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

NONLINEAR COINTEGRATION, MISSPECIFICATION AND BIMODALITY

MARCELO C. MEDEIROS, EDUARDO MENDES, AND LES OXLEY

ABSTRACT. We show that the asymptotic distribution of the ordinary least squares estimator in a cointegration regression may be bimodal. 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 generalises to more complicated nonlinear models involving integrated time series.

KEYWORDS: Cointegration, nonlinearity, bimodality, misspecification, instrumental variables, 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 grateful to Peter C.B. Phillips for many enlightening discussions. The research of the first author is partially supported by CNPq/Brazil and the third by the Marsden Fund of New Zealand.

1. INTRODUCTION

One area of econometrics that has recently expanded is that of nonlinear modeling and nonlinear cointegration in particular see for example, Teräsvirta, Tjøstheim, and Granger (2009), Choi and Saikkonnen (2004a,b), Juhl and Xiao (2005), Bierens and Martins (2009), Cai, Li, and Park (2009), and Xiao (2009). These authors forcefully argue, and richly illustrate, why and how nonlinearities form the basis of a variety of economic theories. Their illustrations include a wide range of disequilibrium models; labour market models; exchange rate models; and production function modelling.

In this paper we discuss several interesting issues that can emerge in possible nonlinear cointegration models. Although we focus on simple nonlinear-in-variables cointegrated regression models, our core results can be easily generalised to more complex nonlinear formulations. We provide conditions under which nonstandard asymptotic distributions arise when estimating the parameters of the model. More specifically, we show that the ordinary least squares (OLS) estimator might be inconsistent and

Date: January 5, 2010.

with an asymptotic distribution which is bimodal. The existence of bimodality has recently been considered by Phillips (2006), Hillier (2006) and Forchini (2006). In their work bimodality typically occurs due to weak instrumentation. This is not the case presented here where bimodality arises analogously to that reported in Phillips and Hajivassiliou (1987) and Fiorio, Hajivassiliou, and Phillips (2008). We derive a Instrumental Variables (IV) estimator which is consistent and asymptotically distributed as a mixed normal.

Bimodality arises also in simpler models. For example, we show that when an intercept is erroneously omitted from a linear cointegrated regression, the distribution of the OLS estimator of the slope parameter is bimodal. This has also an adverse effect in the converge rate of the estimator as well as in the distribution of the t statistic.

The rest of the paper is organised as follows. Section 2 presents the simplest case arising from the erroneous omission of the intercept in the cointegrating relationship. Section 3 presents the general result which permits nonlinearity, potential endogeneity, and a generalised error structure. We also discuss the IV estimator of the model. Section 4 presents the simulation results and Section 5 concludes.

2. A Simple Result

In this section we report the results concerning a very simple misspecification problem: the omission of an intercept in a regression with cointegrated variables. More specifically, consider the following assumption.

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 also that the partial sum processes* $\boldsymbol{S}_{\boldsymbol{v},T}(r) = \sum_{i=1}^{[Tr]} \boldsymbol{v}_i$ *and* $S_{u,T}(r) = \sum_{i=1}^{[Tr]} u_i$, $r \in [0,1]$, *constructed from* $\{\boldsymbol{v}_t\}_{t=1}^{\infty}$ *and* $\{u_t\}_{t=1}^{\infty}$, *respectively, satisfy the multivariate invariance principle*[1]. *Specifically, define* $\boldsymbol{X}_{\boldsymbol{v},T}(r) = \sqrt{T}\boldsymbol{S}_{\boldsymbol{v},T}(r)$ *and* $X_{u,T}(r) = \sqrt{T}S_{u,T}(r)$, *then*

$$\begin{aligned} \boldsymbol{X}_{\boldsymbol{v},T}(r) &\Rightarrow \boldsymbol{B}_{\boldsymbol{v}}(r), \quad \text{as } T \to \infty, \\ X_{u,T}(r) &\Rightarrow \sigma_u^2 W_u(r), \quad \text{as } T \to \infty, \end{aligned} \tag{1}$$

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

where $\boldsymbol{B}_{\boldsymbol{v}}(r) \in \mathbb{R}^{k_x}$ *is a multivariate Brownian motion with covariance matrix*

$$\boldsymbol{\Omega} = \lim_{T \to \infty} \frac{1}{T} \mathbb{E}\left[\boldsymbol{S}_T(r)\boldsymbol{S}_T(r)'\right]$$

and $W_u(r)$ *is a standard Brownian motion. Finally, assume that* $W_u(r)$ *is independent of* $\boldsymbol{B}_{\boldsymbol{v}}(r)$.[2]

Now, suppose that an econometrician estimates the regression described in Assumption 1 by OLS without including the intercept [3]. Proposition 1 presents the asymptotic distribution of the OLS estimator.

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*

$$\sqrt{T}\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \Rightarrow \alpha \left[\int_0^1 \boldsymbol{B}_{\boldsymbol{v}}(r)\boldsymbol{B}_{\boldsymbol{v}}(r)'dr\right]^{-1} \int_0^1 \boldsymbol{B}_{\boldsymbol{v}}(r)dr. \tag{2}$$

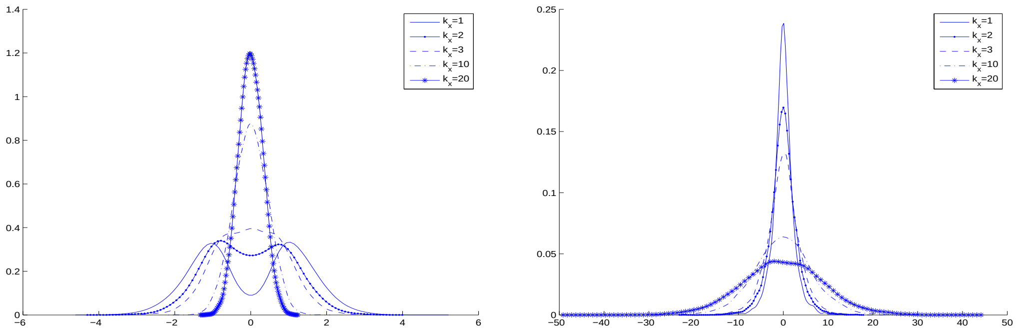
Several interesting features emerge from the above result. First, the OLS estimator is no longer super-consistent, as the convergence rate is $\sqrt{T}$. This will have serious implications in hypothesis testing. Second, the distribution in Proposition (1) may display bimodality in certain cases. For example, Figure 1, panel (a), displays the first marginal component of the asymptotic distribution in (2) for different dimensions, k_x, of the Brownian motion $\boldsymbol{B}_{\boldsymbol{v}}(r)$ [4]. 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, contrary to the previous case, the variance increases as the $k_x \longrightarrow \infty$. Figure 2 displays the variance of the first component of the asymptotic distribution of the OLS estimator as a function of the dimension of the vector Brownian process. Panel (a) refers to the case where $\alpha = 1$ while Panel (b) refers to the case where $\alpha = 0$.

---

[2]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 3.

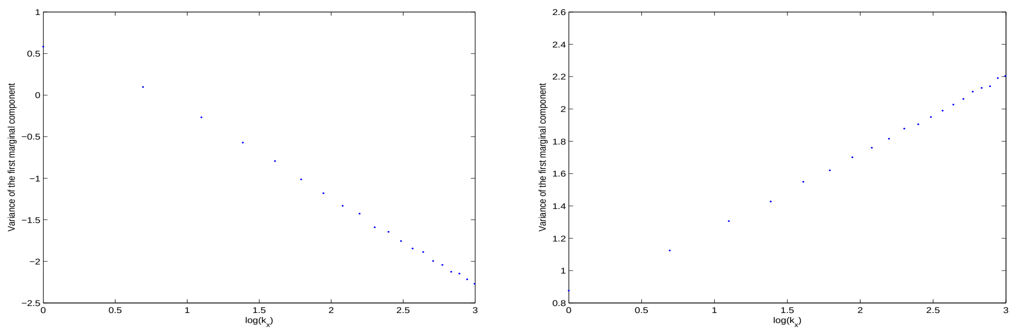
[3]When testing for the purchasing power parity (PPP) hypothesis (where the intercept is zero by definition) or synchronous dynamics among commodity prices, for example, it is not rare to find empirical models omitting the intercept.

[4]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.



(a) (b)

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 incorrectly omitted from the estimated regression. Panel (b) $\alpha = 0$.



(a) (b)

FIGURE 2. Scatter plot of the log variance of the first component of the asymptotic distribution of the OLS estimator versus the log dimension of the Brownian process. Panel (a) $\alpha \neq 0$ and it is incorrectly omitted from the estimated regression. Panel (b) $\alpha = 0$.

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 $\beta = \beta^*$,*

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

TABLE 1. Empirical size of the t-test.

The table shows the rejection rates of the null hypothesis $\mathcal{H}_0 : \beta = 1$ when a t-test is used. The nominal significance level is 5%.

Sample size	$\alpha = 0$	$\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 1$	$\alpha = 10$
50	0.052	0.051	0.056	0.088	0.850	0.991
100	0.044	0.045	0.064	0.124	0.919	0.997
250	0.049	0.049	0.083	0.188	0.953	0.997
500	0.040	0.041	0.104	0.332	0.963	0.995
1000	0.039	0.054	0.189	0.561	0.973	0.996
5000	0.042	0.066	0.626	0.885	0.995	1.000

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.

Although the above results are quite simple, the message is important and direct: Never omit the intercept in a cointegrating regression. Furthermore, to our knowledge this is the first paper addressing this issue.

2.1. **Simulation Evidence.** To illustrate the results above we conduct a simple simulation exercise. Consider the following data generating process (DGP):

$$y_t = \alpha + x_t + u_t$$

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

where $u_t \sim \mathsf{NID}(0, 1)$, $v_t \sim \mathsf{NID}(0, 1)$, and $\mathbb{E}(u_t v_\tau) = 0, \forall\, t, \tau$. The DGP was simulated for different values of α. We consider the estimation of the slope coefficient when the intercept is erroneously omitted from the estimated regression. We study the effects on the empirical size of the t-test for the null hypothesis $\mathcal{H}_0 : \beta = 1$ at the usual 5% significance level. The results are shown in Table 1. The table shows the rejection rates of the null hypothesis when it is in fact true. As expected, when $\alpha = 0$, the rejection rates are close to the nominal size of 5%. However, the distortions tend to be large as the value of the omitted intercept grows. For example, even for a reasonably small value of α, such as $\alpha = 0.05$, the rejection rates can be as high as 8% for 250 observations or almost 19% for 1000 observations.

3. A SIMPLE NONLINEAR-IN-VARIABLES MODEL

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 variables $\boldsymbol{z}_t$. We assume that $\boldsymbol{z}_t$ is observable and second-order stationary. More specifically consider the following assumption.

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

$$y_t = \alpha_0 + \beta_0 x_t + \alpha_1 g(\boldsymbol{z}_t) + \beta_1 x_t g(\boldsymbol{z}_t) + u_t, \tag{4}$$

$$x_t = x_{t-1} + v_t, \tag{5}$$

$$u_t = \sum_{j=0}^{\infty} \pi_{u,j}\varepsilon_{1,t-j} = \pi_u(L)\varepsilon_{1,t}, \tag{6}$$

$$v_t = \sum_{j=0}^{\infty} \pi_{v,j}\varepsilon_{2,t-j} = \pi_v(L)\varepsilon_{2,t}, \text{ and} \tag{7}$$

$$\boldsymbol{z}_t = \sum_{j=0}^{\infty} \boldsymbol{\pi}_{\boldsymbol{z},j}\boldsymbol{\varepsilon}_{3,t-j} = \boldsymbol{\pi}_{\boldsymbol{z}}(L)\boldsymbol{\varepsilon}_{3,t}, \tag{8}$$

where $\pi_u(L)$, $\pi_v(L)$, *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 (4) may arise in a number of situations, as for example, in threshold cointegrating regressions where the threshold is known are special cases of (4). 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, (4) becomes $y_t = \alpha_0 + \beta_0 x_t + \alpha_1 d_t + \beta_1 x_t d_t + u_t$.

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 $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g(\boldsymbol{z}_t) \xrightarrow{d} \mathsf{N}(\mu_g, \omega_g^2)$, 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 $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} g(\boldsymbol{z}_t)u_t \xrightarrow{d} \mathsf{N}(\mu_{gu}, \omega_{gu}^2)$, 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* $2 \leq a < \infty$*;*

(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, such that the partial sum process $\boldsymbol{S}_T(r) = \sum_{i=1}^{[Tr]} \boldsymbol{w}_i$, $r \in [0,1]$, constructed from $\{\boldsymbol{w}_t\}_{t=1}^{\infty}$ satisfies the multivariate invariance principle. Specifically, define $\boldsymbol{X}_T(r) = \sqrt{T}\boldsymbol{S}_T(r)$, then $\boldsymbol{X}_T(r) \Rightarrow \boldsymbol{B}(r)$, 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*

$$\boldsymbol{\Omega} = \begin{pmatrix} \omega_u^2 & \omega_{vu} & \omega_{gu} & \omega_{g^2u} & \omega_{guu} \\ \omega_{vu} & \omega_v^2 & \omega_{gv} & \omega_{g^2v} & \omega_{guv} \\ \omega_{gu} & \omega_{gv} & \omega_g^2 & \omega_{g^2g} & \omega_{ggu} \\ \omega_{g^2u} & \omega_{g^2v} & \omega_{g^2g} & \omega_{g^2}^2 & \omega_{g^2gu} \\ \omega_{guu} & \omega_{guv} & \omega_{ggu} & \omega_{g^2gu} & \omega_{gu}^2 \end{pmatrix} = \boldsymbol{\Sigma} + \boldsymbol{\Lambda} + \boldsymbol{\Lambda}', \tag{9}$$

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 $\boldsymbol{\theta} = (\alpha_0, \beta_1, \alpha_1, \beta_1)'$ and consider the OLS estimator

$$\widehat{\boldsymbol{\theta}} = \begin{bmatrix} \sum_{t=1}^T 1 & \sum_{t=1}^T x_t & \sum_{t=1}^T g(\boldsymbol{z}_t) & \sum_{t=1}^T g(\boldsymbol{z}_t)x_t \\ \cdot & \sum_{t=1}^T x_t^2 & \sum_{t=1}^T g(\boldsymbol{z}_t)x_t & \sum_{t=1}^T g(\boldsymbol{z}_t)x_t^2 \\ \cdot & \cdot & \sum_{t=1}^T g(\boldsymbol{z}_t)^2 & \sum_{t=1}^T g(\boldsymbol{z}_t)^2 x_t \\ \cdot & \cdot & \cdot & \sum_{t=1}^T g(\boldsymbol{z}_t)^2 x_t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^T y_t \\ \sum_{t=1}^T x_t y_t \\ \sum_{t=1}^T g(\boldsymbol{z}_t) y_t \\ \sum_{t=1}^T x_t g(\boldsymbol{z}_t) y_t \end{bmatrix}.$$

The distribution of $\widehat{\boldsymbol{\theta}}$ changes according to the properties of the function $g(\mathbf{z}_t)$ as is illustrated in the following theorem.

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

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

where $\Delta_{vu} = \sigma_{vc} + \lambda_{vu}$ *and*

$$\boldsymbol{\Gamma} = \begin{bmatrix} T^{1/2} & 0 & 0 & 0 \\ 0 & T & 0 & 0 \\ 0 & 0 & T^{1/2} & 0 \\ 0 & 0 & 0 & T^{1/2} \end{bmatrix}.$$

On the other hand, if $\mu_g = 0$

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

where $\Delta_{gv} = \sigma_{gv} + \lambda_{gv}$ *and* $\boldsymbol{\Gamma}$ *is as above.*

Two important features emerge from Theorem 1. First, as expected, the OLS estimate of α_1 is not consistent when $\boldsymbol{z}_t$ is endogenous. Second, the asymptotic distribution of $\widehat{\boldsymbol{\beta}}_1$ may also be bimodal.

3.1. **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 $\mathbb{E}\left(s_t^2\right) = m_s^2 < \infty$. Furthermore, the partial sum process $S_{su,T}(r) = \sum_{i=1}^{[Tr]} s_i u_i$ constructed from $\{s_t u_t\}_{t=1}^{\infty}$ satisfies the invariance principle. Specifically, define $X_{su,T}(r) = \sqrt{T} S_{su,T}(r)$, then $X_{su,T}(r) \Rightarrow \omega_{su} W_{su}(r), \quad$ as $T \to \infty$, where $W_{su}(r)$ 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 IV estimator of $\boldsymbol{\theta}$ is given by

$$\widetilde{\boldsymbol{\theta}} = \begin{bmatrix} \sum_{t=1}^T 1 & \sum_{t=1}^T x_t & \sum_{t=1}^T \widehat{g}_t & \sum_{t=1}^T \widehat{g}_t x_t \\ \cdot & \sum_{t=1}^T x_t^2 & \sum_{t=1}^T \widehat{g}_t x_t & \sum_{t=1}^T \widehat{g}_t x_t^2 \\ \cdot & \cdot & \sum_{t=1}^T \widehat{g}_t^2 & \sum_{t=1}^T \widehat{g}_t^2 x_t \\ \cdot & \cdot & \cdot & \sum_{t=1}^T \widehat{g}_t^2 x_t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=1}^T y_t \\ \sum_{t=1}^T x_t y_t \\ \sum_{t=1}^T \widehat{g}_t y_t \\ \sum_{t=1}^T x_t \widehat{g}_t y_t \end{bmatrix}. \tag{10}$$

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 $\mathbf{\Gamma}$ *is given by*

$$\mathbf{\Gamma} = \begin{bmatrix} T^{1/2} & 0 & 0 & 0 \\ 0 & T & 0 & 0 \\ 0 & 0 & T^{1/2} & 0 \\ 0 & 0 & 0 & T \end{bmatrix}$$

in both cases.

4. Monte Carlo Simulations

In this section we present simulation evidence of the results in the previous sections. Consider two cases of the general model described in equations (4)–(8).

(1) Model 1: Identity function.

$$\begin{aligned} y_t &= \alpha_0 + \alpha_1 z_t + \beta_0 x_t + \beta_1 x_t z_t + u_t \\ &= 1 + x_t + \alpha_1 z_t + x_t z_t + u_t, \\ z_t &= s_t + u_t, \\ x_t &= x_{t-1} + v_t, \end{aligned}$$

where $u_t \sim \mathsf{NID}(0,1)$, $v_t \sim \mathsf{NID}(0,1)$, $s_t \sim \mathsf{NID}(0,1)$, $\mathbb{E}(u_t v_\tau) = 0$, $\forall\, t, \tau$, and $\mathbb{E}(s_t u_\tau) = 0$, $\forall\, t, \tau$.

(2) Model 2: Threshold function.

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

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

For both DGPs, we consider two cases: $\alpha_1 = 0$ and $\alpha_1 = 1$. In Model 1, $g(\boldsymbol{z}_t) = z_t$ whereas in Model 2, $g(\boldsymbol{z}_t) = I(z_t > 0)$. We simulate 5000 observations of each model over 1000 Monte Carlo replications and evaluate the distribution of both the OLS and IV estimators of the parameters α_0, α_1, β_0, and β_1. In Model 1, the first stage regression for the IV estimator consists of regressing z_t on s_t while for the second model, in the first stage the $I(z_t > 0)$ is regressed on $I(s_t > 0)$. The results are shown in Figures 3–6. We also consider the distribution of the t-statistic under the null hypothesis as shown in Figures 7–10.

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 (and for both models), 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. In Model 1, as expected, the OLS estimator of α_1 is not consistent for the true parameter, while the IV counterpart is. When Model 2 is considered, 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.

5. 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 endogeneity and a generalised error structure are considered, 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 Instrumental Variables (IV) in these cases removes both bimodality and inconsistency.

References

Bierens, H., and L. Martins (2009): "Time Varying Cointegration," *Econometric Theory*, forthcoming.

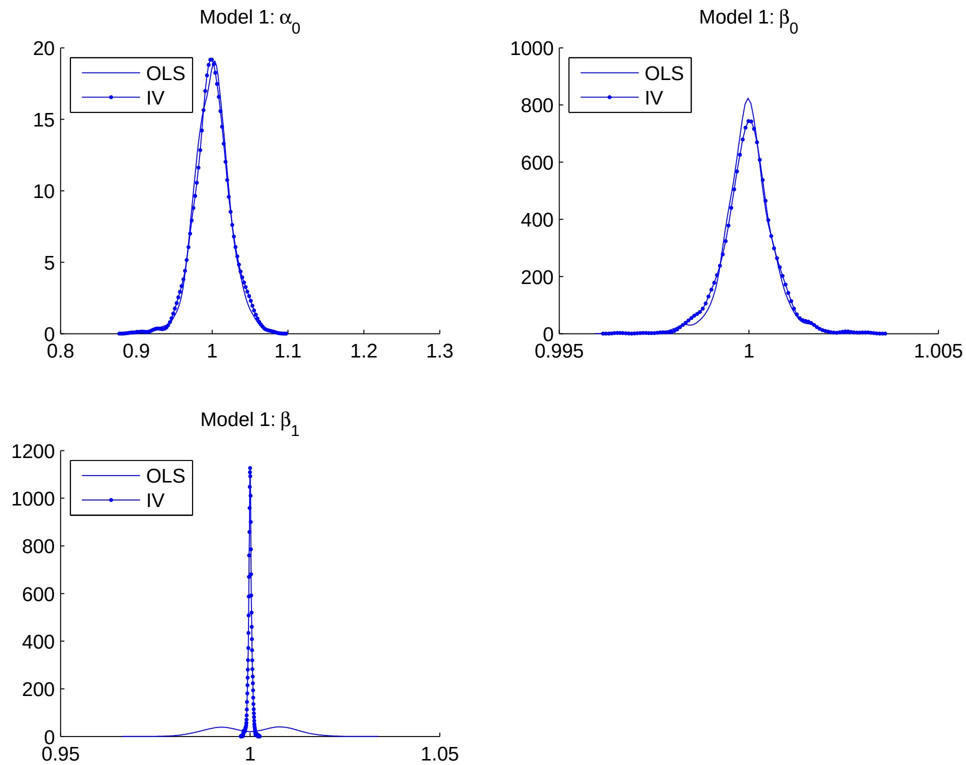


FIGURE 3. Empirical distribution of the OLS and IV estimators in Model 1 with $\alpha_1 = 0$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

CAI, Z., Q. LI, AND J. PARK (2009): "Functional-Coefficient Models for Nonstationary Time Series Data," *Journal of Econometrics*, 148, 101–113.

FIORIO, C., V. HAJIVASSILIOU, AND P. PHILLIPS (2008): "Bimodal t-ratios: The impact of thick tails on inference," Unpublished manuscript, London School of Economics.

FORCHINI, G. (2006): "On the bimodality of the exact distribution of the TSLS estimator," *Econometric Theory*, 22, 932–946.

HILLIER, G. (2006): "Yet more on the exact properties of IV estimators," *Econometric Theory*, 22, 913–931.

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

JUHL, T., AND Z. XIAO (2005): "Testing for cointegration using partially linear models," *Journal of Econometrics*, 124, 363–394.

PHILLIPS, P. (2006): "A remark on bimodality and weak instrumentation in structural equation estimation," *Econometric Theory*, 22, 947–960.

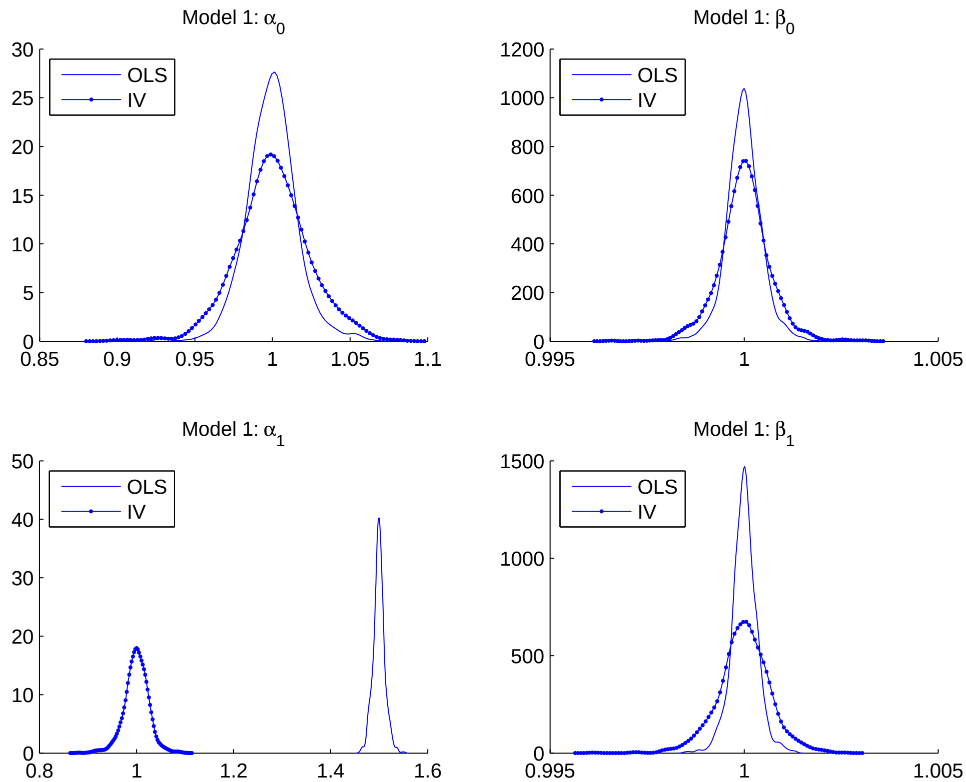


FIGURE 4. Empirical distribution of the OLS and IV estimators in Model 1 with $\alpha_1 = 1$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

PHILLIPS, P., AND V. HAJIVASSILIOU (1987): "Bimodal t-ratios," Dicussion Paper 842, Cowles Foundation for Research in Economic, Yale University.

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.

TERÄSVIRTA, T., T. TJØSTHEIM, AND C. GRANGER (2009): *Nonlinear econometric modeling*. Oxford Univristy Press, forthcoming.

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*

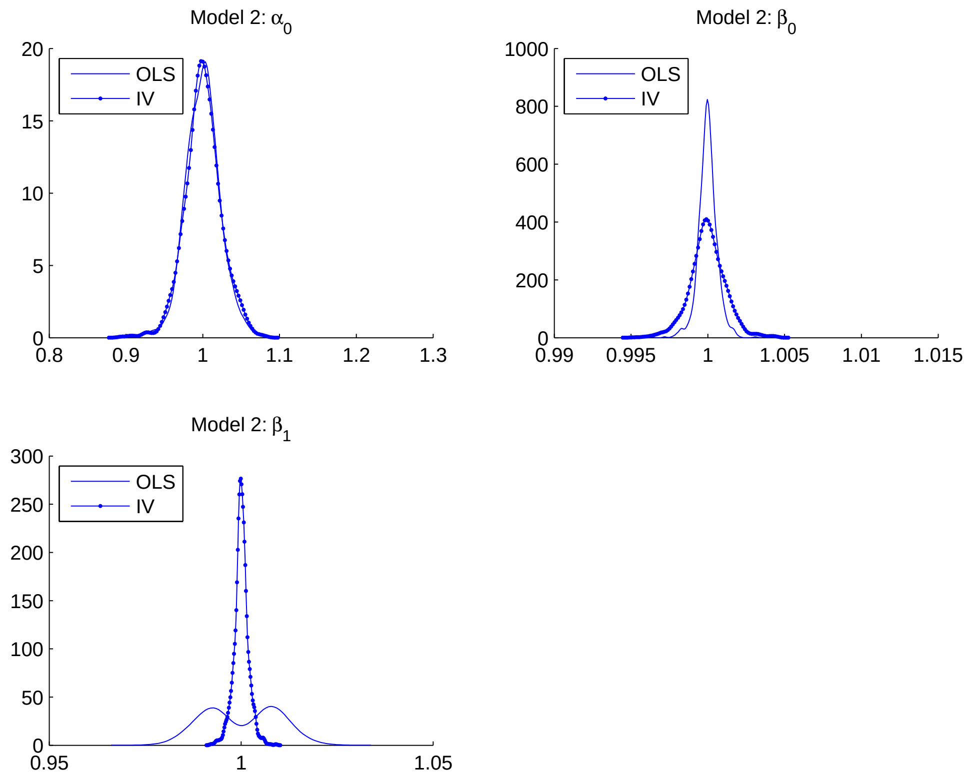


FIGURE 5. Empirical distribution of the OLS and IV estimators in Model 2 with $\alpha_1 = 0$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

that the partial sum process $\boldsymbol{S}_T(r) = \sum_{j=1}^{[Tr]} \boldsymbol{w}_j$, $r \in [0,1]$, *constructed from the stationary innovation process* $\{\boldsymbol{w}_t\}_{t=1}^T$, *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*

$$\boldsymbol{\Omega} = \begin{pmatrix} \omega_u^2 & \omega_{vu} \\ \omega_{vu} & \omega_v^2 \end{pmatrix} = \mathbb{E}(\boldsymbol{w}_1\boldsymbol{w}_1') + \sum_{k=2}^{\infty} \left[\mathbb{E}(\boldsymbol{w}_1\boldsymbol{w}_k') + \mathbb{E}(\boldsymbol{w}_k\boldsymbol{w}_1')\right] = \boldsymbol{\Sigma} + \boldsymbol{\Lambda} + \boldsymbol{\Lambda}'. \tag{11}$$

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* $T^{-3/2}\sum_{t=1}^T x_t u_t \Rightarrow \mu_u \int_0^1 B_v(r)dr$.

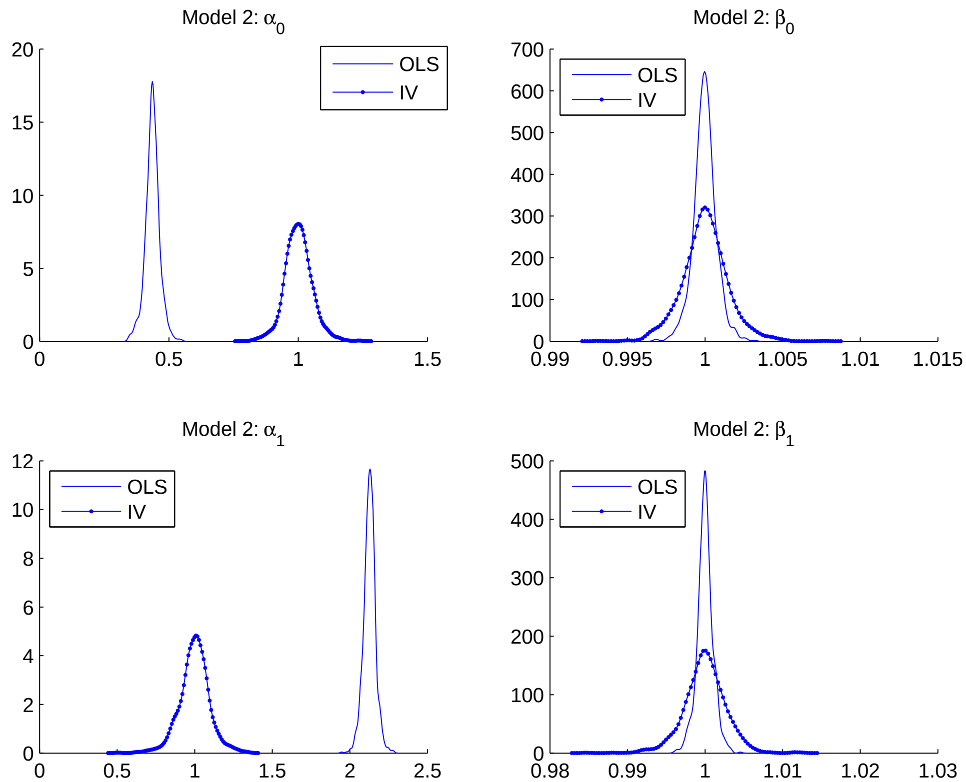


FIGURE 6. Empirical distribution of the OLS and IV estimators in Model 2 with $\alpha_1 = 1$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

Proof. First, define $u_t^* = u_t - \mu_u$ and write

$$\sum_{t=1}^{T} x_t^2 u_t = \sum_{t=1}^{T} x_t^2 \left(\mu_u + u_t^*\right) = \mu_u \sum_{t=1}^{T} x_t^2 + \sum_{t=1}^{T} x_t^2 u_t^*.$$

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 $\sum_{t=1}^{T} x_t u_t = \sum_{t=1}^{T} x_t \left(\mu_u + u_t^*\right) = \mu_u \sum_{t=1}^{T} x_t + \sum_{t=1}^{T} x_t u_t^*$, 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 $\frac{1}{T} \sum_{t=1}^{T} x_t u_t^* \Rightarrow \int_0^1 B_v(r) dB_u(r) dr + \Delta_{uv}$.

■

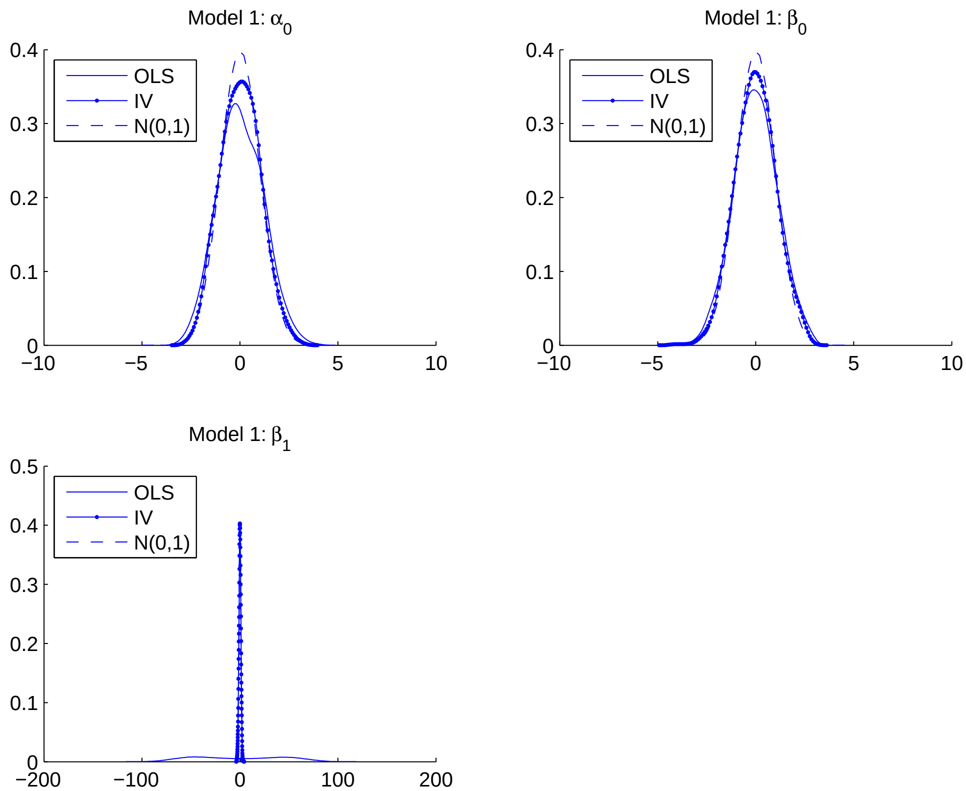


FIGURE 7. Empirical distribution of t-statistic for the OLS and IV estimators in Model 1 with $\alpha_1 = 0$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

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 $T^{-3/2}\sum_{t=1}^{T} \boldsymbol{x}_t u_t \xrightarrow{p} \mathbf{0}$,

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

■

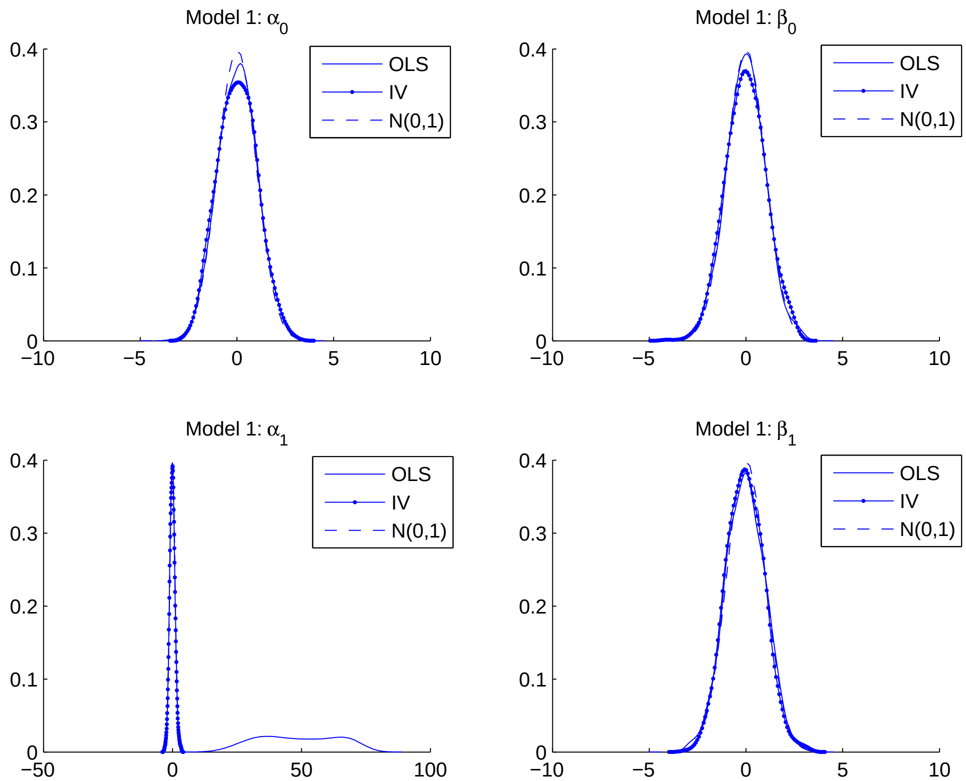


FIGURE 8. Empirical distribution of t-statistic for the OLS and IV estimators in Model 1 with $\alpha_1 = 1$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

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

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

Hence,

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

■

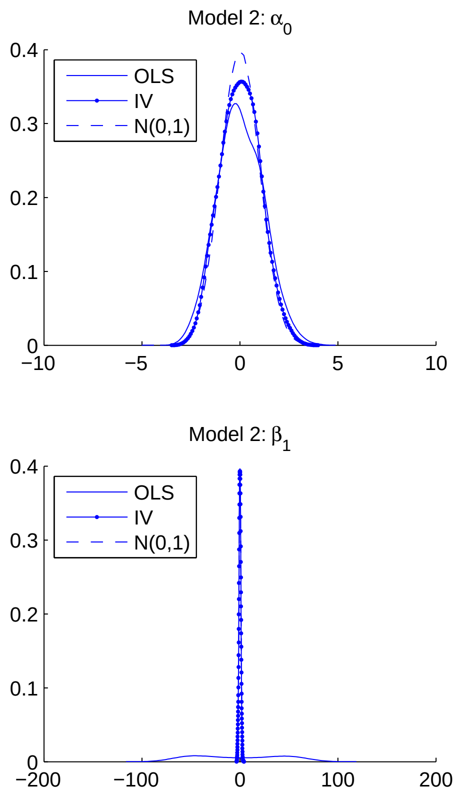


FIGURE 9. Empirical distribution of t-statistic for the OLS and IV estimators in Model 2 with $\alpha_1 = 0$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

B.3. **Proof of Theorem 1.** First, define the following matrices:

$$\boldsymbol{H} = \begin{bmatrix} \sqrt{T} & 0 & 0 & 0 \\ 0 & T & 0 & 0 \\ 0 & 0 & \sqrt{T} & 0 \\ 0 & 0 & 0 & T \end{bmatrix} \text{ and } \boldsymbol{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \sqrt{T} \end{bmatrix}.$$

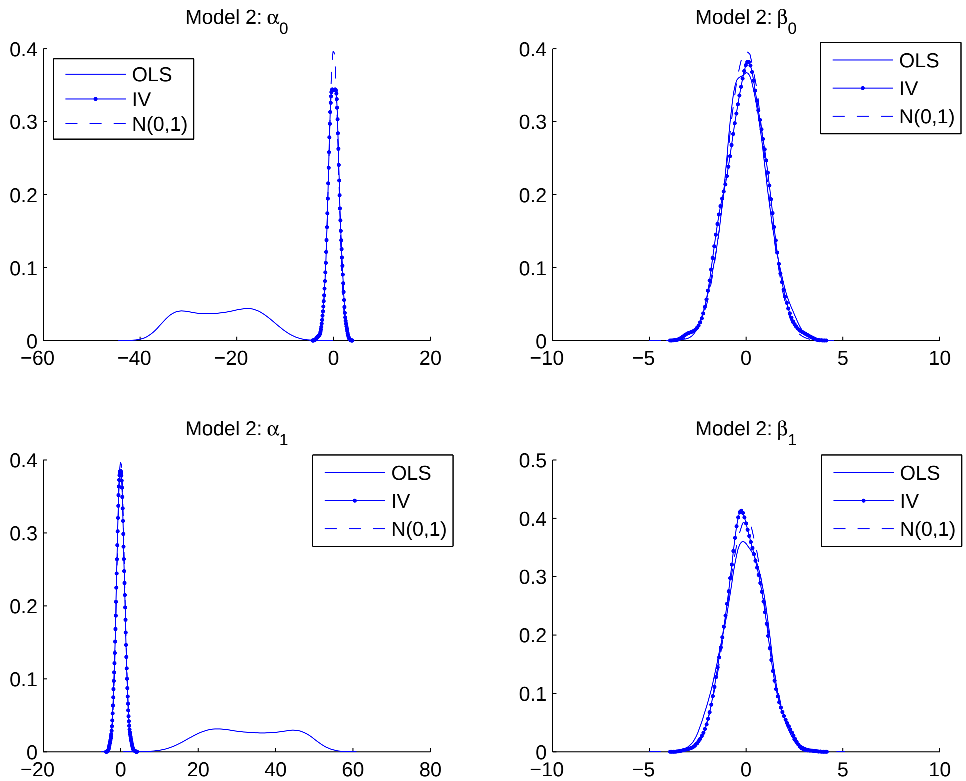


FIGURE 10. Empirical distribution of t-statistic for the OLS and IV estimators in Model 2 with $\alpha_1 = 1$. The data are simulated with 5000 observations and the Monte Carlo is conducted with 1000 replications.

Note that

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

$$\times\,\boldsymbol{D}^{-1}\boldsymbol{H}^{-1}\begin{bmatrix}\sum_{t=1}^{T}u_t\\ \sum_{t=1}^{T}x_tu_t\\ \sum_{t=1}^{T}g(\boldsymbol{z}_t)u_t\\ \sum_{t=1}^{T}x_tg(\boldsymbol{z}_t)u_t\end{bmatrix}.$$

Hence,

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

$$\times\begin{bmatrix}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}u_t\\ \frac{1}{T}\sum_{t=1}^{T}x_tu_t\\ \frac{1}{\sqrt{T}}\sum_{t=1}^{T}g(\boldsymbol{z}_t)u_t\\ \frac{1}{T^{3/2}}\sum_{t=1}^{T}x_tg(\boldsymbol{z}_t)u_t\end{bmatrix}.$$

Therefore, for $\mu_g = 0$ 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

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

From the results in Lemma 1, it follows that

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

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

Now, let's turn to $\sum_{t=1}^{T} g_t^2 x_t$. Again, $\sum_{t=1}^{T} g_t^2 x_t = \sum_{t=1}^{T} \left(m_g^2 + g_t^{2*}\right) x_t = m_g^2 \sum_0^T x_t + O(T)$. Hence, $\frac{1}{T^{3/2}} \sum_{t=1}^{T} g_t^2 x_t \Rightarrow m_g^2 \int_0^1 B_v(r) dr$ and (g) follows.

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

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

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

$$\boldsymbol{\Gamma} = \begin{bmatrix} T^{1/2} & 0 & 0 & 0 \\ 0 & T & 0 & 0 \\ 0 & 0 & T^{1/2} & 0 \\ 0 & 0 & 0 & T \end{bmatrix}$$

and write

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

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