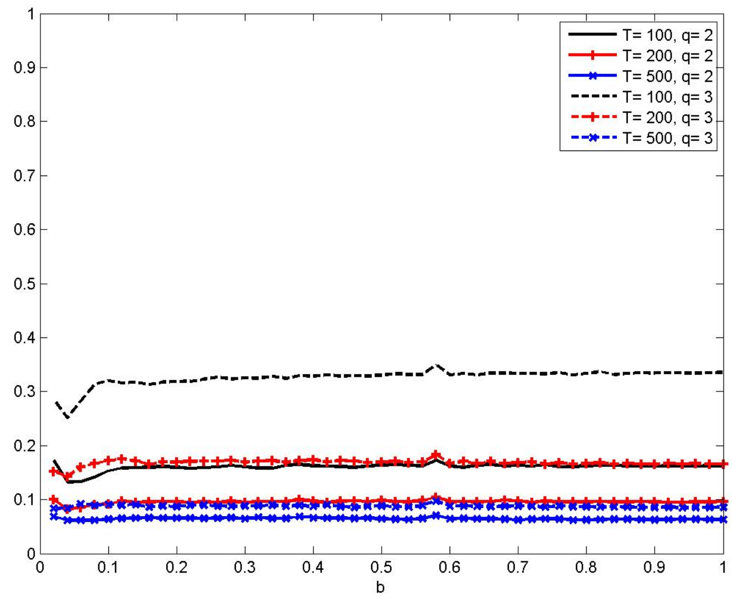


Figure 3: Empirical Null Rejections, IM(fb), $\rho_1 = \rho_2 = .3$
Bartlett Kernel, 5% Nominal Level

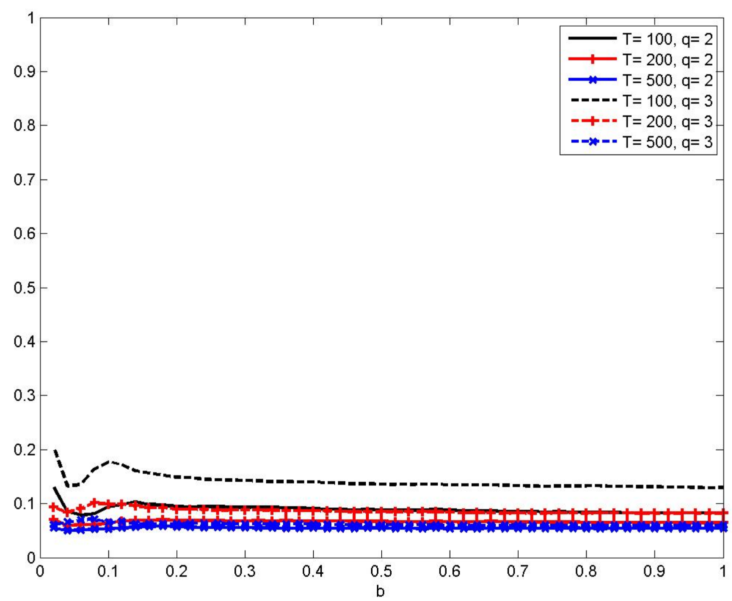


Figure 4: Empirical Null Rejections, IM(fb), $\rho_1 = \rho_2 = .3$
QS Kernel, 5% Nominal Level

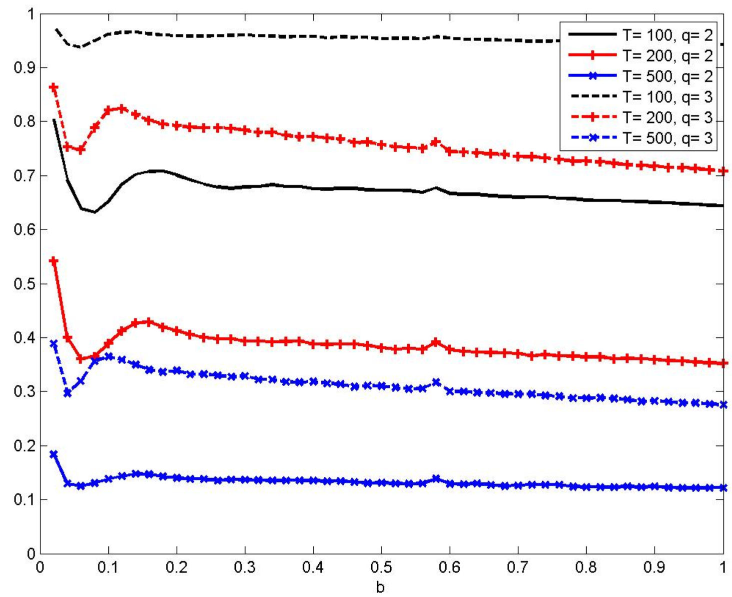


Figure 5: Empirical Null Rejections, IM(fb), $\rho_1 = \rho_2 = .8$
Bartlett Kernel, 5% Nominal Level

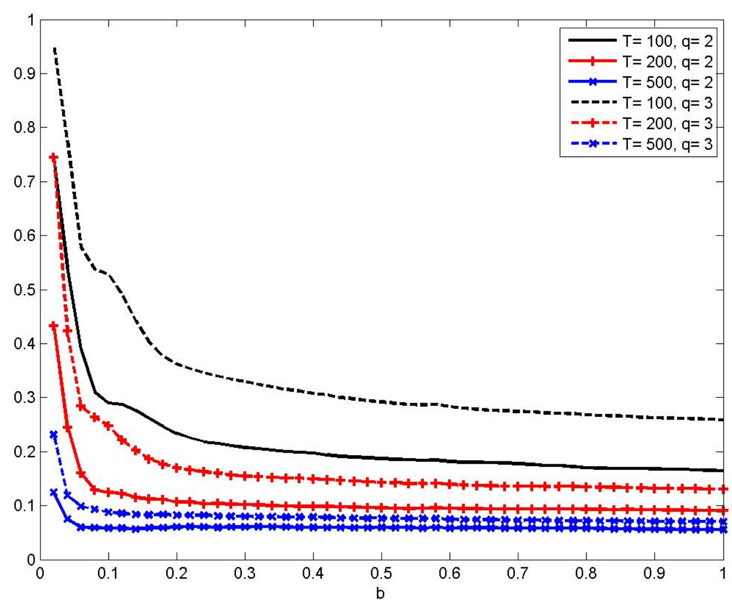


Figure 6: Empirical Null Rejections, IM(fb), $\rho_1 = \rho_2 = .8$
QS Kernel, 5% Nominal Level

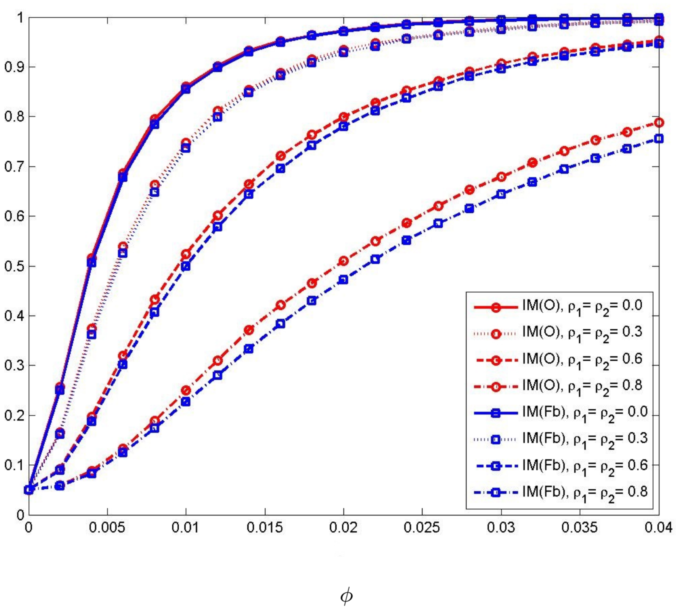


Figure 7: Empirical Size Adjusted Power, $T = 200$, $q = 2$, $G(X_t) = x_{1t}^2$
Bartlett Kernel, Andrews Bandwidth, 5% Nominal Level

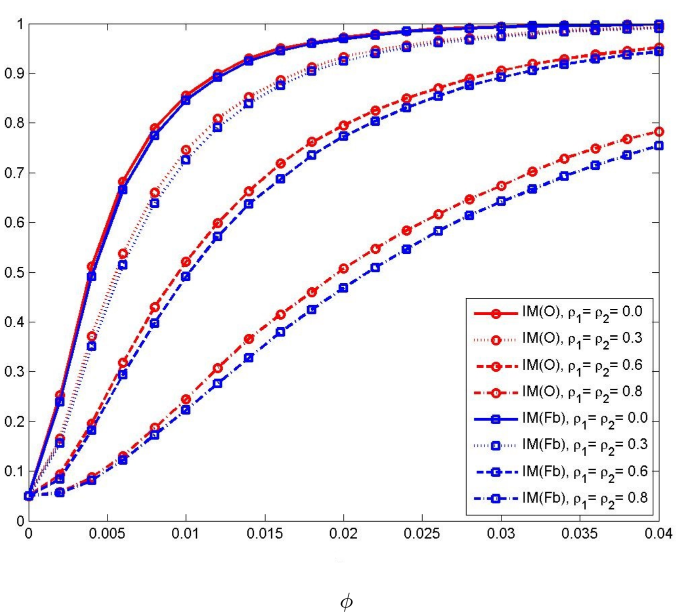


Figure 8: Empirical Size Adjusted Power, $T = 200$, $q = 2$, $G(X_t) = x_{1t}^2$
QS Kernel, Andrews Bandwidth, 5% Nominal Level

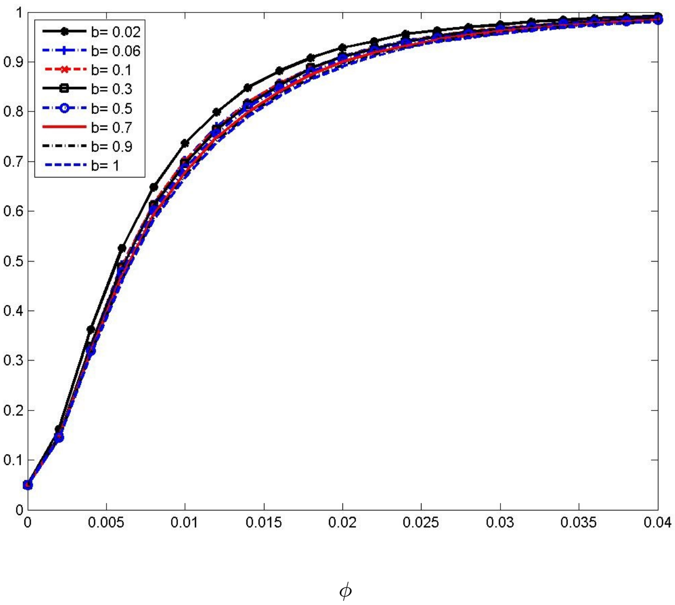


Figure 9: Empirical Size Adjusted Power, IM(fb), $T = 200$, $q = 2$, $\rho_1 = \rho_2 = .3$
$G(X_t) = x_{1t}^2$, Bartlett Kernel, Andrews Bandwidth, 5% Nominal Level

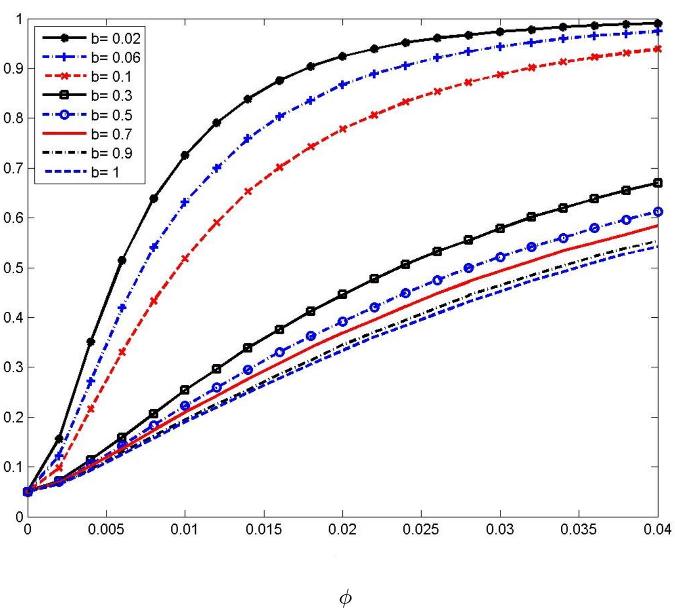
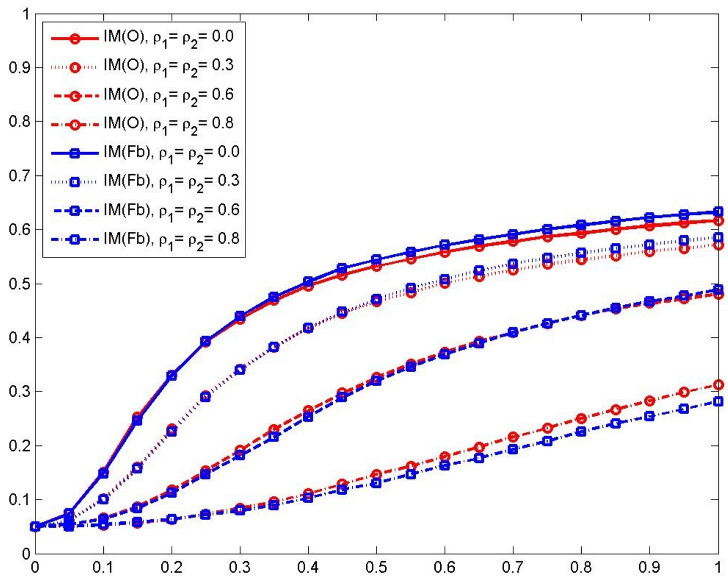
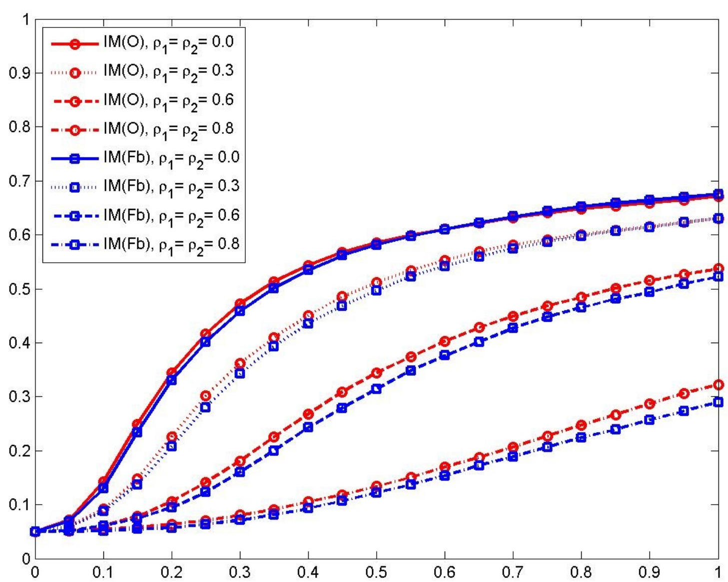


Figure 10: Empirical Size Adjusted Power, IM(fb), $T = 200$, $q = 2$, $\rho_1 = \rho_2 = .3$
$G(X_t) = x_{1t}^2$, QS Kernel, Andrews Bandwidth, 5% Nominal Level



ϕ

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



ϕ

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