LINEAR REGRESSION WITH DETERMINISTIC REGRESSORS AND UNIT ROOT IN THE VARIANCE

ALEXANDRE PETKOVIC*

Received December 13, 2010; revised December 27, 2010

ABSTRACT. The first part of this paper derives the asymptotic distribution of the ordinary least squares estimator in a linear regression model with deterministic regressors when the variance of the innovations is a function of an integrated time series. In the second part of this paper we study the impact of heteroscedasticity on the standard t-test for the slope coefficient in a linear trend model.

1 Introduction The tools for the study of a linear system of an integrated time series were introduced by Phillips (1986,1987) and Phillips and Durlauf (1986). Their results relied on weak convergence in functional spaces, the continuous mapping theorem and on weak convergence of stochastic integrals to martingales. It is with the papers of Phillips and Park (1999, 2001) that the study of the asymptotic behavior of nonlinear functions of an integrated time series started. Phillips and Park (1999, 2001) derived the asymptotic distribution of the average of a nonlinear function of integrated time series. These results where further extended by Chang and Park (2003), Jong and Wang (2005) and Shi and Phillips (2010).

The results obtained by Park and Phillips (1999, 2001) have been applied to various nonlinear econometric models. For example, Park and Phillips (2001) and Shi and Phillips (2010) used them to derive the distribution of the least squares estimator in a nonlinear regression model. Chang and all (2001) considered nonlinear regressions with separably additive regression functions. Park (2002) studied the possibility of modeling assets variance using a nonlinear function of an integrated time series. Studying the USD/DM exchange rate he found out that the conditional variance of the spread can be accurately modeled using the spot rate. Chung and Park (2003) considered nonstationary index models. Hu and Phillips (2004) worked on discrete choice models. Chang and Park (2007) studied the distribution of the ordinary least squares estimator of a linear regression model with integrated or stationary regressors when the error term volatility is a nonlinear function of an integrated time series. They showed that, when the volatility of the error term is a function of an integrated time series, the asymptotic distribution of the ordinary least squares is nonstandard and involves an integral with respect to the local time of a Brownian motion at the origin. Note also that an earl study of the linear regression model where the variance of the innovation is a function of an integrated time series can be found in Hansen (1995).

The objective of this paper is twofold. Firstly, we extend some results of Phillips and Park (1999,2001) by deriving the asymptotic distribution of a temporally weighted average of a function of an integrated time series. Secondly, we use our results to study the distribution of the ordinary least squares estimator of linear regression model when the regressors are time-deterministic and the variance of the error term is a function of an integrated time series. Using these results we also study the asymptotic distribution of the standard t-stat. In this sense our results can be seen as an extension of those derived by Chung and Park (2007). As explained bellow potential applications of our results can be found in macroeconometrics and finance.

The paper is organized as follows: Section 2 presents the model and the assumptions, Section 3 studies the asymptotic distribution of the ordinary least squares estimator, in Section 4 we consider an application

2000 *Mathematics Subject Classification.* 62J05.
Key words and phrases. nonlinear, integrated time series, deterministic regressors, regression with heteroskedastic errors.

of our theory and study the impact of heteroscedasticity on the level of the standard t-test, finally Section 5 concludes. All the proofs can be found in the appendix. Throughout this paper we will use the following notations: $\rightarrow_d$ and $\rightarrow_p$ will mean convergence in distribution and in probability, respectively. $\mathbb{N}^*$ and $\mathbb{N}$ will denote the integer and the non-negative integer, respectively.

2 The model and Assumptions

Consider the regression model given by

(1) $$y_{n,t} = \alpha + g_{1,n}(t)\beta_1 + \cdots + g_{k,n}(t)\beta_k + \epsilon_{n,t}, \qquad t = 0, 1, \ldots, n$$

where $y_{n,t}$ is the depend variable and the $g_{i,n}(t)$'s are deterministic regressors. The error term, $\epsilon_{n,t}$, is modeled as

$$\epsilon_{n,t} = \sigma(z_t)u_{n,t},$$

where $u_{n,t}$ is a martingale difference sequence with mean zero and unit variance with respect to a filtration $\mathcal{F}_{n,t}$, z_t is an integrated time series and σ a function whose properties will be specified bellow. We assume that z_t is measurable with respect to $\mathcal{F}_{n,t-1}$ implying that $(\epsilon_{n,t}, \mathcal{F}_{n,t})$ is a martingale difference sequence satisfying

$$E(\epsilon_{n,t}^2|\mathcal{F}_{n,t-1}) = \sigma^2(z_t).$$

Let $[s]$ be the larger integer smaller than s. Through this paper we will assume that each deterministic regressor $g_{i,n}(t)$ satisfies the following assumption.

Assumption 1. *Let $g_{i,n}(t)$, $t = 0, 1, \ldots, n$, be a sequence of finite valued deterministic regressors. Then there exists a positive function $c_i(n)$, whose limit as $n \rightarrow \infty$ exists in $\bar{\mathbb{R}}$, such that*

$$sup_{r\in[0,1]}|\frac{g_{i,n}([rn])}{c_i(n)} - \breve{g}_i(r)| \rightarrow 0,$$

where $\breve{g}_i(r)$ is piecewise continuous on $[0, 1]$ and satisfies $\int_0^1 |\breve{g}_i(r)|dr \neq 0$.

Assumption 1 is not restrictive since several standard regressors satisfy it. Consider for example the linear trend model in which $g_n(i) = i$, in this case we can set $c(n) = n$ with $\breve{g}(r) = r$. Another example is when the regressors are given $g_n(i) = \cos((2\pi i)/n)$. In this case we can set $c(n) = 1$ and $\breve{g}(r) = \cos(2\pi r)$.

We assume that the process z_t is of the form

(2) $$z_t = z_{t-1} + w_t,$$

where w_t follows the linear process

$$w_t = \psi(L)e_t = \sum_{k=0}^{+\infty} \psi_k e_{t-k},$$

where e_t is an iid sequence of random variables with mean zero. In this paper we set $z_0 = 0$.

In the proofs we will write

$$U_n(r) = \frac{1}{\sqrt{n}}\sum_{t=1}^{[rn]} u_{n,t} \qquad W_n(r) = \frac{1}{\sqrt{n}} z_{[rn]} = \frac{1}{\sqrt{n}}\sum_{t=1}^{[nr]} w_t.$$

We make the following assumptions on the error term.

Assumption 2. *(a) $(U_n, W_n) \to_d (U, W)$ as $n \to \infty$, where (U, W) is a vector Brownian motion.*
(b) $\sum_{k=0}^{+\infty} k|\psi_k| < \infty$, $\psi(1) \neq 1$ and $E(e_t^p) < \infty$ for some $p > 2$.
(c) The distribution of e_t is absolutely continuous with respect to the Lebesgue measure and has characteristic function $\phi(t)$ for which $\lim_{t\to\infty} t^r \phi(t) = 0$ for some $r > 0$.
Moreover for each n, there exists a filtration $(\mathcal{F}_{n,t})$, $t = 0, 1, \ldots, n$, such that
(d) $(u_{n,t}, \mathcal{F}_{n,t})$ is a martingale difference sequence with $E(u_{n,t}^2|\mathcal{F}_{n,t}) = 1$ a.s. for all, $t = 1, 2, ..., n$, and $sup_{1\leq t\leq n} E(|u_{n,t}|^q|\mathcal{F}_{n,t-1}) < \infty$ a.s. for $q > 2$.
(e) z_t is $\mathcal{F}_{n,t-1}$ measurable.

Assumptions $2(a)(d)$ are standard in the literature. Assumption $2(b)$ is a summability condition on the moving average coefficient of w_t and a moment condition on the innovation e_t, finally $2(c)$ is a technical assumption.

We will consider two kinds of volatility functions σ, more specifically:

Definition 1. *We let $\sigma \in \mathcal{I}$ if σ is Riemann integrable, $\int_{\infty}^{\infty} \sigma^2(s)ds < \infty$, $|\sigma(x) - \sigma(y)| \leq c|x - y|^l$, for some constant c and $l > 6/(p - 2)$, where p is defined in Assumption 2(b).*
On the other hand, we write $\sigma \in \mathcal{H}$ if

$$\sigma(\lambda s) = \nu(\lambda)\tau(s) + o(\nu(\lambda)),$$

for large λ uniformly over s and over any compact interval, where τ is locally Riemann integrable. For $\sigma \in \mathcal{H}$, we call ν and τ the asymptotic order and the limit homogenous function of σ, respectively.

The limit distribution of the ordinary least squares estimator involves the Brownian local time of W at the origin, which is defined as

$$L(t, 0) = \lim_{\epsilon\to 0} \frac{1}{2\epsilon} \int_0^t 1_{\{|W(r)|<\epsilon\}} dr,$$

Intuitively $2\epsilon L(t, 0)$ measures the time spent by W in an infinitesimal neighborhood of 0 during the time interval $[0, t]$. $L(t, 0)$ is monotone increasing and almost surely continuous in t, see Revuz and Yor (1990) for a discussion.

3 The Ordinary Least Squares Estimator

3.1 $\mathcal{I}$ regular function In this section we will derive the distribution of the ordinary least squares estimator of $\theta = (\alpha, \beta_1, \ldots, \beta_K)$. To do so we first need to generalize Theorem 3.2 from Phillips and Park (2001).

Theorem 1. *If Assumptions 1 and 2 are satisfied with $p > 4$ and $\sigma \in \mathcal{I}$ then*

$$\frac{1}{n^{1/4}} \sum_{t=1}^{n} \frac{g_n(t)}{c(n)} \sigma(z_t) u_{n,t} \to_d \left(\left(\int_{-\infty}^{+\infty} \sigma(s)ds \right) \left(\int_0^1 \breve{g}^2(s) dL(s, 0) \right) \right)^{1/2} B(1),$$

as $n \to \infty$, where $B(1)$ is a Brownian motion independent of W.

Setting $g_n(t) = c(n) = 1$ Theorem 1 we recover the original result of Phillips and Park (2001). In the appendix we give a proof of Theorem 1 that is based on the proof of Theorem 3.2 in Phillips and Park (2001). However, the difference in our proof with the one of Phillips and Park (2001) is that we define $\rho_n(r)$ using the quadratic variation of the process $\tilde{M}_n(r)$ instead of $M_n(r)$. This modification is necessary since if we had defined the stopping time $\rho_n(r)$ using the process $M_n(r)$ there may be a set of sample path with positive probability for which $\rho_n(r)$ is not well defined for some $r \in [0, 1]$.

Now we can derive the asymptotic distribution of the ordinary least squares estimator of model (1). Consider

$$\begin{pmatrix} \hat{\alpha}-\alpha \\ \hat{\beta}_1-\beta_1 \\ \dots \\ \hat{\beta}_k-\beta_k \end{pmatrix} = \begin{pmatrix} n & \sum_{t=1}^n g_{1,n}(t) \cdots \sum_{t=1}^n g_{k,n}(t) \\ \sum_{t=1}^n g_{1,n}(t) & \sum_{t=1}^n g_{1,n}^2(t) \cdots \sum_{t=1}^n g_{1,n}(t) g_{k,n}(t) \\ \vdots & \ddots \quad \vdots \\ \sum_{t=1}^n g_{k,n}(t) & \sum_{t=1}^n g_{1,n}(t) g_{k,n}(t) \cdots \sum_{t=1}^n g_{k,n}^2(t) \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=1}^n \epsilon_{n,t} \\ \sum_{t=1}^n g_{1,n}(t)\epsilon_{n,t} \\ \vdots \\ \sum_{t=1}^n g_{k,n}(t)\epsilon_{n,t} \end{pmatrix}.$$

If we let D_a be the following diagonal normalization matrix

$$D_a = \begin{pmatrix} n^{3/4} & 0 & \dots & 0 \\ 0 & n^{3/4}c_1(n) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & n^{3/4}c_k(n) \end{pmatrix},$$

then we have the following theorem

Theorem 2. *Consider model* (1) *and assume that Assumptions 1, 2 are satisfied with* $p > 4$ *and* $\sigma \in \mathcal{I}$*, then the ordinary least squares estimator of* θ *satisfies*

$$D_a \begin{pmatrix} \hat{\alpha}-\alpha \\ \hat{\beta}_1-\beta_1 \\ \vdots \\ \hat{\beta}_k-\beta_k \end{pmatrix} \to_d A^{-1}B_a,$$

where

$$A = \begin{pmatrix} 1 & \int_0^1 \breve{g}_1(s)ds & \dots & \int_0^1 \breve{g}_k(s)ds \\ \int_0^1 \breve{g}_1(s)ds & \int_0^1 \breve{g}_2^2(s)ds & \dots & \int_0^1 \breve{g}_1(s)\breve{g}_k(s)ds \\ \vdots & \vdots & \ddots & \vdots \\ \int_0^1 \breve{g}_k(s)ds & \int_0^1 \breve{g}_1(s)\breve{g}_k(s)ds & \dots & \int_0^1 \breve{g}_k^2(s)ds \end{pmatrix},$$

and

$$B_a = \begin{pmatrix} \left(\left(\int_{-\infty}^{+\infty}\sigma(s)ds\right) L(1,0)\right)^{1/2} B(1) \\ \left(\left(\int_{-\infty}^{+\infty}\sigma(s)ds\right)\left(\int_0^1 \breve{g}_1^2(s)dL(s,0)\right)\right)^{1/2} B(1) \\ \vdots \\ \left(\left(\int_{-\infty}^{+\infty}\sigma(s)ds\right)\left(\int_0^1 \breve{g}_k^2(s)dL(s,0)\right)\right)^{1/2} B(1) \end{pmatrix}.$$

Note that the ordinary least squares estimator of β_i will be consistent if $n^{3/4}c_i(n) \to \infty$.

Assume now that we want to test an hypothesis of the form

$$\begin{aligned} H_0 &: \beta_i = \beta_{0,i} \\ H_1 &: \beta_i \neq \beta_{0,i}, \end{aligned} \tag{3}$$

where $i \in \{1, \dots, k\}$ and $\beta_{0,i}$ is some constant. A usual test statistic for this problem is given by the t-stat whose expression is

$$t = \frac{\hat{\beta}_i - \beta_{0,i}}{\sqrt{\hat{\sigma}_\epsilon^2 R \left(\sum_{t=1}^n G_n(t)G_n(t)'\right)^{-1} R'}}, \tag{4}$$

where $G_n(t) = (1, g_{1,n}(t), \ldots, g_{k,n}(t))'$ and R is a $1 \times (k+1)$ vector whose only nonzero component is located in the $(i+1)^{th}$ column and $\hat{\sigma}_\epsilon^2$ is given by

$$\hat{\sigma}_\epsilon^2 = \frac{1}{n}\left(\sum_{t=1}^{n} \epsilon_{n,t}^2 - \left(\sum_{t=1}^{n} G_n'(t)\epsilon_{n,t}\right)\left(\sum_{t=1}^{n} G_n(t)G_n'(t)\right)^{-1}\left(\sum_{t=1}^{n} G_n(t)\epsilon_{n,t}\right)\right).$$

The next proposition derives the asymptotic distribution of t

Proposition 1. *Assume that $\sigma \in \mathcal{I}$ and consider the test statistic* (4) *associated to the hypothesis* (3). *If Assumption 1 and 2 are satisfied, then under H_0*

$$t \to_d \frac{RA^{-1}B_a}{\sqrt{C}},$$

where

$$C = L(1,0)\left(\int_{-\infty}^{\infty} \sigma^2(s)ds\right) RA^{-1}R',$$

while under H_1

$$t \to_d \infty, \quad a.s.$$

if and only if $c_i(n)n^{3/4} \to \infty$.

In particular we see from the above proposition that the t-test is consistent if and only if the ordinary least squares estimator is. In practice one is sometimes also interested in testing an hypothesis on a linear combination of the vector θ, i.e. $R\theta = b$ for $R \in \mathbb{R}^{1\times k+1}$. The usual test statistics in this case is the Wald test statistics. However in our case we could not derive the distribution of the Wald test statistics, the problem stems from the fact that the different components $\hat{\theta}$ do not converge at the same rate. This situation is similar to the testing problem in the homoskedastic linear trend model described in Hamilton (1994). In such a case the hypothesis $R\theta = b$ can be tested using the standard t-test, provided that $R \in \mathbb{R}^{1\times k+1}$. The derivation of the asymptotic distribution of this test can easily be carried along the lines of Hamilton (1994) and thus will be omitted here.

3.2 $\mathcal{H}$ Regular Function The next theorem will allow us to study the asymptotic least squares when the volatility function belongs to $\mathcal{H}$.

Theorem 3. *If Assumptions 1 and 2 are satisfied and $\sigma \in \mathcal{H}$ then*

$$\frac{1}{n^{1/2}\nu(\sqrt{n})}\sum_{t=1}^{n} \frac{g_n(t)}{c(n)}\sigma(z_t)u_{n,t} \to_d \int_0^1 \breve{g}(s)\tau(W(s))dU(s),$$

as $n \to \infty$.

Define

$$D_b = \frac{n^{1/2}}{\nu(n)}\begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & c_1(n) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & c_k(n) \end{pmatrix},$$

then computations similar to those of the previous subsection and the previous theorem leads to the following result.

Theorem 4. *Consider model* (1) *and assume that Assumptions 1 and 2 are satisfied with* $\sigma \in \mathcal{H}$*, then the ordinary least squares estimator satisfies*

$$D_b \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta}_1 - \beta_1 \\ \vdots \\ \hat{\beta}_k - \beta_k \end{pmatrix} \to_d A^{-1} B_b,$$

where

$$B_b = \begin{pmatrix} \int_0^1 \tau(W(s)) dU(s) \\ \int_0^1 \breve{g}_1(s)\tau(W(s)) dU(s) \\ \vdots \\ \int_0^1 \breve{g}_k(s)\tau(W(s)) dU(s) \end{pmatrix}.$$

Note that in this case the consistency of the ordinary least squares estimator depends on the behavior of the ratio $c_i(n)n^{1/2}/\nu(n)$. In particular the ordinary least squares estimator of β_i is consistent if $c_i(n)n^{1/2}/\nu(n) \to \infty$.

Consider now the hypothesis testing problem described in (3). As in the previous subsection consider the testing statistic given by (4), we can prove the following proposition.

Proposition 2. *Consider the test statistic* (4) *associated to the hypothesis* (3)*. If* $\sigma \in \mathcal{H}$ *then under* H_0

$$t \to_d \frac{RA^{-1}B_b}{\sqrt{C}},$$

where

$$C = \left(\int_0^1 \tau^2(W(s)) ds \right) RA^{-1}R',$$

while under H_1

$$t \to_d \infty, \qquad a.s.$$

if and only if $c_i(n)n^{1/2}/\nu(n) \to \infty$.

From the above proposition we can see that t-test is consistent if and only if the ordinary least squares estimator is. Finally the Wald test when $\sigma \in \mathcal{H}$ has exactly the same behavior as when $\sigma \in \mathcal{I}$, thus our previous discussion holds words for words and will not be repeated here.

4 Application: The Linear Trend Model As an application of our previous results consider the linear trend model given by

(5)
$$\begin{aligned} y_t &= \alpha + t\beta + \epsilon_t \qquad t = 0, 1, \dots, n. \\ \epsilon_t &= \sigma(z_t)u_t. \end{aligned}$$

In this case it is easy to see that Assumption 1 is satisfied with $g_n(i) = i$, $c(n) = n$ and $\breve{g}(r) = r$. Such model could be used in macroeconometrics where y_t is the logarithm of the gdp, β its linear trend and z_t could be, for example, the exchange rate. The asymptotic distribution of the ordinary least squares estimator when $\sigma \in \mathcal{I}$ is given by

$$D_a \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix} \rightarrow_d \begin{pmatrix} 4 & -6 \\ -6 & 12 \end{pmatrix} \begin{pmatrix} \left(\left(\int_{-\infty}^{+\infty} \sigma(s)ds\right) L(1,0)\right)^{1/2} B(1) \\ \left(\left(\int_{-\infty}^{+\infty} \sigma(s)ds\right) \left(\int_0^1 s^2 dL(s,0)\right)\right)^{1/2} B(1) \end{pmatrix},$$

where

$$D_a = \begin{pmatrix} n^{3/4} & 0 \\ 0 & n^{7/4} \end{pmatrix}.$$

When the volatility function $\sigma \in \mathcal{H}$ the asymptotic distribution of the ordinary least squares estimator is given by

$$D_b \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix} \rightarrow_d \begin{pmatrix} 4 & -6 \\ -6 & 12 \end{pmatrix} \begin{pmatrix} \int_0^1 \tau(W(s))dU(s) \\ \int_0^1 s\tau(W(s))dU(s) \end{pmatrix},$$

where

$$D_b = \frac{1}{\nu(\sqrt{n})} \begin{pmatrix} n^{1/2} & 0 \\ 0 & n^{3/2} \end{pmatrix}.$$

In comparison if there was no heteroscedasticity in the data, i.e. $\sigma(z_t) = \sigma$, then we would have (see Hamilton, 1994, pp 458)

$$D \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\beta} - \beta \end{pmatrix} \rightarrow_d N(0, \sigma \begin{pmatrix} 4 & -6 \\ -6 & 12 \end{pmatrix}), \tag{6}$$

where

$$D = \begin{pmatrix} \sqrt{n} & 0 \\ 0 & n^{3/2} \end{pmatrix}.$$

Note that this implies that in the homoscedasticity case the t-stat follows a standard normal distribution.

We now study the impact of heteroscedasticity on the size of the t-test. To do so we consider four volatility functions

$$\begin{aligned} \sigma_1(x) &= e^{-\frac{x^2}{2}} \\ \sigma_2(x) &= e^{-|x|} \\ \sigma_3(x) &= \frac{e^x}{1+e^x} \\ \sigma_4(x) &= |x|, \end{aligned}$$

the first two functions belong to $\mathcal{I}$ while the last two belong to $\mathcal{H}$. It can be checked that for σ_3 we have $\nu(\lambda) = 1$ and $\tau(s) = 1\{s \geq 0\}$, while for σ_4 we have $\nu(\lambda) = |\lambda|$ and $\tau(s) = |s|$. Furthermore since $n^{3/2}/\nu(\sqrt{n}) \rightarrow \infty$ as $n \rightarrow \infty$, the ordinary least squares estimator is consistent. In our simulations we generate the w_t and $u_{n,t}$ as iid normal with mean 0 and variance 1 and the series z_t is generated according to (2). We set the sample size $n = 1000$ and run 100000 repetitions.

To illustrate the impact of heteroscedasticity we compute the level of the test that consists in rejecting H_0 given by (3) whenever t is larger or lower than the $1 - \alpha/2$ or $\alpha/2$ quantile of the standard normal distribution. The following table shows the level of this test for our four volatility functions.

Volatility	α=0.01	α=0.05	α=0.1
σ_1	0.053	0.131	0.203
σ_2	0.005	0.031	0.071
σ_3	0.032	0.085	0 .141
σ_4	0.014	0.053	0.096

It can be seen from the table that the impact of heteroscedasticity on the size of the test depends on the functional form of the variance.

5 Conclusion This paper contributes to the litterature on nonlinear time series in two ways. Firstly, it extends some results of Phillips and Park (1999,2001) on nonlinear integrated time series. Secondly, it shows how these results can be used to derive the distribution of the ordinary least squares estimator when the regressors are deterministic and the volatility is a function of a unit root process. Future work should focus on the estimation of the volatility function σ.

Appendix

In what follows we will often use of the following lemma

Lemma 1. *Let $g_n(t)$, $t = 1, \dots, n$, and $c(n)$ satisfy Assumption 1, then*

$$sup_{r\in[0,1]}|\frac{g_n([rn])}{c(n)}| < M, \quad \forall n$$

for some $M \in \mathbb{R}$.

Proof of Lemma 1
The result follows from the boundness of $\breve{g}(r)$ and the assumption of uniform convergence. □

Define the following random variable

$$V(s,r) = \sum_{t=1}^{n} \frac{g_n(t)}{c(n)} 1_{\{s\le z_t<r\}} \qquad s < r,$$

where z_t is defined in (2) and $g_n(t)$, $c(n)$ satisfy Assumption 1. Then we have the following lemma.

Lemma 2. *If Assumptions 1, 2(b) and 2(c) are satisfied then there exists a constant C such that for any $\delta \in (0,\infty)$ and n such that $n \ge 2$, $\delta\sqrt{n} \ge 1$*
(a) $E(V(0,\delta) - \frac{1}{K}V(\delta,(K+1)\delta))^2 \le C\delta\sqrt{n}(1 + K\delta^2 log(n))$.
(b) $E(V(K\delta,(K+1)\delta) - \frac{1}{K}V(\delta,(K+1)\delta))^2 \le C\delta\sqrt{n}(1 + K\delta^2 log(n))$.
(c) $E(V(0,\delta) - V(K\delta,(K+1)\delta))^2 \le C\delta\sqrt{n}(1 + K\delta^2 log(n))$.
where $k = 1, \dots, K$ with $K \in \mathbb{N}^*$.

Proof of Lemma 2
(a) and (b) can be deduced from Akonom (1993) upon noticing that the sequence $g_n(t)/c(n)$ is bounded by Lemma 1.
(c) This follows from (a), (b) and the triangular inequality.
□

To prove Theorem 1 we will also need the following result.

Theorem 5. *Suppose Assumption 2(b) and 2(c) hold with $p > 4$. If $\sigma \in \mathcal{I}$, then*

$$\frac{1}{n^{1/2}c(n)}\sum_{t=1}^{n} g_n(t)\sigma(z_i) \to_d \left(\int_{-\infty}^{+\infty}\sigma(s)ds\right)\left(\int_0^1 \breve{g}(s)dL(s,0)\right),$$

as $n \to \infty$ for any sequence $g_n(i)$ satisfying Assumption 1 .

Proof of Theorem 5
Start by writing

$$\frac{1}{n^{1/2}}\sum_{t=1}^{n}\frac{g_n(t)}{c(n)}\sigma(z_t) = n^{1/2}\int_0^1 \frac{g_n([rn])}{c(n)}\sigma(n^{1/2}W_n(r))dr + \frac{g(n)}{c(n)}\frac{\sigma(n^{1/2}W_n(1))}{\sqrt{n}} - \frac{g(0)}{c(n)}\frac{\sigma(n^{1/2}W_n(0))}{\sqrt{n}},$$

in the right hand side of the above equality the second and third terms are $o_p(1)$, thus we just need to study the asymptotic distribution of the first term. Define then

$$\kappa_n = n^a \text{ and } \delta_n = n^{-b}, \tag{7}$$

with $a, b > 0$ and

$$\begin{aligned}
\sigma_n(x) &= \sigma(x)1_{\{-\kappa_n\delta_n \le x < \kappa_n\delta_n\}}\\
\sigma'_n(x) &= \sigma(x)1_{\{\kappa_n\delta_n \le x\}}\\
\sigma''_n(x) &= \sigma(x)1_{\{x < -\kappa_n\delta_n\}}.
\end{aligned}$$

As in Phillips and Park (1999) we will assume that the following inequalities hold

$$a - (1+l)b < 0 \tag{8}$$
$$(6b-1)p + 2 < 0 \tag{9}$$
$$2a - 1 < 0 \tag{10}$$
$$4a - 4b - 1 < 0 \tag{11}$$
$$(a-b)p - 1 > 0, \tag{12}$$

it can be checked that under our assumptions the above system defines a non empty set when $p > 4$. The theorem will be proved if we show that

$$\begin{aligned}
n^{1/2}\int_0^1 \frac{g_n([rn])}{c(n)}\sigma_n(n^{1/2}W_n(r))dr &= \left(\int_{-\infty}^{+\infty}\sigma(s)ds\right)\left(\int_0^1 \breve{g}(s)dL(s,0)\right) + o_p(1)\\
n^{1/2}\int_0^1 \frac{g_n([rn])}{c(n)}\sigma'_n(n^{1/2}W_n(r))dr &= o_p(1)\\
n^{1/2}\int_0^1 \frac{g_n([rn])}{c(n)}\sigma''_n(n^{1/2}W_n(r))dr &= o_p(1).
\end{aligned}$$

Define

$$\sigma_{\delta_n}(x) = \sum_{k=-\kappa_n}^{\kappa_n-1}\sigma(k\delta_n)1_{\{k\delta_n \le x < (k+1)\delta_n\}},$$

then the Lipschitz condition on σ implies that sup$|\sigma_n(x) - \sigma_{\delta_n}(x)| \le c\delta_n^l$. This with Lemma 1, (7) and (8) leads to

$$|n^{1/2}\int_0^1 \frac{g_n([rn])}{c(n)}\sigma_n(n^{1/2}W_n(r))dr - n^{1/2}\int_0^1 \frac{g_n([rn])}{c(n)}\sigma_{\delta_n}(n^{1/2}W_n(r))dr| = o_p(1),$$

Now

$$n^{1/2}\int_0^1 \frac{g_n([rn])}{c(n)}\sigma_{\delta_n}(n^{1/2}W_n(r))dr = A_n + B_n, \tag{13}$$

where

$$\begin{aligned} A_n &= n^{1/2}\sum_{k=-\kappa_n}^{\kappa_n-1}\frac{1}{n}\left(\sum_{t=1}^{n}\frac{g_n(t)}{c(n)}\sigma(k\delta_n)1_{\{0\le n^{1/2}W_n(\frac{t}{n})<\delta_n\}}\right)\\ B_n &= n^{1/2}\sum_{k=-\kappa_n}^{\kappa_n-1}\frac{1}{n}\left(\sum_{t=1}^{n}\frac{g_n(t)}{c(n)}\sigma(k\delta_n)(1_{\{k\delta_n\le n^{1/2}W_n(\frac{t}{n})<(k+1)\delta_n\}} - 1_{\{0\le n^{1/2}W_n(\frac{t}{n})<\delta_n\}})\right). \end{aligned}$$

Using the Cauchy-Schwartz inequality we can write

$$\begin{aligned} E(B_n^2) &\le n\left(\sum_{k=-\kappa_n}^{\kappa_n-1}\sigma^2(k\delta_n)\right)\sum_{k=-\kappa_n}^{\kappa_n-1}E\left(\frac{1}{n}V(k,(k+1)\delta_n)-\frac{1}{n}V(0,\delta_n)\right)^2\\ &= \frac{n}{\delta_n}\left(\int_{-\infty}^{+\infty}\sigma^2(s)ds+o(1)\right)\sum_{k=-\kappa_n}^{\kappa_n-1}E\left(\frac{1}{n}V(k,(k+1)\delta_n)-\frac{1}{n}V(0,\delta_n)\right)^2\\ &= o(1), \end{aligned}$$

where we used Lemma 2, (7), (10) and (11). This implies that $E(B_n^2) = o(1)$, thus to determine the asymptotic behavior of (13) we only need to study A_n.

A_n can be rewritten as

$$A_n = \frac{n^{1/2}}{\delta_n}\left(\int_{-\infty}^{+\infty}\sigma(r)dr+o(1)\right)\left(\int_0^1\frac{g_n([rn])}{c(n)}1_{\{0\le n^{1/2}W_n(r)<\delta_n\}}dr\right).$$

Define $s_m(r) = c_0 1_{[0.a_1)} + \sum_{i=1}^{h-1} c_i 1_{[a_i,a_{i+1})} + c_h 1_{[a_{h-1},1]}$ with $a_i \in (0,1)$ and $\sup_{r\in[0,1]}|\breve{g}(r)-s_m(r)| < 1/m$ where $m\in\mathbb{N}$. The triangular inequality implies that $\lim_{n\to\infty}\sup_{r\in[0,1]}|g_n([rn])/c(n) - s_m(r)| < 1/m$. Consider now

$$\begin{aligned} C_1(n) &= \frac{n^{1/2}}{\delta_n}\int_0^1\frac{g_n([rn])}{c(n)}1_{\{0\le n^{1/2}W_n(r)<\delta_n\}}dr\\ C_{2,m}(n) &= \frac{n^{1/2}}{\delta_n}\int_0^1 s_m(r)1_{\{0\le n^{1/2}W_n(r)<\delta_n\}}dr\\ C_3 &= \int_0^1\breve{g}(r)dL(r,0). \end{aligned}$$

Since $L(r,0)$ is almost surely monotone and $\breve{g}(r)$ piecewise continuous C_3 is well defined with probability one. The triangular inequality implies that

$$\{\omega : |C_1(n)-C_{2,m}(n)| > \frac{\eta}{2}\}\cup\{\omega : |C_{2,m}(n)-C_3| > \frac{\eta}{2}\} \supseteq \{\omega : |C_1(n)-C_3| > \eta\}. \tag{14}$$

We will now show that the probability of the left hand size of the above expression goes to zero.

Firstly, for a n large enough we can write

$$|C_1(n)-C_{2,m}(n)| \le \frac{1}{m}\frac{n^{1/2}}{\delta_n}\int_0^1 1_{\{0\le n^{1/2}W_n(r)<\delta_n\}}dr.$$

The left-hand side of the above inequality can be rewritten as

$$\frac{1}{m}(L(1,0)+o_p(1)), \tag{15}$$

where $o_p(1)$ and thus goes to zero as $n\to\infty$ by Lemma 2.5(b) of Phillips and Park (1999) and (9). Thus another application of the triangular inequality yields

$$P\left(\{\omega:\frac{1}{m}L(1,0)>\frac{\eta}{4}\}\right)+P\left(\{\omega:\frac{1}{m}o_p(1)>\frac{\eta}{4}\}\right)\geq P\left(\{\omega:|C_1(n)-C_{2,m}(n)|>\frac{\eta}{2}\}\right),$$

taking first the limit for $n\to\infty$ and then for $m\to\infty$ we obtain that

$$\lim_{m\to\infty}\lim_{n\to\infty}P\left(\{\omega:|C_1(n)-C_{2,m}(n)|>\frac{\eta}{2}\}\right)=0.$$

Secondly, we have the following bound

$$\begin{aligned}|C_{2,m}(n)-C_3| &\leq |\int_0^1 s_m(r)dL(r,0)-\int_0^1\breve{g}(r)dL(r,0)|+|o_p(1)| \\ &\leq \frac{1}{m}L(1,0)+o_p(1),\end{aligned} \tag{16}$$

where $o_p(1)$ goes to zero in probability if $n\to\infty$. In (16) we use Theorem 4 of Akanom (1993) and (9) to write

$$C_{2,m}(n)=\int_0^1 s_m(r)dL(r,0)+o_p(1),$$

with $o_p(1)$ goes to zero as $n\to\infty$. An argument similar to the one above leads to

$$\lim_{m\to\infty}\lim_{n\to\infty}P\left(\{\omega:|C_{2,m}(n)-C_3|>\frac{\eta}{2}\}\right)=0.$$

Using our previous results we can write

$$\begin{aligned}0= \quad \lim_{m\to\infty}\lim_{n\to\infty} \quad &\left(P\left(\{\omega:|C_1(n)-C_{2,m}(n)|>\frac{\eta}{2}\}\right)+P\left(\{\omega:|C_{2,m}(n)-C_3|>\frac{\eta}{2}\}\right)\right) \\ &\geq \lim_{n\to\infty}P\left(\{\omega:|C_1(n)-C_3|>\eta\}\right),\end{aligned}$$

for every positive η. Thus we have just proved that

$$A_n\to_p\left(\int_{-\infty}^{+\infty}\sigma(r)dr\right)\left(\int_0^1\breve{g}(r)dL(r,0)\right).$$

Finally to complete the proof note that as in Phillips and Park we can alway assume that σ_n' and σ_n'' are monotone decreasing. It then follows that

$$\begin{aligned}|n^{1/2}\int_0^1\frac{g_n(\lfloor rn\rfloor)}{c(n)}\sigma_n'(n^{1/2}W_n(r))dr| &\leq Kn^{1/2}\int_0^1|\sigma_n'(n^{1/2}W_n(r))|dr=o_p(1) \\ |n^{1/2}\int_0^1\frac{g_n(\lfloor rn\rfloor)}{c(n)}\sigma_n''(n^{1/2}W_n(r))dr| &\leq Kn^{1/2}\int_0^1|\sigma_n''(n^{1/2}W_n(r))|dr=o_p(1),\end{aligned}$$

where for the first inequality we used Lemma 1 and the $o_p(1)$ equality follows from Park and Phillips (1999). The argument for negative σ_n' and σ_n'' can be deduced in a similar way be noticing that $-\sigma_n'$ and $-\sigma_n''$ are positive. □

Lemma 3. *Consider $g_n(t)$, $t = 1, \ldots, n$, and $c(n)$ satisfy the Assumptions of Lemma 1 then*

$$sup_{r\in[0,1]}|\frac{g_n^2([rn])}{c^2(n)} - \breve{g}^2(r)| \to 0.$$

Proof of Lemma 3
Simply notice that for n large enough we can write

$$\begin{aligned}
& \sup_{r\in[0,1]}|\frac{g_n^2([rn])}{c^2(n)} - \breve{g}^2(r)| \\
= \; & \sup_{r\in[0,1]}|\left(\frac{g_n([rn])}{c(n)} - \breve{g}(r)\right)\left(\frac{g_n([rn])}{c(n)} + \breve{g}(r)\right)| \\
\leq \; & \sup_{r\in[0,1]}|\frac{g_n([rn])}{c(n)} - \breve{g}(r)|\sup_{r\in[0,1]}|\frac{g_n([rn])}{c(n)} + \breve{g}(r)| \\
\leq \; & \sup_{r\in[0,1]}|\frac{g_n([rn])}{c(n)} - \breve{g}(r)|M,
\end{aligned}$$

for some constant M. The last inequality is due to the assumption of uniform convergence. The lemma follows by letting $n \to \infty$. □

Proof of Theorem 1
Define the following martingale

$$\begin{aligned}
M_n(r) &= n^{1/4}\sum_{t=1}^{k-1}\frac{g_n(t)}{c(n)}\sigma(\sqrt{n}W_n(\frac{t}{n}))\left(U(\frac{\tau_{n,t}}{n}) - U(\frac{\tau_{n,t-1}}{n})\right) \\
&+ n^{1/4}\frac{g_n(k)}{c(n)}\sigma(\sqrt{n}W_n(\frac{k}{n}))\left(U(r) - U(\frac{\tau_{n,k-1}}{n})\right),
\end{aligned}$$

where $\tau_{n,k-1}/n < r \leq \tau_{n,k}/n$. The $\tau_{n,k}$ are stopping the times defined in Park and Phillips (2001).

Consider now the two following martingales

$$\tilde{M}_n(r) = \begin{cases} M_n(r) & r \in [0, \frac{\tau_{n,n}}{n}] \\ M_n(r) + B_1\left(r - \frac{\tau_{n,n}}{n}\right) & r > \frac{\tau_{n,n}}{n}. \end{cases}$$

and

$$\tilde{W}(r) = \begin{cases} W(r) & r \in [0,1] \\ W(1) + B_2(r-1) & r > 1. \end{cases}$$

where $B_1(r)$ and $B_2(r)$ are two independent Brownian motions.

The quadratic variation of $\tilde{M}_n$ is given by

$$[\tilde{M}_n]_r = n^{1/4}\int_0^{\min(r,\frac{\tau_{n,n}}{n})}\left(\frac{g_n([sn])}{c(n)}\right)^2\sigma^2(\sqrt{n}W_n(s))ds(1+o_p(1)) + \left(r - \frac{\tau_{n,n}}{n}\right)1_{\{r>\frac{\tau_{n,n}}{n}\}},$$

therefore from our previous theorem and Lemma 3 we have that

$$[\tilde{M}_n]_r \to_d \left(\int_{-\infty}^{+\infty}\sigma^2(s)ds\right)\left(\int_0^{\min(r,1)}\breve{g}^2(s)dL(s,0)\right) + (r-1)1_{\{r>1\}}.$$

Here we also used a result from Park and Phillips (2001) that if $\sigma \in \mathcal{I}$ then $\sigma^2 \in \mathcal{I}$.

The independence of B_1, B_2, Lemma 1 and computations similar to those in Phillips and Park (2001) yield

$$[\tilde{M}_n, \tilde{W}]_r \to_p 0, \tag{17}$$

$r \in [0,\infty)$. Define now $\rho_n(r) = \inf\{s \in [0,\infty) : [\tilde{M}_n]_s > r\}$, note that since $[\tilde{M}_n]_s \to \infty$ as $s \to \infty$, $\rho_n(r)$ is well defined for $r \in [0,\infty)$. Define now the Dambis, Dubins-Schwarz Brownian motion

$$B_n(r) = \tilde{M}_n(\rho_n(r)),$$

this is a well defined Brownian motion over $[0,\infty)$. It then follows that $(B_n(r), \tilde{W}(r))$ converges to $(B(r), \tilde{W}(r))$, where $B(r)$ and $\tilde{W}(r)$ are two independent Brownian motions by (17) (see Revuz and Yor (1994), Theorem 1.6, page 173). Therefore

$$\tilde{M}_n\left(\frac{\tau_{n,n}}{n}\right) \to_d \left(\left(\int_{-\infty}^{\infty} \sigma^2(s)ds\right)\int_0^1 \breve{g}^2(s)dL(s,0)\right)^{1/2} B(1),$$

where B is a brownian motion independent of W. □

Proof of Proposition 1
Note that t can be rewritten as

$$\frac{n^{3/4}c_i(n)(\hat{\beta}_i - \beta_{0,i})}{\sqrt{\hat{\sigma}_\epsilon^2 R D_a (\sum_{t=1}^n G_n(t)G_n(t)')^{-1} D_a R'}}.$$

Under H_0 it holds that

$$n^{3/4}c_i(n)(\hat{\beta}_i - \beta_{0,i}) \to_d RA^{-1}B_a.$$

From Chung and Park (2007) we have

$$n^{1/2}\hat{\sigma}_\epsilon^2 \to_d L(1,0)\left(\int_{-\infty}^{\infty} \sigma^2(s)ds\right).$$

Finally it is straightforward to see that

$$\sqrt{n}D_a^{-1}\left(\sum_{t=1}^n G_n(t)G_n(t)'\right)D_a^{-1} \to A.$$

The distribution under H_1 can be derived in the same way. Note that under H_1 the numerator can be written as

$$\sqrt{n}c_i(n)(\hat{\beta}_i - \beta_i) + \sqrt{n}c_i(n)(\beta_i - \hat{\beta}_{i,0}),$$

the above expression almost surely diverges to infinity if and only if $\sqrt{n}c_i(n) \to \infty$. □

Before proving Theorem 3 we first need to extend Lemma 2 of Phillips and Park (1999).

Lemma 4. *Let τ be a locally Riemann integrable function. If Assumptions 2(a)(d) and (e) holds then*

$$\frac{1}{\sqrt{n}}\sum_{t=1}^n g_n(t)\tau\left(\frac{z_t}{\sqrt{n}}\right)u_{n,t} \to_d \int_0^1 \breve{g}(s)\tau(V(s))dU(s).$$

Proof of Lemma 4
Since the proof of this lemma is very similar to the one of Lemma 2 of Phillips and Park we will just highlight the necessary modifications. Fix a compact set $K = [-c, c]$ with c defined in Phillips and Park (1999). Then, we can find two continuous functions $\overline{\tau}_\epsilon$ and $\underline{\tau}_\epsilon$ such that $\overline{\tau}_\epsilon \leq \overline{\tau} \leq \underline{\tau}_\epsilon$ on K. Since $\underline{\tau}_\epsilon$ is continuous then by the Skorokhod (see Billingsley (1968)) embeding theorem we can assume that $\underline{\tau}_\epsilon(V_n(r)) \to_p \underline{\tau}_\epsilon(V(r))$. It then follows from theorem 4.6 of Kurtz and Protter (1991) or Theorem 2.1 of Hansen (1992) that

$$\int_0^1 \frac{g_n([rn])}{c(n)} \underline{\tau}_\epsilon(V_n(r))dU_n(r) \to_d \int_0^1 \breve{g}(r)\underline{\tau}_\epsilon(V(r))dU(r).$$

As in Phillips and Park (1999) to prove the lemma it is enough to show that

$$|\int_0^1 \frac{g_n([rn])}{c(n)} \tau(V_n(r))dU_n(r) - \int_0^1 \breve{g}(r)\underline{\tau}_\epsilon(V_n(r))dU_n(r)| \to_p 0. \tag{18}$$

as $\epsilon \to 0$. This however follows from Phillips and Park (2001) and the boundess of the array $g_n(t)/c(n)$ □

Proof of Theorem 4
Simply note that

$$\frac{1}{\sqrt{n}\nu(\sqrt{n})} \sum_{t=1}^{n} \frac{g_n(t)}{c(n)} \sigma(z_t)u_{n,t} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{g_n(t)}{c(n)} \tau(z_t)u_{n,t} + o(1)$$

where the equality is due to the boundeness of the array $g_n(t)/c(n)$ and Lemma A5(b) of Phillips and Park (2001). □

Proof of Proposition 2
¿From equation (12) of Chung and Park (2007) we have that

$$\frac{1}{\nu^2(\sqrt{n})n} \sum_{t=1}^{n} \epsilon_{n,t}^2 \to_d \int_0^1 \tau^2(V(r))dr.$$

The rest of the proof is similar to the proof of Proposition 1 and thus will not be repeated here. □

Acknowledgement The author would like to thank Professor Masanobu Taniguchi and an anonymous referee for their helpful comments and suggestions.

REFERENCES

[1] Akonom, J. (1993): Comportement asymptotique du tems d occupation du processus des sommes partielles, *Annales de l I H P*, section B, tome 29, n1, p 57-81.

[2] Chang Y. and Park J. (2003): Index Models with Integrated Time Series, *J. Econometrics*, vol 114, 73-103.

[3] Chang Y., Park J. and Phillips (2001): Nonlinear Econometric models with cointegrated and deterministically trending regressors, *Econometrics J.*, vol 4, 1-36.

[4] Chung H. and Park J. (2007): Nonstationary Nonlinear Heteroskedasticity in Regression, *J. Econometrics*, vol 137, 230-259.

[5] Hamilton J.D. (1999): Time Series Analysis, Princeton University Press.

[6] Hansen B. E. (1995): Regression with Nonstationary Volatility, *Econometrica*, vol 63, 1113-1132.

[7] Hansen B. E. (1996): Stochastic Equicontinuity for Unbounded Dependent Heterogeneous Arrays, *Econometric Theory*, vol 12, 347-359.

[8] Hu, L. and Phillips, P.C.B (2004): Nonstationary Discrete Choice, *J. Econometrics*, vol. 141, 103-138.

[9] Jong, R. and Wang, C. (2005): Further Results on the Asymptotics for Nonlinear Transformation of Integrated Time Series, *Econometric Theory*, vol 21, 413-430.

[10] Meyer-Brandis, T. and Tankov P. (2008): Multi-factor jump-diffusion models of electricity prices, *Int. J. Theoretical Appl. Finance* , vol 11, 503 - 528.

[11] Park, J. (2002): Nonstationary Nonlinear Heteroskedasticity, *J. Econometrics*, vol 110, 383-415.

[12] Park, J. and Phillips, P.C.B (1999): Asymptotics for Nonlinear Transformations of Time Series, *Econometric Theory*, vol 15, 269-298.

[13] Park, J. and Phillips, C.B.P. (2001): Nonlinear Regression with Integrated Time Series, *Econometrica*, vol 69, 117-161.

[14] Phillips, P.C.B. (1986): Understanding Spurious Regressions in Econometrics, *J. Econometrics*, vol 33, 311-340.

[15] Phillips, P.C.B. (1987): Time Series Regression with a Unit Root, *Econometrica*, vol 55, 277-301.

[16] Phillips, P.C.B. and S. N. Durlauf (1986): Multiple Time Series with Integrated Variables, *Review of Economic Studies*, vol 53, 473-496.

[17] Shi X. and Phillips P. C. B. (2010): Nonlinear Cointegrating Regression Under Weak Identification, Cowles Foundation Working Paper.

CENTER FOR ENGLISH LANGUAGE EDUCATION IN SCIENCE AND ENGINEERING, FACULTY OF SCIENCE AND ENGINEERING, WESEDA UNIVERSITY, 3-4-1, OKUBO SHINJUKU, TOKYO (169-8555)

E-mai : apetkovi@aoni.waseda.jp