TEXTO PARA DISCUSSÃO

No. 577

Nonlinear Cointegration, Misspecification and Bimodality

Marcelo C Medeiros
Eduardo Mendes
Les Oxley

DEPARTAMENTO DE ECONOMIA
www.econ.puc-rio.br

A NOTE ON NONLINEAR COINTEGRATION, MISSPECIFICATION AND BIMODALITY

MARCELO C. MEDEIROS, EDUARDO MENDES, AND LES OXLEY

ABSTRACT. We derive the asymptotic distribution of the ordinary least squares estimator in a regression with cointegrated variables under misspecification and/or nonlinearity in the regressors. We show that, under some circumstances, the order of convergence of the estimator changes and the asymptotic distribution is non-standard. The t-statistic might also diverge. A simple case arises when the intercept is erroneously omitted from the estimated model or in nonlinear-in-variables models with endogenous regressors. In the latter case, a solution is to use an instrumental variable estimator. The core results in this paper also generalise to more complicated nonlinear models involving integrated time series.

KEYWORDS: Cointegration, nonlinearity, bimodality, misspecification, asymptotic theory.

ACKNOWLEDGMENTS: This work was carried out while the first author was visiting the Department of Economics at the University of Canterbury and its kind hospitality is gratefully acknowledged. The authors are very thankful to Peter C.B. Phillips for many enlightening discussions. The suggestions of the editor, Essie Maasoumi, and an Associate Editor are also gratefully acknowledged. The research of the first author is partially supported by CNPq/Brazil and the third by the Marsden Fund of New Zealand.

1. INTRODUCTION

In this paper we discuss a few interesting issues that can emerge in nonlinear and misspecified cointegration models. More precisely, we show that under certain conditions, the classical Ordinary Least Squares (OLS) estimator has a non-standard, possibly bimodal, distribution. Furthermore, the OLS estimator does not converge at usual rate T and this may have a harmful effect on hypothesis testing. To overcome this problem we propose a Two Stage Least Square (2SLS) estimator which is both consistent and has the classical distribution found in the cointegration literature. Although the

Date: October 23, 2012.

results in the paper are simple and trivially obtained from standard cointegration asymptotics[1], the implications for statistical inference are significant and should be investigated.

Slower rates of convergence and atypical distributions, possibly bimodal, are not rare in the cointegration literature. Hansen (1992) showed that the OLS estimator in heteroskedastic cointegration models converges at $\sqrt{T}$ rate, while Park and Phillips (1999, 2001) found slower rates of convergence than the usual T when nonlinear regressions with integrated variables are considered. Hansen (1992) and Chang and Park (2011) also found that under endogeneity, the distribution of the parameter of interest has an unusual form and might even be bimodal. We show that these phenomena can also happen in other classes of cointegration models.

All the results in this paper are derived for a simple, but interesting, nonlinear-in-variables framework. However, the results found for this trivial model can be straightforwardly generalised to more complex ones, such as smooth transition cointegration as in Choi and Saikkonnen (2004a,b) when the transition variable is stationary (Medeiros, Mendes, and Oxley, 2011), and functional coefficient cointegration models as in Xiao (2009), where the coefficients are allowed to vary according to some stationary variable.

Nonstandard distributions can also arise in even simpler situations. We show that when an intercept is erroneously omitted from a linear regression with cointegrated variables, the distribution of the OLS estimator of the slope parameter is not mixed normal anymore. This omission changes the converge rate of the estimator as well as the distribution of the t-statistic. This result also extends to other linear estimators such as the dynamic OLS (DOLS). Although is quite unusual to have the omission of the constant term in a standard regression, in cointegration analysis there are several applications where the intercept is naturally omitted. When testing for the purchasing power parity (PPP) or unbiased forward rate hypotheses (where the intercept is zero by definition), consumption function theory, or synchronous dynamics among commodity prices, for example, it is not rare to find empirical models omitting the intercept. Furthermore, due to the super-consistency property of the OLS estimator when the variables are cointegrated, is reasonable to imagine that the omission of the intercept will not cause

[1]For references on classical results in cointegration theory see e.g. Park and Phillips (1988) and Chapters 17, 18 and 19 in Hamilton (1994).

any harm. We show that this statement is wrong. The estimator will still be consistent, however, this simple misspecification will have serious consequences in terms of inference.

The rest of the paper is organised as follows. Section 2 present the classical OLS results for cointegrating regression, with focus on the issue we deal in the next sections. Section 3 presents the simplest case arising from the erroneous omission of the intercept in the cointegrating relationship. Section 4 presents the general result which permits nonlinearity and potential endogeneity. We also discuss the IV estimator of the model. Section 5 presents the simulation results and Section 6 concludes.

2. Overview

Consider a random vector $\boldsymbol{Y}_t = (y_{1t}, \ldots, y_{kt})' \in \mathbb{R}^k$, which consists of possibly non-stationary and cointegrated variables and write a expression for the OLS estimator $\widehat{\boldsymbol{\beta}}_{OLS} \in \mathbb{R}^q$:

$$\widehat{\boldsymbol{\beta}}_{OLS} = [\boldsymbol{A}(\boldsymbol{Y})]^{-1}\,\boldsymbol{H}(\boldsymbol{Y}), \tag{1}$$

where $\boldsymbol{Y} = (\boldsymbol{Y}'_1, \ldots, \boldsymbol{Y}'_T)'$, T is the sample size, and $\boldsymbol{A}(\boldsymbol{Y})$ and $\boldsymbol{H}(\boldsymbol{Y})$ are $q \times q$ and $q \times 1$ matrices, respectively. The usual cointegration theory postulates that, under certain regularity conditions, the estimator in (1) converges to a mixed normal random vector at rate T.

In this note, we show that under some circumstances, such as simple misspecification or nonlinearity, this convergence rate does not hold. More specifically, some components of the random vector $\boldsymbol{H}(\boldsymbol{Y})$ will be written as $\sum_{t=1}^{T} y_{it}v_t$, where y_{it} is an integrated variable and v_t is a weakly stationary random term with mean different from zero. Hence, writing a Brownian motion as $B(r)$, $\sum_{t=1}^{T} y_{it}v_t$ will converge to $\int_0^1 B(r)dr$ at rate $T^{3/2}$ instead of $\int_0^1 B(r)dB(r)$ at rate T, which is the case when $\mathbb{E}(v_t) = 0$.[2] These terms may reduce the order of converge of the estimator, change the respective asymptotic distribution, and make the t-statistic divergent.

We consider two specific cases of the above mentioned phenomenon. In the first one, $\mathbb{E}(v_t) \neq 0$ due to the omission of an intercept in a linear cointegration model. In the second case, the model is a correctly specified nonlinear-in-variables equation which consists of integrated and stationary variables. In this situation, the source of the problem is different: $\mathbb{E}(v_t) \neq 0$ whenever the stationary variable is endogenous. This is an interesting case to be analysed as usual corrections for cointegration

[2] See Hamilton (1994, p. 486)

models with endogenous regressors, such as DOLS, will not work, and an IV estimator must be considered instead.

3. A SIMPLE RESULT

In this section we report the consequence of a simple misspecification problem: the omission of an intercept in a regression with cointegrated variables. Although it is quite rare to have the omission of the intercept in a linear regression with stationary variables, this is not the case when the variables are cointegrated. It is not unfrequent to see applications of two-step-estimators, as in the Engle-Granger methodology, where the first step consists in estimating a cointegration relation without an intercept. In the latter, the intercept is usually included in the error-correction model.

Consider the following assumption about the data generating process.

ASSUMPTION 1. *Let $\boldsymbol{x}_t = \boldsymbol{x}_{t-1} + \boldsymbol{v}_t$, where $\boldsymbol{x}_t \in \mathbb{R}^{k_x}$ and $\boldsymbol{v}_t \sim \mathsf{IID}(\mathbf{0}, \boldsymbol{\Omega})$. Furthermore, $y_t = \alpha + \boldsymbol{\beta}'\boldsymbol{x}_t + u_t$, where $\alpha \neq 0$, $u_t \sim \mathsf{IID}(\mathbf{0}, \sigma_u^2)$, and $\mathbb{E}(\boldsymbol{v}_t u_\tau) = \mathbf{0}$, $\forall\, t, \tau$. Assume that the processes $\boldsymbol{S}_{\boldsymbol{v},T}(r) = \frac{1}{T}\sum_{i=1}^{[Tr]} \boldsymbol{v}_i$ and $S_{u,T}(r) = \frac{1}{T}\sum_{i=1}^{[Tr]} u_i$, $r \in [0,1]$, satisfy[3]:*

$$\boldsymbol{X}_{\boldsymbol{v},T}(r) \equiv \sqrt{T}\boldsymbol{S}_{\boldsymbol{v},T}(r) \Rightarrow \boldsymbol{B}_{\boldsymbol{v}}(r), \text{ and } X_{u,T}(r) \equiv \sqrt{T} S_{u,T}(r) \Rightarrow \sigma_u^2 W_u(r), \quad \text{as } T \to \infty, \tag{2}$$

where $\boldsymbol{B}_{\boldsymbol{v}}(r) \in \mathbb{R}^{k_x}$ is a multivariate Brownian motion with covariance matrix $\boldsymbol{\Omega}$ and $W_u(r)$ is a standard Brownian motion. Finally, assume that $W_u(r)$ is independent of $\boldsymbol{B}_{\boldsymbol{v}}(r)$.[4]

Now, suppose that an econometrician estimates the regression described in Assumption 1 by OLS ommiting the intercept. In this case, the following trivial result holds.

PROPOSITION 1. *Define $\widehat{\boldsymbol{\beta}} = \left(\sum_{t=1}^{T} \boldsymbol{x}_t\boldsymbol{x}_t'\right)^{-1} \sum_{t=1}^{T} \boldsymbol{x}_t y_t$, which is the OLS estimator when the intercept is erroneously omitted from the estimated equation. Under Assumption 1*

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \Rightarrow \alpha \left[\int_0^1 \boldsymbol{B}_{\boldsymbol{v}}(r)\boldsymbol{B}_{\boldsymbol{v}}(r)' dr\right]^{-1} \int_0^1 \boldsymbol{B}_{\boldsymbol{v}}(r) dr. \tag{3}$$

Many interesting features emerge from the above result. First, the OLS estimator is consistent but no longer super-consistent as the convergence rate is $\sqrt{T}$ as opposed to T. This change in the

[3] $[X]$ denotes the integer part of X.

[4] This last assumption excludes the case where the multivariate random walk is endogenous with respect to $\boldsymbol{\beta}$. Generalizing our results to the case of endogenous $\boldsymbol{x}_t$ is considered in Section 4.

convergence rate has serious implications in hypothesis testing. Second, the distribution in (3) is not a mixed normal anymore, even with exogenous regressors and IID errors. In some specific cases, the distribution in (3) can be bimodal. Figure 1, panel (a), displays the first marginal component of the asymptotic distribution in (3) for different dimensions, k_x, of the Brownian motion $\boldsymbol{B}_{\boldsymbol{v}}(r)$ [5]. The distribution is clearly bimodal for $k_x = 1$ and $k_x = 2$. However, the bimodality disappears as the dimension of $\boldsymbol{B}(r)$ increases. Third, there is a variance reduction as k_x grows. In order to compare with the standard result in cointegration theory, in panel (b) we consider the case where the intercept is zero in the cointegration relationship, such that the usual result holds, i.e., $T\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \Rightarrow \left[\int_0^1 \boldsymbol{B}_{\boldsymbol{v}}(r)\boldsymbol{B}_{\boldsymbol{v}}(r)'dr\right]^{-1} \int_0^1 \boldsymbol{B}_{\boldsymbol{v}}(r)dW_u(r)$. As we can see, the distribution is, as expected, always unimodal and mixed normal and, contrary to the previous case, the variance increases as the $k_x \longrightarrow \infty$.

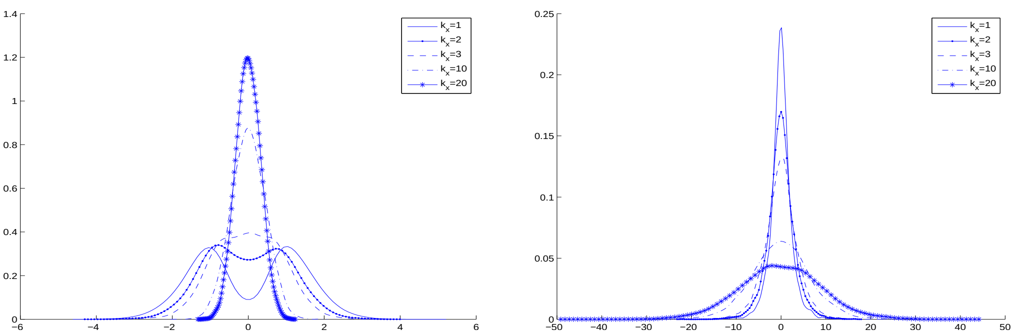

FIGURE 1. Asymptotic distribution of the OLS estimator of in a multiple cointegrating regression for different number of regressors. Panel (a) $\alpha \neq 0$ and it is omitted from the estimated regression. Panel (b) $\alpha = 0$.

The reason why bimodality seems to vanish as the number of regressors increase is simple. Bimodality arises because the scalar distribution

$$\frac{\int_0^1 B(r)dr}{\int_0^1 B^2(r)dr}$$

is bimodal as verified in our simulations. Note that the numerator and the denominator in the above expression consists of the same scalar Brownian motion. On the other hand, in a multivariate setting

[5]In order to simulate the distributions we consider that $\boldsymbol{\Omega}$ in Assumption 1 is an identity matrix and $\alpha = 1$. The Brownian motions are generated from 10,000 observations and the simulations repeated 10,000 times.

the term $\int_0^1 \boldsymbol{B_v}(r)\boldsymbol{B_v}(r)'dr$ involves sums of cross-products of different scalar Brownian motions, such that, component-wise, the asymptotic distribution of the parameters is a ratio where the denominator and the numerator are not functions of the same scalar Brownian motion.

To evaluate the effects of the above result in terms of inference, we consider the simple case of a single regressor, i.e., $k_x = 1$. Under the misspecified model without an intercept, the distribution of the t-statistic for $\mathcal{H}_0 : \beta = \beta^*$ is given in the following proposition.

PROPOSITION 2. *Suppose that Assumption 1 holds with $k_x = 1$, such that $\boldsymbol{B_v} \equiv \sigma_v W_v(r)$, where $W_v(r)$ is a standard Brownian process. Under the null hypothesis $\beta = \beta^*$,*

$$\frac{1}{\sqrt{T}}t_\beta = \frac{1}{\sqrt{T}}\frac{\widehat{\beta}-\beta^*}{\widehat{\sigma}_u\left(\sum_{t=1}^T x_t^2\right)^{-1/2}} \Rightarrow \frac{\alpha}{\sigma_u}\frac{\int_0^1 W_v(r)dr}{\left[\int_0^1 W_v(r)^2dr\right]^{-1/2}}. \tag{4}$$

As the denominator of the t-statistic is $O(T)$ and the numerator is $O(\sqrt{T})$, the ratio will diverge as $T \longrightarrow \infty$, such that it should be scaled by $\sqrt{T}$. Furthermore, the distribution of the scaled t-statistic is not free from nuisance parameters as both α and σ_u appear in the asymptotic distribution.

The consequences of the omission of an intercept in a cointegrating regression are quite different from the ones in a model involving solely stationary variables. In the latter case, the OLS estimator is inconsistent but the asymptotic distribution is still Gaussian and the order of convergence is preserved. More specifically, assuming that $\boldsymbol{x}_t$ is a vector of second-order stationary random variables and $\underset{T\longrightarrow\infty}{\mathsf{plim}} \frac{1}{T}\sum_{t=1}^T \boldsymbol{x}_t\boldsymbol{x}_t' = \boldsymbol{Q}$, then

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right) \xrightarrow{d} \mathsf{N}\left[\alpha\boldsymbol{Q}^{-1}\boldsymbol{\mu}, \boldsymbol{\Omega}\right],$$

where $\boldsymbol{\Omega}$ is a positive-definite covariance matrix[6].

4. A SIMPLE NONLINEAR-IN-VARIABLES MODEL

Nonstandard results also arise in simple nonlinear models with nonstationary and stationary variables. In this section we consider a cointegration regression with time-varying parameters. Our model has a key feature that the cointegration relationship changes according to an observed state vector of

[6]A second-order stationary random variable x_t has the following features: (a) $\mathbb{E}(x_t) = \mu$, $-\infty < \mu < \infty$; (b) $\mathbb{E}[(x_t - \mu)^2] = \sigma^2 < \infty$; and $\mathbb{E}[(x_t-\mu)(x_{t-j}-\mu)] = \gamma_j$, $-\infty < \gamma_j < \infty$.

variables $\boldsymbol{z}_t$. We assume that $\boldsymbol{z}_t$ is observable and second-order stationary. If $\boldsymbol{z}_t$ is endogenous the asymptotic distribution of the OLS estimator will be changed.

Consider the following assumption.

ASSUMPTION 2. *The vector $\boldsymbol{Y}_t = (y_t, x_t, \boldsymbol{z}_t')'$ satisfy*

$$y_t = \alpha_0 + \beta_0 x_t + \alpha_1 g(\boldsymbol{z}_t) + \beta_1 x_t g(\boldsymbol{z}_t) + u_t, \tag{5}$$

$x_t = x_{t-1} + v_t$, $u_t = \sum_{j=0}^{\infty} \pi_{u,j}\varepsilon_{1,t-j} = \pi_u(L)\varepsilon_{1,t}$, $v_t = \sum_{j=0}^{\infty} \pi_{v,j}\varepsilon_{2,t-j} = \pi_v(L)\varepsilon_{2,t}$, $\boldsymbol{z}_t = \sum_{j=0}^{\infty} \boldsymbol{\pi}_{\boldsymbol{z},j}\boldsymbol{\varepsilon}_{3,t-j} = \boldsymbol{\pi}_{\boldsymbol{z}}(L)\boldsymbol{\varepsilon}_{3,t}$, $\pi_u(L)$, $\pi_v(L)$, *and* $\boldsymbol{\pi}_{\boldsymbol{z}}(L)$ *are lag polynomials,* $\sum_{j=0}^{\infty} j|\pi_{u,j}| < \infty$, $\sum_{j=0}^{\infty} j\|\pi_{v,j}\| < \infty$, *and* $\sum_{j=0}^{\infty} j\|\boldsymbol{\pi}_{\boldsymbol{z},j}\| < \infty$. *Set* $\boldsymbol{\varepsilon}_t = \left(\varepsilon_{1,t}, \varepsilon_{2,t}, \boldsymbol{\varepsilon}_{3,t}'\right)'$ *such that* $\mathbb{E}(\boldsymbol{\varepsilon}_t) = \mathbf{0}$ *and* $\mathbb{E}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') = \boldsymbol{\Omega}_{\boldsymbol{\varepsilon}}$, *where*

$$\boldsymbol{\Omega}_{\boldsymbol{\varepsilon}} = \begin{pmatrix} \omega_1^2 & \omega_{12} & \boldsymbol{\omega}_{13}' \\ \omega_{12} & \omega_2 & \boldsymbol{\omega}_{23}' \\ \boldsymbol{\omega}_{13} & \boldsymbol{\omega}_{23} & \boldsymbol{\Omega}_3 \end{pmatrix}.$$

Assume also that $x_0 = 0$ *or is randomly drawn from a density independent of* t. *Finally,* $g(\boldsymbol{z}_t) : \mathbb{R}^{k_z} \to \mathbb{R}$ *is a known function of the stationary vector process* $\boldsymbol{z}_t \in \mathbb{R}^{k_z}$.

Model (5) may arise in a number of situations. Threshold cointegration models where the change-point is known are special cases of (5). Such kind of models are relevant when, for instance, the long-run equilibrium changes according to the business cycle. Suppose that $g(\boldsymbol{z}_t) = d_t$ is a dummy variable indicating recessions, such as, for example the NBER recession indicator. In this case, (5) becomes $y_t = \alpha_0 + \beta_0 x_t + \alpha_1 d_t + \beta_1 x_t d_t + u_t$. Of course, in most practical applications $g(\boldsymbol{z}_t)$ will be indexed by a vector of unknown parameters, i.e., $g(\boldsymbol{z}_t) \equiv g(\boldsymbol{z}_t; \boldsymbol{\theta})$. We will not consider this case explicitly here, nevertheless, the consequences of our main results below will also extend to the nonlinear least squares estimator and a correction similar to the one proposed here should be considered.

We show that, under endogeneity of $\boldsymbol{z}_t$, the order of convergence of the parameters related to solely I(1) or I(0) variables will be preserved. However, the parameter related to the mixed variable $x_t g(\boldsymbol{z}_t)$ will converge at a slower rate. This is a new result compared to the recent work of Chang and Park (2011), where only I(1) regressors are considered in a general nonlinear cointegration model.

With respect to the random variable $g(\boldsymbol{z}_t)$ consider the assumption:

ASSUMPTION 3. *The stochastic process* $g(\boldsymbol{z}_t)$ *is such that* $\mathbb{E}\left[g(\boldsymbol{z}_t)\right] = \mu_g < \infty$ *and* $\mathbb{E}\left[g(\boldsymbol{z}_t)^2\right] = m_g^2 < \infty$*. Furthermore,* $\frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{z}_t) \xrightarrow{p} \mu_g$ *and* $\sqrt{T}\left[\frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{z}_t) - \mu_g\right] \xrightarrow{d} \mathsf{N}(0, \omega_g^2)$*, where* ω_g^2 *is the long-run variance of* $g(\boldsymbol{z}_t)$*. Assume also that* $\mathbb{E}\left[g(\boldsymbol{z}_t)u_t\right] = \mu_{gu} < \infty$ *and* $\mu_{gu} \neq 0$*. Finally,* $\frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{z}_t)u_t \xrightarrow{p} \mu_{gu}$ *and* $\sqrt{T}\left[\frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{z}_t)u_t - \mu_{gu}\right] \xrightarrow{d} \mathsf{N}(0, \omega_{gu}^2)$*, where* $\omega_{gu}^2 < \infty$*.*

Define the following stationary zero-mean process

$$\boldsymbol{w}_t' = \left[u_t, v_t, g(\boldsymbol{z}_t) - \mu_g, g(\boldsymbol{z}_t)^2 - m_g^2, g(\boldsymbol{z}_t)u_t - \mu_{gu}\right]' \in \mathbb{R}^k.$$

We make the following assumptions about $\boldsymbol{w}_t$.

ASSUMPTION 4. *Each element of the process* $\{\boldsymbol{w}_t\}_{t=1}^{\infty}$*, satisfies: (a)*$\mathbb{E}|w_{it}|^a < \infty$*,* $i = 1, \ldots, k$*, for* $2 \leq a < \infty$*; and (b)* $\{\omega_{it}\}_{t=1}^{\infty}$*,* $i = 1, \ldots, k$*, is either uniform mixing of size* $-a/(2a-2)$ *or strong mixing of size* $-a/(a-2)$*, for* $a > 2$*.*

ASSUMPTION 5. *The process* $\boldsymbol{w}_t$ *has a continuous spectral density function* $\boldsymbol{f}_{\boldsymbol{ww}}(\lambda)$ *which is bounded away from zero. Define* $\boldsymbol{X}_T(r) = \sqrt{T}\boldsymbol{S}_T(r)$*, where* $\boldsymbol{S}_T(r) = \frac{1}{T}\sum_{i=1}^{[Tr]} \boldsymbol{w}_i$*,* $r \in [0,1]$*. Hence,* $\boldsymbol{X}_T(r) \Rightarrow \boldsymbol{B}(r),\quad$ *as* $T \to \infty$*, where* $\boldsymbol{B}(r) = \left[B_u(r), B_v(r), B_g(r), B_{g^2}(r), B_{gu}(r)\right]'$ *is a multivariate Brownian process with covariance matrix* $\boldsymbol{\Omega} = \lim_{T\to\infty} T^{-1}\mathbb{E}\left[\boldsymbol{S}_T(r)\boldsymbol{S}_T(r)'\right]$ *defined as*

$$\boldsymbol{\Omega} = \begin{pmatrix} \omega_u^2 & \omega_{vu} & \omega_{gu} & \omega_{g^2u} & \omega_{guu} \\ \omega_{vu} & \omega_v^2 & \omega_{gv} & \omega_{g^2v} & \omega_{guv} \\ \omega_{gu} & \omega_{gv} & \omega_g^2 & \omega_{g^2g} & \omega_{ggu} \\ \omega_{g^2u} & \omega_{g^2v} & \omega_{g^2g} & \omega_{g^2}^2 & \omega_{g^2gu} \\ \omega_{guu} & \omega_{guv} & \omega_{ggu} & \omega_{g^2gu} & \omega_{gu}^2 \end{pmatrix} = \boldsymbol{\Sigma} + \boldsymbol{\Lambda} + \boldsymbol{\Lambda}', \tag{6}$$

accordingly to the partitions of $\boldsymbol{w}_t$*, where* $\boldsymbol{\Sigma} = \mathbb{E}(\boldsymbol{w}_1\boldsymbol{w}_1')$ *and* $\boldsymbol{\Lambda} = \lim_{T\to\infty} \frac{1}{T}\sum_{t=2}^{\infty} \mathbb{E}(\boldsymbol{w}_1\boldsymbol{w}_t')$*.*

Set $\boldsymbol{\theta} = (\alpha_0, \beta_1, \alpha_1, \beta_1)'$. The OLS estimator $\widehat{\boldsymbol{\theta}}$ is distributed as:

THEOREM 1. *Under Assumptions 2–5 and the additional assumption that* $\mu_g \neq 0$,

$$\mathbf{\Gamma}\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right) \Rightarrow \begin{bmatrix} 1 & \int_0^1 B_v(r)dr & \mu_g & \mu_g \int_0^1 B_v(r)dr \\ \cdot & \int_0^1 B_v(r)^2 dr & \mu_g \int_0^1 B_v(r)dr & \mu_g \int_0^1 B_v(r)^2 dr \\ \cdot & \cdot & m_g^2 & m_g^2 \int_0^1 B_v(r)dr \\ \cdot & \cdot & \cdot & m_g^2 \int_0^1 B_v(r)^2 dr \end{bmatrix}^{-1} \begin{bmatrix} \mathsf{N}\left[0, \omega_1^2 \pi_u(1)^2\right] \\ \int_0^1 B_v(r)dB_u(r) + \Delta_{vu} \\ \mathsf{N}(\mu_{gu}, \omega_{gu}^2) \\ \mu_{gu} \int_0^1 B_v(r)dr \end{bmatrix},$$

where $\Delta_{vu} = \sigma_{vu} + \lambda_{vu}$ *and* $\mathbf{\Gamma} = \mathsf{diag}\left(T^{1/2}, T, T^{1/2}, T^{1/2}\right)$. *On the other hand, if* $\mu_g = 0$

$$\mathbf{\Gamma}\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right) \Rightarrow \begin{bmatrix} 1 & \int_0^1 B_v(r)dr & 0 & 0 \\ \cdot & \int_0^1 B_v(r)^2 dr & 0 & 0 \\ \cdot & \cdot & m_g^2 & m_g^2 \int_0^1 B_v(r)dr \\ \cdot & \cdot & \cdot & m_g^2 \int_0^1 B_v(r)^2 dr \end{bmatrix}^{-1} \begin{bmatrix} \mathsf{N}\left[0, \omega_1^2 \pi_u(1)^2\right] \\ \int_0^1 B_v(r)dB_u(r) + \Delta_{vu} \\ \mathsf{N}(\mu_{gu}, \omega_{gu}^2) \\ \mu_{gu} \int_0^1 B_v(r)dr \end{bmatrix}.$$

The endogeneity of $\boldsymbol{z}_t$ plays an important role here as, similar to the case described in Section 3, $T^{-3/2}\sum_{t=1}^{T} x_t g(\boldsymbol{z}_t)u_t$ converges to $\mu_{gu}\int_0^1 B_v(r)dr$. On the other hand, if $\boldsymbol{z}_t$ is exogenous, the usual convergence $T^{-1}\sum_{t=1}^{T} x_t g(\boldsymbol{z}_t)u_t \Rightarrow \int_0^1 B_v(r)dB_{gu}(r)$ holds, where $B_{gu}(r)$ is the Brownian motion associated to the process $g(\boldsymbol{z}_t)u_t$. As before, this simple result has severe consequences on parameter inference.

Consider a simple version of (5) where

$$y_t = \beta x_t g(z_t) + u_t. \tag{7}$$

In this case, the OLS estimator for the parameter β will have the following asymptotic distribution

$$\sqrt{T}(\widehat{\beta}-\beta) \Rightarrow \frac{\mu_{gu}}{m_g^2}\left(\int_0^1 B(r)^2 dr\right)^{-1}\int_0^1 B(r)dr, \tag{8}$$

when z_t is endogenous. The same problems discussed in Section 3 will affect parameter inference.

4.1. **A Simple Solution.** In this section we show how IV may be used in the present context. To simplify the exposition, consider the case where $\boldsymbol{x}_t$ is exogenous, such that $\boldsymbol{\omega}_{vu} = \mathbf{0}$.

ASSUMPTION 6. $s_t \in \mathbb{R}$ *is a stochastic process such that* $\mathbb{E}\left[s_t g(\boldsymbol{z}_t)\right] \neq 0$, $\mathbb{E}(s_t u_t) = 0$, $\mathbb{E}(s_t) = \mu_s < \infty$, *and* $\mathbb{E}\left(s_t^2\right) = m_s^2 < \infty$. *Define* $X_{su,T}(r) = \sqrt{T}S_{su,T}(r)$, *where* $S_{su,T}(r) = \frac{1}{T}\sum_{i=1}^{[Tr]} s_i u_i$,

$r \in [0,1]$. *Hence,* $X_{su,T}(r) \Rightarrow \omega_{su} W_{su}(r), \quad$ *as* $T \to \infty$. $W_{su}(r)$ *is a standard Brownian motion and* ω_{su}^2 *is the long-run variance of the process* $s_t u_t$.

Define $\widehat{g}_t = \widehat{\lambda} s_t$, where $\widehat{\lambda} = \left(\sum_{t=1}^{T} s_t^2\right)^{-1} \sum_{t=1}^{T} s_t g(\boldsymbol{z}_t)$. The distribution of the IV estimator $\widetilde{\boldsymbol{\theta}}$ is given by the following theorem.

THEOREM 2. *Under Assumptions 2–6 and the additional assumption that* $\omega_{vu} = 0$, *if* $\mu_s \neq 0$ *then*

$$\boldsymbol{\Gamma}\left(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \Rightarrow \begin{bmatrix} 1 & \int_0^1 B_v(r)dr & \lambda\mu_s & \lambda\mu_s \int_0^1 B_v(r)dr \\ \cdot & \int_0^1 B_v(r)^2 dr & \lambda\mu_s \int_0^1 B_v(r)dr & \lambda\mu_s \int_0^1 B_v(r)^2 dr \\ \cdot & \cdot & \lambda^2 m_s^2 & \lambda^2 m_s^2 \int_0^1 B_v(r)dr \\ \cdot & \cdot & \cdot & \lambda^2 m_s^2 \int_0^1 B_v(r)^2 dr \end{bmatrix}^{-1} \begin{bmatrix} \mathsf{N}\left[0, \omega_1^2 \pi_u(1)^2\right] \\ \int_0^1 B_v(r) dB_u(r) \\ \mathsf{N}(0, \lambda^2 \omega_{su}^2) \\ \lambda \int_0^1 B_v(r) dW_{su}(r) \end{bmatrix}.$$

Otherwise, if $\mu_s = 0$, *then*

$$\boldsymbol{\Gamma}\left(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \Rightarrow \begin{bmatrix} 1 & \int_0^1 B_v(r)dr & 0 & 0 \\ \cdot & \int_0^1 B_v(r)^2 dr & 0 & 0 \\ \cdot & \cdot & \lambda^2 m_s^2 & \lambda^2 m_s^2 \int_0^1 B_v(r)dr \\ \cdot & \cdot & \cdot & \lambda^2 m_s^2 \int_0^1 B_v(r)^2 dr \end{bmatrix}^{-1} \begin{bmatrix} \mathsf{N}\left[0, \omega_1^2 \pi_u(1)^2\right] \\ \int_0^1 B_v(r) dB_u(r) \\ \mathsf{N}(0, \lambda^2 \omega_{su}^2) \\ \lambda \int_0^1 B_v(r) dW_{su}(r) \end{bmatrix}.$$

The matrix $\boldsymbol{\Gamma}$ *is given by* $\boldsymbol{\Gamma} = \mathsf{diag}\left(T^{1/2}, T, T^{1/2}, T\right)$ *in both cases.*

5. MONTE CARLO SIMULATIONS

Consider the following special case of the general model described in Assumption 2:

$$\begin{aligned} y_t &= \alpha_0 + \alpha_1 I(z_t > 0) + \beta_0 x_t + \beta_1 x_t I(z_t > 0) + u_t \\ &= 1 + x_t + \alpha_1 I(z_t > 0) + x_t I(z_t > 0) + u_t, \\ z_t &= s_t + u_t, \quad \text{and} \quad x_t = x_{t-1} + v_t. \end{aligned}$$

$I(A)$ is an indicator function which equals one if the event A occurs or zero otherwise, $u_t \sim \mathsf{NID}(0,1)$, $v_t \sim \mathsf{NID}(0,1)$, $s_t \sim \mathsf{NID}(0,1)$, $\mathbb{E}(u_t v_\tau) = 0$, and $\mathbb{E}(s_t u_\tau) = 0$, $\forall\, t, \tau$.

We consider two cases: $\alpha_1 = 0$ and $\alpha_1 = 1$. We simulate 5000 observations over 1000 replications and evaluate the distribution of both the OLS and IV estimators of α_0, α_1, β_0, and β_1. In the first

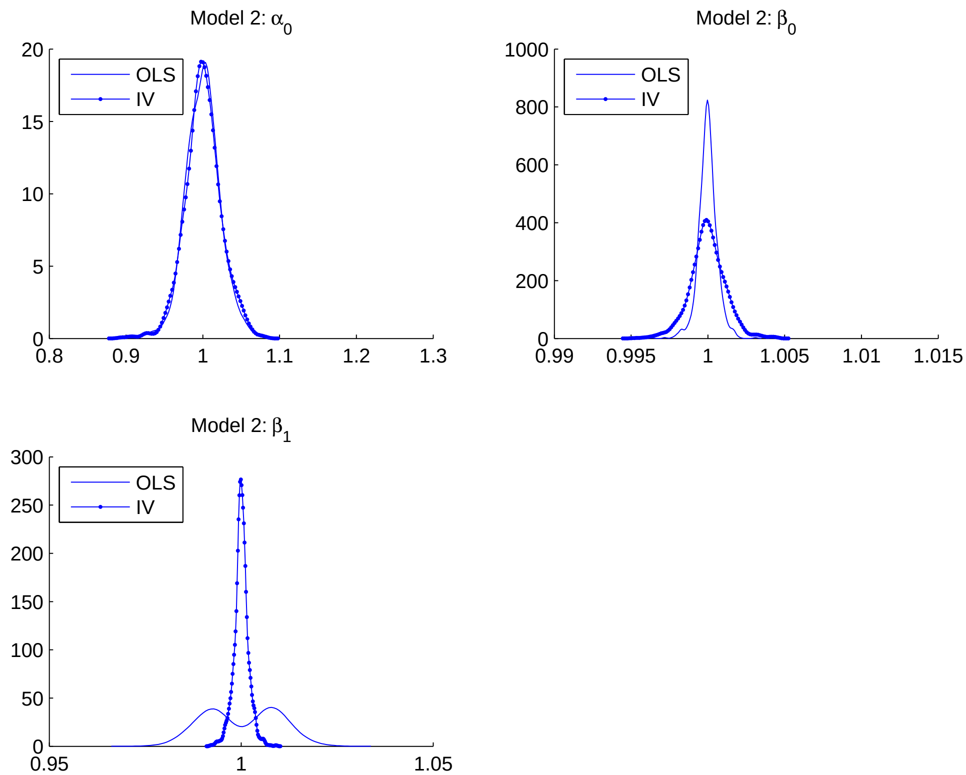

FIGURE 2. Empirical distribution of the OLS and IV estimators with $\alpha_1 = 0$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

stage the variable $I(z_t > 0)$ is regressed on $I(s_t > 0)$. The results are shown in Figures 2–3. We also consider the distribution of the t-statistic under the null hypothesis as shown in Figures 4–5.

Several features emerge from the graphs. First, depending on the value of α_1, bimodality may or may not be present. When $\alpha_1 = 1$, the OLS estimator of β_1 is always bimodal, while the IV estimator is not. Furthermore, in this specific case, the IV estimator has lower variance than the OLS estimator. The t-statistics for the OLS estimators display bimodality, whereas the the ones for the IV estimators are, as expected, normally distributed. Second, the OLS estimator of β_1 is always consistent. The OLS delivers inconsistent estimators for both α_0 and α_1 while the IV estimator is always consistent. The t-statistic for the IV estimators are always distributed as a standard normal random variable.

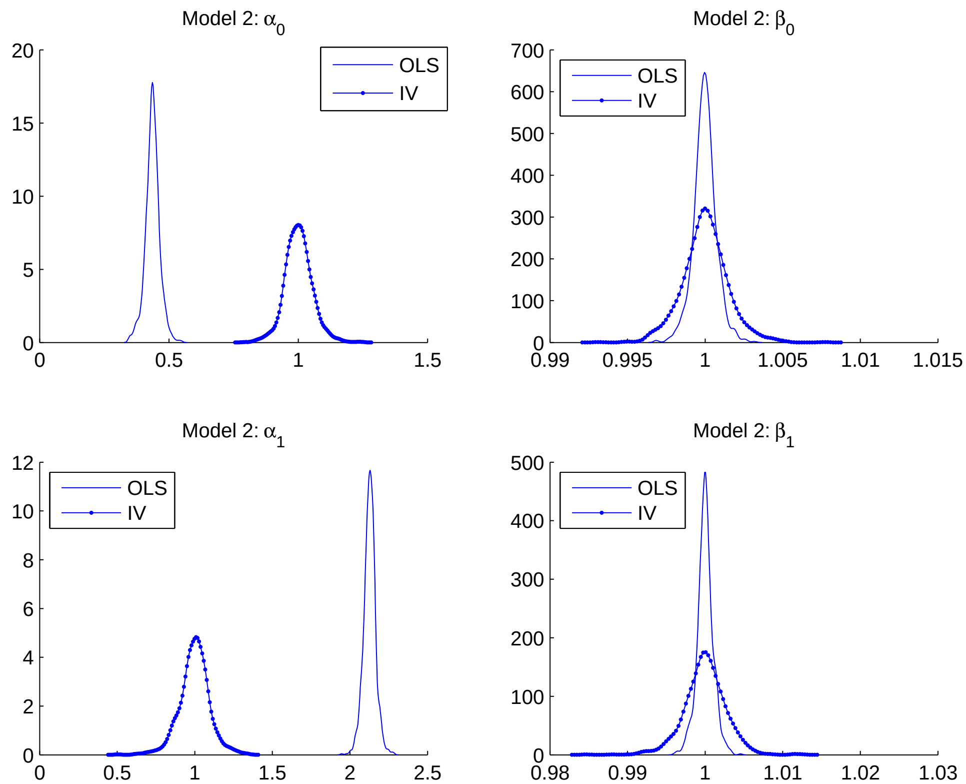

FIGURE 3. Empirical distribution of the OLS and IV estimators with $\alpha_1 = 1$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

6. CONCLUSION

The paper identifies a number of interesting cases that can arise in cointegration models. Bimodality is one such case. We show how bimodality arises; the consequences, including the loss of super-consistency of the estimates in a simple case; and how the addition of regressors leads to disappearance of the phenomena. Inclusion of an intercept removes both bimodality and inference related problems arising from using a non-scaled t-statistic. Secondly, in the more general nonlinear case, where, as expected, endogeneity leads to the possibility of inconsistent OLS estimates, but also the potential for the asymptotic distribution to be bimodal. The use of IV in these cases removes both bimodality and inconsistency.

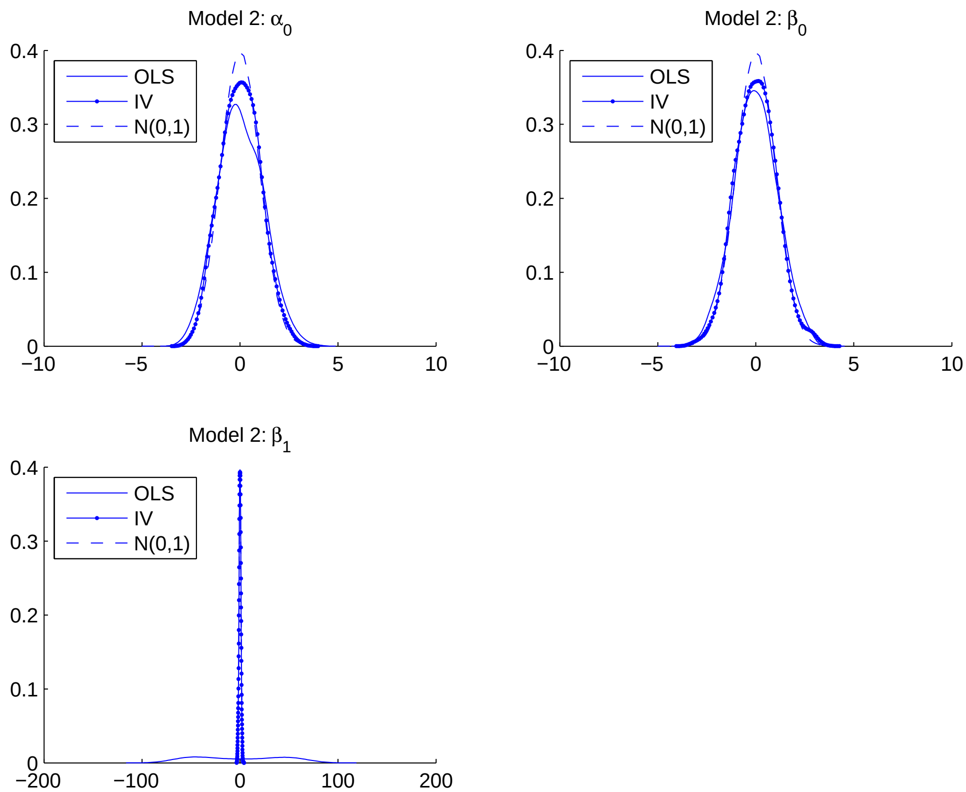

FIGURE 4. Empirical distribution of t-statistic for the OLS and IV estimators with $\alpha_1 = 0$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

REFERENCES

CHANG, Y., AND J. PARK (2011): "Endogeneity in Nonlinear Regressions with Integrated Time Series," *Econometric Reviews*, 30, 51–87.

HAMILTON, J. (1994): *Time series analysis*. Princeton Univ Press.

HANSEN, B. E. (1992): "Heteroskedastic Cointegration," *Journal of Econometrics*, 54, 139–158.

IBRAGIMOV, R., AND P. PHILLIPS (2008): "Regression asymptotics using martingale convergence methods," *Econometric Theory*, 24, 888–947.

MEDEIROS, M., E. MENDES, AND L. OXLEY (2011): "Cointegrating Smooth Transition Regressions with a Stationary Transition Variable," Working paper, Pontifical Catholic University of Rio de Janeiro.

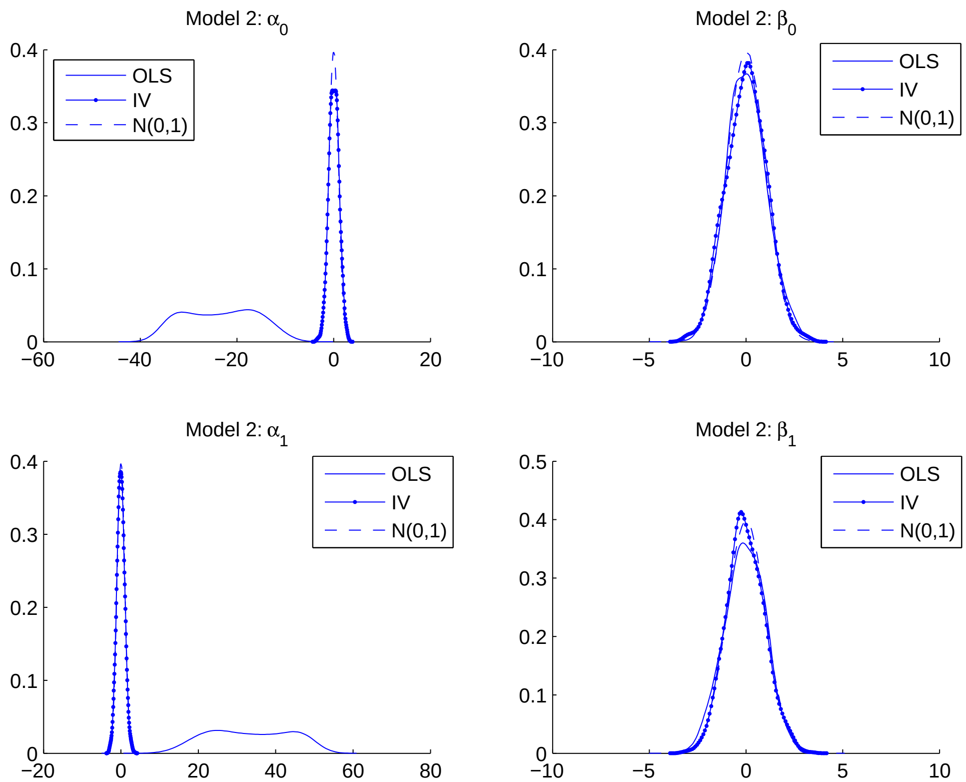

FIGURE 5. Empirical distribution of t-statistic for the OLS and IV estimators with $\alpha_1 = 1$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

PARK, J., AND P. PHILLIPS (1988): "Statistical inference in regressions with integrated processes: Part I," *Econometric Theory*, 4, 468–497.

——— (1999): "Asymptotics fof Nonlinear Transformation of Integrated Time Series," *Econometric Theory*, 15, 269–298.

——— (2001): "Nonlinear Regression with Integrated Time Series," *Econometrica*, 69, 1452–1498.

SAIKKONEN, P., AND I. CHOI (2004a): "Cointegrating Smooth Transition Regressions," *Econometric Theory*, 20, 301–340.

——— (2004b): "Testing Linearity in Cointegrating Smooth Transition Regressions," *Econometrics Journal*, 7, 341–365.

XIAO, Z. (2009): "Functional-Coefficient Cointegration Models," *Journal of Econometrics*, 152, 81–92.

APPENDIX A. LEMMA

LEMMA 1. *Let* $\{x_t\}_{t=1}^T$ *be a stochastic process satisfying* $x_t = x_{t-1} + v_t$, *where* $\mathbb{E}(v_t) = 0$. *Define* $\boldsymbol{w}_t = (u_t - \mu_u, v_t)'$, *where* u_t *is a stationary process with* $\mathbb{E}(u_t) = \mu_u < \infty$. *Assume that the process* $\boldsymbol{S}_T(r) = \frac{1}{T}\sum_{j=1}^{[Tr]} \boldsymbol{w}_j$, $r \in [0,1]$, *satisfies the multivariate invariance principle. More specifically, define* $\boldsymbol{X}_T(r) = \sqrt{T}\boldsymbol{S}_T(r)$, *such that* $\boldsymbol{X}_T(r) \Rightarrow \boldsymbol{B}(r)$, *as* $T \to \infty$, *where* $\boldsymbol{B}(r) = [B_u(r), B_v(r)]' \in \mathbb{R}^2$ *is a vector Brownian process with covariance matrix*

$$\boldsymbol{\Omega} = \begin{pmatrix} \omega_u^2 & \omega_{vu} \\ \omega_{vu} & \omega_v^2 \end{pmatrix} = \mathbb{E}(\boldsymbol{w}_1\boldsymbol{w}_1') + \sum_{k=2}^{\infty}\left[\mathbb{E}(\boldsymbol{w}_1\boldsymbol{w}_k') + \mathbb{E}(\boldsymbol{w}_k\boldsymbol{w}_1')\right] = \boldsymbol{\Sigma} + \boldsymbol{\Lambda} + \boldsymbol{\Lambda}'. \tag{9}$$

Define $\Delta_{uv} = \sigma_{uv} + \lambda_{uv}$. *Under the assumptions above, the following results hold:*

(a) *if* $\mu_u \neq 0$, *then* $T^{-2}\sum_{t=1}^T x_t^2 u_t \Rightarrow \mu_u \int_0^1 B_v^2 dr$;
(b) *if* $\Delta_{vu} \neq 0$ *and* $\mu_u = 0$, *then* $T^{-3/2}\sum_{t=1}^T x_t^2 u_t \Rightarrow \int_0^1 B_v(r)^2 dB_u(r) + \Delta_{vu}\int_0^1 B_v(r)dr$;
(c) *if* $\Delta_{vu} = 0$ *and* $\mu_u = 0$, *then* $T^{-3/2}\sum_{t=1}^T x_t^2 u_t \Rightarrow \int_0^1 B_v(r)^2 dB_u(r)$;
(d) *and, if* $\mu_u \neq 0$, *then* $T^{-3/2}\sum_{t=1}^T x_t u_t \Rightarrow \mu_u \int_0^1 B_v(r)dr$.

Proof. First, define $u_t^* = u_t - \mu_u$ and write $\sum_{t=1}^T x_t^2 u_t = \mu_u \sum_{t=1}^T x_t^2 + \sum_{t=1}^T x_t^2 u_t^*$. It is well-known that $\mu_u \frac{1}{T^2}\sum_{t=1}^T x_t^2 \Rightarrow \mu_u \int_0^1 B_v(r)^2 dr$. Direct application of the results in Theorem 3.1 in Ibragimov and Phillips (2008) implies that

$$\frac{1}{T^{3/2}}\sum_{t=1}^T x_t^2 u_t^* \Rightarrow \int_0^1 B_v(r)^2 dB_u(r) + \Delta_{vu}\int_0^1 B_v(r)dr.$$

Hence, (a), (b), and (c) follow from the above convergence limits.

To prove (d) is enough to write $\sum_{t=1}^T x_t u_t = \sum_{t=1}^T x_t(\mu_u + u_t^*) = \mu_u \sum_{t=1}^T x_t + \sum_{t=1}^T x_t u_t^*$, and note that $\mu_u \frac{1}{T^{3/2}}\sum_{t=1}^T x_t \Rightarrow \int_0^1 B_v(r)dr$ and $\frac{1}{T}\sum_{t=1}^T x_t u_t^* \Rightarrow \int_0^1 B_v(r)dB_u(r)dr + \Delta_{uv}$.

■

APPENDIX B. PROOF OF PROPOSITIONS AND THEOREMS

B.1. **Proof of Proposition 1.** The proof is very simple. First, note that

$$\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) = \left(\sum_{t=1}^T \boldsymbol{x}_t\boldsymbol{x}_t'\right)^{-1}\sum_{t=1}^T \boldsymbol{x}_t(\alpha + u_t) = \left(\sum_{t=1}^T \boldsymbol{x}_t\boldsymbol{x}_t'\right)^{-1}\left(\alpha\sum_{t=1}^T \boldsymbol{x}_t + \sum_{t=1}^T \boldsymbol{x}_t u_t\right).$$

It is clear that $\frac{1}{T^2}\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t' \Rightarrow \int_0^1 \boldsymbol{B}(r)\boldsymbol{B}(r)'dr$, $\frac{1}{T^{3/2}}\sum_{t=1}^{T}\boldsymbol{x}_t \Rightarrow \int_0^1 \boldsymbol{B}(r)dr$, and $\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_t u_t \Rightarrow \int_0^1 \boldsymbol{B}(r)dW(r)$. Hence, as $T^{-3/2}\sum_{t=1}^{T}\boldsymbol{x}_t u_t \xrightarrow{p} \mathbf{0}$,

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right) = \left(\frac{1}{T^2}\sum_{t=1}^{T}\boldsymbol{x}_t\boldsymbol{x}_t'\right)^{-1}\left(\alpha\frac{1}{T^{3/2}}\sum_{t=1}^{T}\boldsymbol{x}_t + \frac{1}{T^{3/2}}\sum_{t=1}^{T}\boldsymbol{x}_t u_t\right)$$
$$\Rightarrow \alpha\left(\int_0^1 \boldsymbol{B}(r)\boldsymbol{B}(r)'dr\right)^{-1}\int_0^1 \boldsymbol{B}(r)dr.$$

■

B.2. **Proof of Proposition 2.** Write the t-statistic as

$$t_\beta = \frac{\sum_{t=1}^{T} x_t(\alpha+u_t)}{\sum_{t=1}^{T} x_t^2} \div \widehat{\sigma}_u\left(\sum_{t=1}^{T} x_t^2\right)^{-1/2} = \frac{\alpha\sum_{t=1}^{T} x_t}{\widehat{\sigma}_u\left(\sum_{t=1}^{T} x_t^2\right)^{1/2}} + \frac{\sum_{t=1}^{T} x_t u_t}{\widehat{\sigma}_u\left(\sum_{t=1}^{T} x_t^2\right)^{1/2}}.$$

Hence,

$$\frac{1}{\sqrt{T}}t_\beta = \frac{\alpha\frac{1}{T^{3/2}}\sum_{t=1}^{T} x_t}{\widehat{\sigma}_u\left(\frac{1}{T^2}\sum_{t=1}^{T} x_t^2\right)^{1/2}} + \frac{\frac{1}{T^{3/2}}\sum_{t=1}^{T} x_t u_t}{\widehat{\sigma}_u\left(\frac{1}{T^2}\sum_{t=1}^{T} x_t^2\right)^{1/2}} \Rightarrow \frac{\alpha}{\sigma_u}\frac{\int_0^1 W(r)dr}{\left[\int_0^1 W(r)^2 dr\right]^{1/2}}.$$

■

B.3. **Proof of Theorem 1.** First, define the following matrices: $\boldsymbol{H} = \mathsf{diag}\left(\sqrt{T}, T, \sqrt{T}, T\right)$ and $\boldsymbol{D} = \mathsf{diag}\left(1,1,1,\sqrt{T}\right)$. Note that

$$\boldsymbol{D}^{-1}\boldsymbol{H}\left(\widehat{\boldsymbol{\theta}}-\boldsymbol{\theta}\right) = \begin{bmatrix} 1 & \frac{1}{T^{3/2}}\sum_{t=1}^{T} x_t & \frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{z}_t) & \frac{1}{T^{3/2}}\sum_{t=1}^{T} g(\boldsymbol{z}_t)x_t \\ \cdot & \frac{1}{T^2}\sum_{t=1}^{T} x_t^2 & \frac{1}{T^{3/2}}\sum_{t=1}^{T} g(\boldsymbol{z}_t)x_t & \frac{1}{T^2}\sum_{t=1}^{T} g(\boldsymbol{z}_t)x_t^2 \\ \cdot & \cdot & \frac{1}{T}\sum_{t=1}^{T} g(\boldsymbol{z}_t)^2 & \frac{1}{T^{3/2}}\sum_{t=1}^{T} g(\boldsymbol{z}_t)^2 x_t \\ \cdot & \cdot & \cdot & \frac{1}{T^2}\sum_{t=1}^{T} g(\boldsymbol{z}_t)^2 x_t^2 \end{bmatrix}^{-1}$$
$$\times \begin{bmatrix} \frac{1}{\sqrt{T}}\sum_{t=1}^{T} u_t \\ \frac{1}{T}\sum_{t=1}^{T} x_t u_t \\ \frac{1}{\sqrt{T}}\sum_{t=1}^{T} g(\boldsymbol{z}_t)u_t \\ \frac{1}{T^{3/2}}\sum_{t=1}^{T} x_t g(\boldsymbol{z}_t)u_t \end{bmatrix}.$$

Therefore, for $\mu_g \neq 0$ we need to show the following: (a) $\frac{1}{T^{3/2}}\sum_{t=1}^T x_t \Rightarrow \int_0^1 B_v(r)dr$; (b) $\frac{1}{T}\sum_{t=1}^T g(\boldsymbol{z}_t) \xrightarrow{p} \mu_g$; (c) $\frac{1}{T^{3/2}}\sum_{t=1}^T g(\boldsymbol{z}_t)x_t \Rightarrow \mu_g\int_0^1 B_v(r)dr$; (d) $\frac{1}{T^2}\sum_{t=1}^T x_t^2 \Rightarrow \int_0^1 B_v(r)^2dr$; (e) $\frac{1}{T^2}\sum_{t=1}^T g(\boldsymbol{z}_t)x_t^2 \Rightarrow \mu_g\int_0^1 B_v(r)^2dr$; (f) $\frac{1}{T}\sum_{t=1}^T g(\boldsymbol{z}_t)^2 \xrightarrow{p} m_g^2$; (g) $\frac{1}{T^{3/2}}\sum_{t=1}^T g(\boldsymbol{z}_t)^2x_t \Rightarrow m_g^2\int_0^1 B_v(r)dr$; (h) $\frac{1}{T^2}\sum_{t=1}^T g(\boldsymbol{z}_t)^2x_t^2 \Rightarrow m_g^2\int_0^1 B_v(r)^2dr$; (i) $\frac{1}{\sqrt{T}}\sum_{t=1}^T u_t \xrightarrow{d} \mathsf{N}\left[0,\omega_1^2\pi_u(1)^2\right]$; (j) $\frac{1}{T}\sum_{t=1}^T x_tu_t \Rightarrow \int_0^1 B_v(r)dB_u(r)+\Delta_{vu}$; (k) $\frac{1}{\sqrt{T}}\sum_{t=1}^T g(\boldsymbol{z}_t)u_t \xrightarrow{d} \mathsf{N}(\mu_{gu},\omega_{gu}^2)$; and finally (l)$\frac{1}{T^{3/2}}\sum_{t=1}^T x_tg(\boldsymbol{z}_t)u_t \Rightarrow \mu_{gu}\int_0^1 B_v(r)dr$.

In the case $\mu_g = 0$, (c) and (e) should be replaced by the following: (c') $\frac{1}{T^{3/2}}\sum_{t=1}^T g_tx_t \xrightarrow{p} 0$ and (e')$\frac{1}{T^2}\sum_{t=1}^T g(\boldsymbol{z}_t)x_t^2 \xrightarrow{p} 0$.

First, define $g_t \equiv g(\boldsymbol{z}_t)$, $g_t^* = g(\boldsymbol{z}_t)-\mu_g$, $g_t^2 \equiv g(\boldsymbol{z}_t)^2$, and $g_t^{2*} = g(\boldsymbol{z}_t)^2 - m_g^2$. It is clear that (a), (d), and (j) follow from standard results in the literature and (b), (f), (i), and (k) are trivially satisfied. Next, write $\sum_{t=1}^T g_tx_t = \sum_{t=1}^T (m_g+g_t^*)\,x_t = \mu_g\sum_{t=1}^T x_t + \sum_{t=1}^T g_t^*x_t$.

It is clear that

$$\mu_g\frac{1}{T^{3/2}}\sum_{t=1}^T x_t \Rightarrow \mu_g\int_0^1 B_v(r)dr \text{ and } \frac{1}{T}\sum_{t=1}^T g_t^*x_t \Rightarrow \int_0^1 B_v(r)dB_g(r)+\Delta_{gv}.$$

Hence, if $\mu_g \neq 0$, $\frac{1}{T^{3/2}}\sum_{t=1}^T g_tx_t \Rightarrow \mu_g\int_0^1 B_v(r)dr$ and (c) is proved. Otherwise, if $\mu_g = 0$, $\frac{1}{T}\sum_{t=1}^T g_tx_t \Rightarrow \int_0^1 B_v(r)dB_g(r)+\Delta_{gv}$, such that $\frac{1}{T^{3/2}}\sum_{t=1}^T g_tx_t \xrightarrow{p} 0$ and (c') is proved.

Following the same reasoning, write

$$\sum_{t=1}^T g_tx_t^2 = \sum_{t=1}^T (\mu_g+g_t^*)\,x_t^2 = \mu_g\sum_{t=1}^T x_t^2 + \sum_{t=1}^T g_t^*x_t^2.$$

From the results in Lemma 1, it follows that

$$\mu_g\frac{1}{T^2}\sum_{t=1}^T x_t^2 \Rightarrow \mu_g\int_0^1 B_v(r)^2dr \text{ and}$$
$$\frac{1}{T^{3/2}}\sum_{t=1}^T g_t^*x_t^2 \Rightarrow \int_0^1 B_v(r)^2dB_g(r)+\Delta_{gv}\int_0^1 B_v(r)dr.$$

Therefore, if $\mu_g \neq 0$, $\frac{1}{T^2}\sum_{t=1}^T g_tx_t^2 \Rightarrow \mu_g\int_0^1 B_v(r)^2dr$ and (e) is proved. If $\mu_g = 0$, $\frac{1}{T^{3/2}}\sum_{t=1}^T g_tx_t^2 \Rightarrow \int_0^1 B_v(r)^2dB_g(r)+\Delta_{gv}\int_0^1 B_v(r)dr$, such that $\frac{1}{T^2}\sum_{t=1}^T g_tx_t^2 \xrightarrow{p} 0$ and (e') follows.

Now, let's turn to $\sum_{t=1}^T g_t^2x_t$. Again, $\sum_{t=1}^T g_t^2x_t = \sum_{t=1}^T \left(m_g^2+g_t^{2*}\right)x_t = m_g^2\sum_0^T x_t + O(T)$.

Hence, $\frac{1}{T^{3/2}}\sum_{t=1}^T g_t^2x_t \Rightarrow m_g^2\int_0^1 B_v(r)dr$ and (g) follows.

Following similar arguments, it is straightforward to prove (h). To prove (l), define $\eta_t = g_t u_t - \mu_{gu}$

$$\sum_{t=1}^{T} x_t g_t u_t = \sum_{t=1}^{T} (\mu_{gu} + \eta_t) \mu_{gu} \sum_{t=1}^{T} x_t + \sum_{t=1}^{T} x_t \eta_t.$$

From Lemma 1, it follows that

$$\mu_{gu} \frac{1}{T^{3/2}} \sum_{t=1}^{T} x_t \Rightarrow \mu_{gu} \int_0^1 B_v(r)dr \text{ and } \frac{1}{T} \sum_{t=1}^{T} x_t \eta_t \Rightarrow \int_0^1 B_v(r)dB_{gu}(r) + \Delta_{guv}.$$

Therefore, if $\mu_{gu} \neq 0$,

$$\frac{1}{T^{3/2}} \sum_{t=1}^{T} x_t g_t u_t \Rightarrow \mu_{gu} \int_0^1 B_v(r)dr$$

else

$$\frac{1}{T} \sum_{t=1}^{T} x_t g_t u_t \Rightarrow \int_0^1 B_v(r)dB_{gu}(r) + \Delta_{guv}.$$

■

B.4. **Proof of Theorem 2.** Define $\mathbf{\Gamma} = \text{diag}\left(\sqrt{T}, T, \sqrt{T}, T\right)$ and write

$$\mathbf{\Gamma}\left(\widetilde{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) = \left\{ \mathbf{\Gamma}^{-1} \begin{bmatrix} \sum_{t=1}^{T} 1 & \sum_{t=1}^{T} x_t & \sum_{t=1}^{T} \widehat{g}_t & \sum_{t=1}^{T} \widehat{g}_t x_t \\ \cdot & \sum_{t=1}^{T} x_t^2 & \sum_{t=1}^{T} \widehat{g}_t x_t & \sum_{t=1}^{T} \widehat{g}_t x_t^2 \\ \cdot & \cdot & \sum_{t=1}^{T} \widehat{g}_t^2 & \sum_{t=1}^{T} \widehat{g}_t^2 x_t \\ \cdot & \cdot & \cdot & \sum_{t=1}^{T} \widehat{g}_t^2 x_t^2 \end{bmatrix} \mathbf{\Gamma}^{-1} \right\}^{-1} \mathbf{\Gamma}^{-1} \begin{bmatrix} \sum_{t=1}^{T} u_t \\ \sum_{t=1}^{T} x_t u_t \\ \sum_{t=1}^{T} \widehat{g}_t u_t \\ \sum_{t=1}^{T} x_t \widehat{g}_t u_t \end{bmatrix}.$$

Therefore, for $\mu_s = 0$ we need to show: (a) $\frac{1}{T^{3/2}} \sum_{t=1}^{T} x_t \Rightarrow \int_0^1 B_v(r)dr$; (b) $\frac{1}{T} \sum_{t=1}^{T} \widehat{g}_t \xrightarrow{p} \lambda\mu_s$; (c) $\frac{1}{T^{3/2}} \sum_{t=1}^{T} \widehat{g}_t x_t \Rightarrow \lambda\mu_s \int_0^1 B_v(r)dr$; (d) $\frac{1}{T^2} \sum_{t=1}^{T} x_t^2 \Rightarrow \int_0^1 B_v(r)^2 dr$; (e) $\frac{1}{T^2} \sum_{t=1}^{T} \widehat{g}_t x_t^2 \Rightarrow \lambda\mu_s \int_0^1 B_v(r)^2 dr$; (f) $\frac{1}{T} \sum_{t=1}^{T} \widehat{g}_t^2 \xrightarrow{p} \lambda^2 \omega_{su}^2$; (g) $\frac{1}{T^{3/2}} \sum_{t=1}^{T} \widehat{g}_t^2 x_t \Rightarrow \lambda^2 m_s^2 \int_0^1 B_v(r)dr$; (h) $\frac{1}{T^2} \sum_{t=1}^{T} \widehat{g}_t^2 x_t^2 \Rightarrow \lambda^2 m_s^2 \int_0^1 B_v(r)^2 dr$; (i) $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_t \xrightarrow{d} \mathsf{N}\left[0, \omega_1^2 \pi_u(1)^2\right]$; (j) $\frac{1}{T} \sum_{t=1}^{T} x_t u_t \Rightarrow \int_0^1 B_v(r)dB_u(r)$; (k) $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \widehat{g}_t u_t \xrightarrow{d} \mathsf{N}(0, \lambda^2 m_s^2 \sigma_u^2)$; and (l) $\frac{1}{T} \sum_{t=1}^{T} x_t \widehat{g}_t u_t \Rightarrow \lambda \int_0^1 B_v(r)dW_{su}(r)$.

In the case $\mu_s = 0$, (c) and (e) should be replaced by (c') $\frac{1}{T^{3/2}} \sum_{t=1}^{T} \widehat{g}_t x_t \xrightarrow{p} 0$ and (e') $\frac{1}{T^2} \sum_{t=1}^{T} \widehat{g}_t x_t^2 \xrightarrow{p} 0$. Items (a), (d), (i), and (j) follow trivially as in the proof of Theorem 1. Writing $g_t = \widehat{\lambda} s_t$ and noting that $\underset{T\longrightarrow\infty}{plim}\, \widehat{\lambda} = \lambda$, it is trivial to prove items (b), (f), and (k). The proof of the remaining items are similar to the ones in Theorem 1.

■

(M. C. Medeiros) DEPARTMENT OF ECONOMICS, PONTIFICAL CATHOLIC UNIVERSITY OF RIO DE JANEIRO, RIO DE JANEIRO, RJ, BRAZIL.

E-mail address: mcm@econ.puc-rio.br

(E. F. Mendes) DEPARTMENT OF STATISTICS, NORTHWESTERN UNIVERSITY, EVANSTON, IL, U.S.A.

(L. Oxley) DEPARTMENT OF ECONOMICS, CANTERBURY UNIVERSITY, CHRISTCHURCH, NZ

Departamento de Economia PUC-Rio
Pontifícia Universidade Católica do Rio de Janeiro
Rua Marques de Sâo Vicente 225 - Rio de Janeiro 22453-900, RJ
Tel.(21) 35271078 Fax (21) 35271084
www.econ.puc-rio.br
flavia@econ.puc-rio.br