Econometric Theory, **30**, 2014, 509–535.
doi:10.1017/S026646661300039X

MARTINGALE LIMIT THEOREM REVISITED AND NONLINEAR COINTEGRATING REGRESSION

QIYING WANG
University of Sydney

For a certain class of martingales, convergence to a mixture of normal distributions is established under convergence in distribution for the conditional variance. This is less restrictive in comparison with the classical martingale limit theorem, where one generally requires convergence in probability. The extension partially removes a barrier in the applications of the classical martingale limit theorem to nonparametric estimation and inference with nonstationarity and enhances the effectiveness of the classical martingale limit theorem as one of the main tools to investigate asymptotics in statistics, econometrics, and other fields. The main result is applied to investigate limit behavior of the conventional kernel estimator in a nonlinear cointegrating regression model, which improves existing works in the literature.

1. INTRODUCTION

Let $\{M_{n,i}, \mathcal{F}_{n,i}, 1 \le i \le k_n\}$ be a zero-mean square integrable martingale array, having difference $y_{n,i}$ and nested σ-fields structure, that is, $\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$ for $1 \le i \le k_n, n \ge 1$. Suppose that, as $n \to \infty$, $k_n \to \infty$,

$$\sum_{i=1}^{k_n} \mathbb{E}\left[y_{n,i}^2 I\left(|y_{n,i}| \ge \epsilon\right) \mid \mathcal{F}_{n,i-1}\right] \to_P 0, \tag{1.1}$$

for all $\epsilon > 0$, and the conditional variance

$$\sum_{i=1}^{k_n} \mathbb{E}\left[y_{n,i}^2 \mid \mathcal{F}_{n,i-1}\right] \to_P M^2, \tag{1.2}$$

where M^2 is an almost surely (a.s.) finite random variable. The classical martingale limit theorem (MLT) shows that $M_{n,k_n} = \sum_{i=1}^{k_n} y_{n,i} \to_D Z$, where the random variable (r.v.) Z has characteristic function $\mathbb{E}e^{itZ} = \mathbb{E}e^{-M^2t^2/2}$, $t \in \mathbf{R}$. If M^2 is a constant, the nested structure of the σ-fields $\mathcal{F}_{n,i}$ is not necessary. See, e.g., Chapter 3 of Hall and Heyde (1980).

The author thanks Prof. Phillips, Prof. Robinson, and four referees for their very helpful comments on the original version. The author also acknowledges research support from the Australian Research Council. Address correspondence to Qiying Wang, School of Mathematics and Statistics, The University of Sydney, NSW 2006, Australia; e-mail: qiying@maths.usyd.edu.au.

© Cambridge University Press 2013

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

As one of the main conventional tools, the classical MLT is widely used in statistics, econometrics, and other fields. In many applications, however, the convergence in probability in (1.2) seems to be too restrictive. To illustrate, consider a functional of the form

$$M_{n,i} = \left(nh^2\right)^{-1/4} \sum_{k=1}^{i} g(x_k/h)\epsilon_{k+1}, \qquad i = 1, 2, ..., n, \tag{1.3}$$

where g is a real integrable function on $\mathbf{R}$, ϵ_k is a stationary process, x_t is an integrated process (i.e., $I(1)$ process) such as a random walk, and $h \equiv h_n$ is a certain sequence of positive constants satisfying $h \to 0$ and $nh^2 \to \infty$. Write $\mathcal{F}_k = \sigma\left(x_1, ..., x_k; \epsilon_1, ..., \epsilon_k\right)$. If $\mathbb{E}\left(\epsilon_{k+1} \mid \mathcal{F}_k\right) = 0$ and $\mathbb{E}\left(\epsilon_{k+1}^2 \mid \mathcal{F}_k\right) = 1$, $\{M_{n,i}, \mathcal{F}_i\}_{i=1}^n$ forms a zero-mean martingale array with conditional variance $V_n^2 = \frac{1}{\sqrt{nh}} \sum_{k=1}^n g^2(x_k/h)$. Unfortunately, the asymptotics of $M_{n,n}$ cannot be obtained by the classical MLT. Indeed, under certain conditions on x_t such that $x_{[ns]}/\sqrt{n} \Rightarrow W(s)$ on $D[0,1]$, where $W(s)$ is a standard Brownian motion, we may prove

$$V_n^2 \to_D \int_{-\infty}^{\infty} g^2(x)dx\, L_W(1,0), \tag{1.4}$$

where $L_W(t,s)$ is a local time process defined as in the next section, but it is difficult or impossible to replace the convergence in distribution in (1.4) by convergence in probability. See, e.g., Wang and Phillips (2009a) for instance. It should be pointed out that, given (1.4), one may enlarge the probability space in which $x_1, ..., x_t$ are equipped so that, in this new space, an equivalent process V_n^{*2} to V_n^2 (that is, $V_n^{*2} =_D V_n^2$ for each $n \geq 1$) exists and

$$V_n^{*2} \to_P \int_{-\infty}^{\infty} g^2(x)dx\, L_W(1,0). \tag{1.5}$$

This fact cannot be used to establish the asymptotics of $M_{n,n}$ as well since, without independence between ϵ_t and x_t, this enlarged probability space may destroy the nested σ-fields structure (i.e., $\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$) required in the classical MLT. See Remark 4 in Section 2 for more details.

Such functionals $M_{n,n}$ having a martingale structure commonly arise in nonlinear regression with integrated time series (Park and Phillips, 1999, 2001), nonparametric estimation in relation to nonlinear cointegration models (Phillips and Park, 1998; Karlsen and Tjøstheim, 2001; Karlsen, Myklebust, and, Tjøstheim, 2007; and Wang and Phillips, 2009a, 2009b, 2011) and many other related research areas (Marmer, 2008; Gao, King, Lu, and Tjøstheim, 2009a, 2009b; Choi and Saikkonen, 2010; Kasparis, 2008, 2011; Kasparis and Phillips, 2009; Wang and Phillips, 2012). In such cases, g may be a kernel function K or a squared kernel function K^2, and the sequence h is the bandwidth used in nonparametric regression. As illustrated above, the classical MLT is limited to establishing the asymptotics of $M_{n,n}$, which, as is well-known, plays a key role in the development of nonparametric estimation and inference with nonstationarity.

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

The main aim of this paper is to provide an extension of the classical MLT. It is shown that, for a certain class of martingales, the condition (1.2), that is, convergence in probability for the conditional variance in the classical MLT, can be reduced to the less restrictive

$$\sum_{i=1}^{k_n} \mathbb{E}[y_{n,i}^2 \mid \mathcal{F}_{n,i-1}] \to_D M^2. \tag{1.6}$$

As discussed above, this kind of extension removes a major barrier in the application of the classical MLT to nonparametric estimation and inference with nonstationarity. It is expected that our results will enhance the effectiveness of the classical MLT as one of the main tools in the investigation of asymptotics in statistics, econometrics, and other fields. We finally remark that there are few results in the literature concerning the establishment of a central limit theorem for martingales under the less restrictive condition (1.6). Theorem 3.4 of Hall and Heyde (1980) and an unpublished manuscript by Jeganathan (2006) provided a partial result in this direction. Park and Phillips (2001) considered a class of martingales having the shape (1.3). Unlike these existing works, for the martingales defined as in (2.1) below, our conditions in investigating the asymptotics are neat, and the verification of the conditions is quite straightforward.

This paper is organized as follows: In the next section, we present our main results. Theorem 2.1 provides a framework in an extension to the classical MLT. It is shown that, for a certain class of martingales, a minor and easily verified additional condition, together with convergence in distribution for the conditional variance, is sufficient to establish the convergence to a mixture of normal distributions. Generally speaking, it is impossible to replace convergence in probability for the conditional variance by convergence in distribution without any additional conditions. This kind of additional condition might be close to necessary to investigate the asymptotics for the specified martingales defined as in (2.1). Theorem 2.2 gives an important application of Theorem 2.1 to general linear processes, where we investigate the asymptotics for the functionals having the shape (1.3). Our result includes the x_k being a partial sum of the auto regressive moving average (ARMA) process and fractionally integrated processes, which are most commonly used in practice. Using Theorem 2.2 as a main tool, Theorems 2.3 and 2.4 investigate the asymptotics for the conventional kernel estimators in a nonlinear cointegrating regression model. These results improve those in the existing literature. This section also presents several remarks for possible further extensions of our results. All technical proofs are given in Section 3.

Throughout the paper, we denote by $C, C_1, \ldots$ constants, which may change at each appearance. The notation $\to_D$ ($\to_P$, respectively) denotes convergence in distribution (in probability, respectively) for a sequence of random variables (or vectors). The notation $=_D$ denotes equality in distribution. If $\alpha_n^{(1)}$, $\alpha_n^{(2)}$,..., $\alpha_n^{(k)}$ $(1 \le n \le \infty)$ are random elements on $D[0,1]$ or $D[0,\infty)$ or $\mathbf{R}$, we will understand the condition

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

$$\left(\alpha_n^{(1)}, \alpha_n^{(2)}, ..., \alpha_n^{(k)}\right) \Rightarrow \left(\alpha_\infty^{(1)}, \alpha_\infty^{(2)}, ..., \alpha_\infty^{(k)}\right)$$

to mean that for all $\alpha_\infty^{(1)}$, $\alpha_\infty^{(2)}$,..., $\alpha_\infty^{(k)}$-continuity sets A_1, A_2,...,A_k,

$$P\left(\alpha_n^{(1)} \in A_1, \alpha_n^{(2)} \in A_2, ..., \alpha_n^{(k)} \in A_k\right) \to P\left(\alpha_\infty^{(1)} \in A_1, \alpha_\infty^{(2)} \in A_2, ..., \alpha_\infty^{(k)} \in A_k\right).$$

(see Billingsley, 1968, Thm. 3.1, or Hall, 1977). As usual, $D[0,1]$ ($D[0,\infty)$, respectively) denotes the space of cadlag functions on $[0,1]$ ($[0,\infty)$, respectively), which is equipped with Skorohod topology.

2. MAIN RESULTS AND APPLICATIONS

2.1. Main Results

Motivated by nonparametric estimation and inference with nonstationarity, this section considers a class of statistics defined by

$$S_n = \sum_{k=1}^{n} x_{n,k}\epsilon_{n,k+1}, \tag{2.1}$$

where $\left\{\epsilon_{n,k}, \mathcal{F}_{n,k}\right\}_{1\le k\le n}$ forms a martingale difference and $x_{n,k}$ is adapted to $\mathcal{F}_{n,k}$ for $1 \le k \le n, n \ge 1$. Let $f_n(...)$ be a real function of its components, and $\eta_{n,k}$ and ξ_{nj} be two sequences of random variables. We specify the $x_{n,k}$ to have a form

$$x_{n,k} = f_n\left(\epsilon_{n,1}, ..., \epsilon_{n,k};\ \eta_{n,1}, ..., \eta_{n,k};\ \xi_{n1}, \xi_{n2}, ...\right),$$

and, for each $n \ge 1$, the $\mathcal{F}_{n,k}$ to be a sequence of increasing σ-fields, that is,

$$\cdots \subseteq \mathcal{F}_{n,k-1} \subseteq \mathcal{F}_{n,k} \subseteq \mathcal{F}_{n,k+1} \subseteq \cdots.$$

We do not require the nested structure of $\mathcal{F}_{n,k}$ (i.e., $\mathcal{F}_{n,k} \subseteq \mathcal{F}_{n+1,k}$) in this paper, and this kind of $x_{n,k}$ and $\mathcal{F}_{n,k}$ is quite natural in many applications.

We make use of the following assumptions in developing the asymptotics of S_n.

Assumption 1. Let $\left\{(\eta_{n,k}, \epsilon_{n,k}), \mathcal{F}_{n,k}\right\}_{1\le k\le n}$ form a martingale difference, satisfying (i) as $n \to \infty$,

$$\max_{1\le k\le n} |\mathbb{E}\left(\eta_{n,k+1}^2 \mid \mathcal{F}_{n,k}\right) - 1| \to 0, \quad \max_{1\le k\le n} |\mathbb{E}\left(\epsilon_{n,k+1}^2 \mid \mathcal{F}_{n,k}\right) - 1| \to 0, \qquad \text{a.s.},$$

(ii) for any constant sequence $0 < A_n \to \infty$,

$$\max_{1\le k\le n} \Big[\mathbb{E}(\eta_{n,k+1}^2 I\left(|\eta_{n,k+1}| \ge A_n\right) \mid \mathcal{F}_{n,k}) + \mathbb{E}\left(\epsilon_{n,k+1}^2 I\left(|\epsilon_{n,k+1}| \ge A_n\right) \mid \mathcal{F}_{n,k}\right)\Big] = o_P(1). \tag{2.2}$$

Assumption 2. Let $\max_{1\le k\le n} |x_{n,k}| = o_P(1)$ and

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} |x_{n,k}|\, |\mathbb{E}\left(\eta_{n,k+1}\epsilon_{n,k+1} \mid \mathcal{F}_{n,k}\right)| = o_P(1). \tag{2.3}$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

In the following, write $G_n^2 = \sum_{k=1}^n x_{n,k}^2$, $W_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{[nt]} \eta_{n,j+1}$ and $\Xi_n(t) = \frac{1}{d_n} \sum_{j=1}^{[nt]} \xi_{nj}$ for some $0 < d_n \to \infty$. Further, let $W(t)$ be a standard Wiener process and $\Xi(s)$ be a continuous random process that is independent of $W(t)$.

Assumption 3. There exists an a.s. finite functional $g_1^2(W)$ of $W(s), 0 \le s \le 1$, such that

$$\left\{ W_n(t),\ G_n^2 \right\} \Rightarrow \left\{ W(t), g_1^2(W) \right\}. \tag{2.4}$$

Assumption 4. (i) ξ_{nj}, $j \ge 1$, are $\mathcal{F}_{n,1}$-measurable for each $n \ge 1$. (ii) There exists an a.s. finite functional $g^2(W, \Xi)$ of $W(s), 0 \le s \le 1$ and $\Xi(s), 0 \le s < \infty$, such that

$$\left\{ W_n(t),\ \Xi_n(t),\ G_n^2 \right\} \Rightarrow \left\{ W(t),\ \Xi(t),\ g^2(W, \Xi) \right\}. \tag{2.5}$$

The following is our main result.

THEOREM 2.1. *Under Assumptions 1–3, we have*

$$\left\{ S_n,\ G_n^2 \right\} \to_D \left\{ g_1(W) N,\ g_1^2(W) \right\}, \tag{2.6}$$

where N is a standard normal variate independent of $g_1(W)$. If Assumptions 1–2 and 4 hold, then

$$\left\{ S_n,\ G_n^2 \right\} \to_D \left\{ g(W, \Xi) N,\ g^2(W, \Xi) \right\}, \tag{2.7}$$

where N is a standard normal variate independent of $g(W, \Xi)$. Consequently, in both situations, we have

$$S_n / G_n \to_D N(0, 1). \tag{2.8}$$

Remark 1. The result (2.6) states that $x_{n,k}$ may take any sequence of random variables $\xi_{n,j}, 1 \le j \le k$, as its components, if $\xi_{n,j} \in \mathcal{F}_{n,k}$ and $\xi_{n,j}$ does not have a contribution to the limit distribution of $\sum_{k=1}^n x_{n,k}^2$. In many applications, we may have $x_{n,k} = f_n(M_{n,k})$ with $M_{n,k} = \sum_{j=1}^k \eta_{n,j} + \xi_{n,k}$, where $\eta_{n,k}$ is a martingale difference and $\xi_{n,k}$ is $\mathcal{F}_{n,k}$-measurable and negligible (that is, $M_{n,k}$ is not a martingale, but may be approached by a martingale. As for the martingale approximation, we refer to Borovskikh and Korolyuk (1997), Wu and Woodroofe (2004), and the references therein). Theorem 2.1 is applicable in this situation only if the $\eta_{n,k}$ and $x_{n,k}$ satisfy the conditions required.

Remark 2. If $G_n^2 \to_P g^2(W, \Xi)$, the asymptotics of S_n may be investigated by the classical MLT. See Hall and Heyde (1980), Jacod and Shiryaev (2003), and Prigent (2003), for instance. Unfortunately, except for the situation where $g^2(W, \Xi)$ is a constant, it is generally difficult to prove $G_n^2 \to_P g^2(W, \Xi)$ in many applications such as nonparametric estimation and inference in nonlinear

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

cointegrating regression. Our condition on G_n^2 is less restrictive and accessible, and therefore our results provide a useful extension to the classical MLT.

Remark 3. If $\eta_{n,k}$ is independent of $\epsilon_{n,k}$, (2.3) holds true automatically. A simple sufficient condition for (2.3) is $\frac{1}{\sqrt{n}}\sum_{k=1}^n |x_{n,k}| = o_P(1)$ and

$$\max_{\substack{1\le k\le n\\ n\ge 1}} \left[\mathbb{E}\left(|\eta_{n,k+1}|^2 \mid \mathcal{F}_{n,k}\right) + \mathbb{E}\left(|\epsilon_{n,k+1}|^2 \mid \mathcal{F}_{n,k}\right)\right] < \infty.$$

The condition (2.3) moderates the covariance between $x_{n,k}$ and $\epsilon_{n,k+1}$. The following example shows that this condition cannot be removed.

Example

Let $\epsilon_k, k \ge 1$, be a sequence of independent and identically distributed (i.i.d.) random variables with $\mathbb{E}\epsilon_1 = 0$ and $\mathbb{E}\epsilon_1^2 = 1$. Write $S_n = \sum_{k=1}^n x_{n,k}\epsilon_{k+1}$, where $x_{n,k} = n^{-1}\sum_{j=1}^k \epsilon_j$.

Obviously, $\{\epsilon_k, \mathcal{F}_k\}_{k\ge 1}$, where $\mathcal{F}_k = \sigma(\epsilon_1, ..., \epsilon_k)$, forms a martingale difference, Assumption 1 holds with $\eta_k = \epsilon_k$ and $\max_{1\le k\le n}|x_{n,k}| = o_P(1)$. Furthermore, we have that

$$S_n \to_D \int_0^1 W(t)dW(t) \quad \text{and} \quad \left\{\frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}\eta_k, \sum_{k=1}^n x_{n,k}^2\right\} \Rightarrow \left\{W(t), \int_0^1 W^2(s)ds\right\},$$

where $W(t)$ is a standard Wiener process. However, (2.6) is not true for this example, since $\int_0^1 W(t)dW(t) \ne_D \left(\int_0^1 W^2(t)dt\right)^{1/2} N$. On the other hand, (2.3) cannot be satisfied, due to

$$\frac{1}{\sqrt{n}}\sum_{k=1}^n |x_{n,k}|\left|\mathbb{E}\left(\epsilon_{k+1}^2 \mid \mathcal{F}_k\right)\right| = \frac{1}{n^{3/2}}\sum_{k=1}^n \left|\sum_{j=1}^k \epsilon_j\right| \to_D \int_0^1 |W(t)|dt.$$

Remark 4. Given $G_n^2 \to_D g^2(W, \Xi)$, one may enlarge the probability space in which $x_{n,1}, ..., x_{n,n}$ are equipped so that, in this new space, an equivalent process V_n^2 to G_n^2 (that is, $V_n^2 =_D G_n^2$ for each $n \ge 1$) exists and $V_n^2 \to_P g^2(W, \Xi)$. This kind of fact fails to be useful in investigating the asymptotics for S_n by the classical MLT, since, without independence between $\epsilon_{n,t}$ and $x_{n,t}$, this enlarged probability space may destroy the nested σ-fields structure (i.e., $\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i}$) required in the classical MLT.

To illustrate, we consider the $S_n = \sum_{k=1}^n x_{n,k}\epsilon_{k+1}$, where $x_{n,k} = \sum_{j=1}^k \epsilon_j/n$, defined in the example above. Using Skorohod's embedding scheme, there exists a probability space with a Wiener process $\{W(t), 0 \le t < \infty\}$ and a triangular sequence $\{\tau_1^{(n)}, ..., \tau_n^{(n)}\}$ of nonnegative r.v. defined on it such that

(i) $\tau_j^{(n)}, 1 \le j \le n$, are i.i.d. random variables for each $n \ge 1$ and with

$$\tau_1^{(n)} =_D \tau_1^{(1)}, n = 1, 2, ...$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

(ii) $\{x_{n,k}, k = 1, ..., n\} =_D \{\widetilde{x}_{n,k}, k = 1, .., n\}$, for each $n \geq 1$, where

$$\widetilde{x}_{n,k} = n^{-1/2} W\left(\frac{\tau_1^{(n)} + ... + \tau_k^{(n)}}{n}\right).$$

Note that $\sup_{0\leq t\leq 1} |\sqrt{n}\widetilde{x}_{n,[nt]} - W(t)| = o_P(1)$. See, e.g., Theorem 2.1.2 of Csörgö and Révész (1981). Simple calculation shows that $\sum_{k=1}^n \widetilde{x}_{n,k}^2 \to_P \int_0^1 W^2(t)dt$ and $S_n =_d \widetilde{S}_{n,n}$, for each $n \geq 1$, where $\widetilde{S}_{n,j} = \sum_{k=1}^j \widetilde{x}_{n,k}\widetilde{\epsilon}_{n,k+1}$, $1 \leq j \leq n$, and

$$\widetilde{\epsilon}_{n,k+1} = n^{1/2}\left[W\left(\frac{\tau_1^{(n)} + ... + \tau_{k+1}^{(n)}}{n}\right) - W\left(\frac{\tau_1^{(n)} + ... + \tau_k^{(n)}}{n}\right)\right].$$

Furthermore, $\widetilde{x}_{n,k} \in \mathcal{F}_{n,k}$ and $\{\widetilde{\epsilon}_{n,k}, \mathcal{F}_{n,k}\}$ form a martingale difference, where

$$\mathcal{F}_{n,k} = \sigma\left[W(t), t \leq \left(\tau_1^{(n)} + ... + \tau_k^{(n)}\right)/n\right].$$

In this extended probability space, the conditional variance $\sum_{k=1}^n \widetilde{x}_{n,k}^2$ of the martingale array $\{\widetilde{S}_{n,j}, \mathcal{F}_{n,j}\}_{1\leq j\leq n}$ converges in probability to a random variable $\int_0^1 W^2(t)dt$, but the nested structure of the σ-fields $\mathcal{F}_{n,k}$ fails to hold, due to

$$\left(\tau_1^{(n)} + ... + \tau_k^{(n)}\right)/n \not\leq \left(\tau_1^{(n+1)} + ... + \tau_k^{(n+1)}\right)/(n+1).$$

As a consequence, the classical MLT cannot be employed to investigate the limit distribution of $\widetilde{S}_{n,n}$ (hence that of S_n).

Remark 5. Let $\sigma_{1n}^2 = \sum_{k=1}^n x_{n,k}^2 \epsilon_{n,k+1}^2$ and $\sigma_{2n}^2 = \sum_{k=1}^n x_{n,k}^2 \mathbb{E}(\epsilon_{n,k+1}^2 \mid \mathcal{F}_{n,k})$. Using Assumption 1, $G_n^2 = \sum_{k=1}^n x_{n,k}^2 \to_D g^2(W, \Xi)$ and $\max_{1\leq k\leq n} |x_{n,k}| = o_P(1)$, we have that $|\sigma_{2n}^2 - G_n^2| = o_P(1)$, σ_{2n}^2 is tight due to $\sigma_{2n}^2 \to_D g^2(W, \Xi)$ and for any $\epsilon > 0$,

$$\begin{aligned}
&\sum_{k=1}^n x_{n,k}^2 \mathbb{E}\left(\epsilon_{n,k+1}^2 I(|x_{n,k}\epsilon_{n,k+1}| \geq \epsilon) \mid \mathcal{F}_{n,k}\right) \\
&\quad \leq G_n^2\left[\max_{1\leq k\leq n} \mathbb{E}\left(\epsilon_{n,k+1}^2 I(|\epsilon_{n,k+1}| \geq \epsilon) \mid \mathcal{F}_{n,k}\right) + \max_{1\leq k\leq n} |x_{n,k}|\right] \\
&\quad = o_P(1).
\end{aligned}$$

Now, by Theorem 2.23 of Hall and Heyde (1980), we also have $|\sigma_{1n}^2 - G_n^2| = o_P(1)$. Hence, Theorem 2.1 still holds if we replace G_n^2 by σ_{1n}^2 or σ_{2n}^2. S_n/σ_{1n} is the so-called self-normalized sum for a martingale.

Remark 6. If, in addition, $S_{[nt]}$ is uniformly integrable for each $0 \leq t \leq 1$ and Assumption 4(ii) is strengthened to: There exists a continuous stochastic process

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

$g_t^2(W,\Xi)$ in t of $W(s), 0 \le s \le 1$ and $\Xi(s), 0 \le s < \infty$, such that

$$\left\{W_n(t),\ \Xi_n(t),\ G^2_{[nt]}\right\} \Rightarrow \left\{W(t),\ \Xi(t),\ g_t^2(W,\Xi)\right\}, \tag{2.9}$$

the result (2.7) can be improved to the following weak convergence on

$$\left\{S_{[nt]},\ G_n^2\right\} \Rightarrow \left\{B\left[g_t^2(W,\Xi)\right],\ g^2(W,\Xi)\right\}, \tag{2.10}$$

where $B(s), s \ge 0$, is a standard Brownian motion independent of $g_t(W,\Xi)$. The proof of (2.10) can be found in Section 3.7.

Let $X_t = B\left[g_t^2(W,\Xi)\right]$. By the notation in Chapter IX of Jacod and Shiryaev (2003) (also see Appendix B of Ibragimov and Phillips, 2008), the second predictable characteristics of $S_{[nt]}$ and X_t are $C_n(t) = \sum_{k=1}^{[nt]} x_{n,k}^2 \epsilon_{n,k+1}^2$ and $C(t,X) \equiv g_t^2(W,\Xi)$, respectively. As a consequence, the condition $[\gamma_7 - D]$ used in the semi-martingale asymptotic theorems of Jacod and Shiryaev (2003, Thm. IX.3.21, for instance) states

$$C_n(t) - C(t,S_n) = \sum_{k=1}^{[nt]} x_{n,k}^2 \epsilon_{n,k+1}^2 - g_t^2(W,\Xi) \to_P 0, \tag{2.11}$$

for each $0 < t \le 1$. Unless $g_t(W,\Xi)$ is a deterministic function of t, this is hard to verify in many applications, as stated in Remark 2. In comparison, our condition (2.9) again has the advantages, as we only require convergence in distribution.

For a certain class of semi-martingales $X_n(t)$, if the limit process $X(t)$ has the second predictable characteristic $C(t,X) = \int_0^t g\left[X(s)\right]ds$ where g is a real function satisfying certain conditions (for instance, $X(t)$ can be denoted as a solution of a diffusion process), it is generally easier to verify the condition $\left[\gamma_7 - D\right]$ of Jacod and Shiryaev (2003):

$$C_n(t) - C\left(t, X_n\right) = \left[X_n, X_n\right](t) - \int_0^t g\left[X_n(s)\right]ds \to_P 0, \tag{2.12}$$

where $\left[X_n, X_n\right](t)$ is the quadratic variation process of $X_n(t)$, because, as noticed in Ibragimov and Phillips (2008), the two components in (2.12) are defined on the same probability space and only involve $X_n(t)$ and not the limit process. It should be mentioned that, except for the trivial situation where $g_t(W,\Xi)$ is a deterministic function of t, the second predictable characteristic $g_t^2(W,\Xi)$ of the limit process $X_t = B\left[g_t^2(W,\Xi)\right]$ in (2.10) cannot be denoted as a functional of X_t due to the independence of $B(s)$ and $g_t(W,\Xi)$.

2.2. An Application to Linear Processes

In what follows we consider an application of Theorem 2.1 to general linear processes. Let $\{\zeta_j, j \ge 1\}$ be a linear process defined by

$$\zeta_j = \sum_{k=0}^{\infty} \phi_k \nu_{j-k}, \tag{2.13}$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

where $\{v_j, -\infty < j < \infty\}$ is a sequence of i.i.d. random variables with $\mathbb{E}v_0 = 0$, $\mathbb{E}v_0^2 = 1$, $\mathbb{E}|v_0|^{2+\delta} < \infty$ for some $\delta > 0$ and characteristic function $\varphi(t)$ of v_0 satisfying $\int_{-\infty}^{\infty} |\varphi(t)|dt < \infty$. Throughout the section, the coefficients $\phi_k, k \geq 0$, are assumed to satisfy one of the following conditions:

Condition 1 $\phi_k \sim k^{-\mu}\rho(k)$, where $1/2 < \mu < 1$ and $\rho(k)$ is a function slowly varying at ∞.

Condition 2 $\sum_{k=0}^{\infty} |\phi_k| < \infty$ and $\phi \equiv \sum_{k=0}^{\infty} \phi_k \neq 0$.

Put $y_k = \sum_{j=1}^{k} \zeta_j$ and $y_{n,k} = y_k/d_n + c_n'$, where $c_n' \to 0$ and $d_n^2 = \mathbb{E}y_n^2$. It is well-known that, with $c_\mu = \frac{1}{(1-\mu)(3-2\mu)} \int_0^\infty x^{-\mu}(x+1)^{-\mu}dx$,

$$d_n^2 = \mathbb{E}y_n^2 \sim \begin{cases} c_\mu n^{3-2\mu}\rho^2(n), & \text{under } \textbf{Condition 1} \\ \phi^2 n, & \text{under } \textbf{Condition 2} \end{cases} \tag{2.14}$$

See, e.g., Wang, Lin, and Gulati (2003) for instance. We consider the limit behavior of sample functions of the form

$$S_{1n} = \left(\frac{c_n}{n}\right)^{1/2} \sum_{k=1}^{n} f\left(c_n\, y_{n,k}\right)\epsilon_{k+1}, \tag{2.15}$$

when $c_n \to \infty$ and $c_n/n \to 0$, under the following conditions.

Assumption 5. Assume f is a real function on $\mathbf{R}$ satisfying $\int_{-\infty}^{\infty} \left[|f(x)| + |f(x)|^{4+\gamma}\right]dx < \infty$ for some $\gamma > 0$ and $\int_{-\infty}^{\infty} f^2(x)dx \neq 0$.

Assumption 6. (i) $\mathcal{F}_k$ is a sequence of increasing σ-fields such that $v_k \in \mathcal{F}_k$ and v_{k+1} is independent of $\mathcal{F}_k$ for all $k \geq 1$, and $v_k \in \mathcal{F}_1$ for all $k \leq 0$. (ii) $\{\epsilon_k, \mathcal{F}_k\}_{k\geq 1}$ forms a martingale difference satisfying, as $n \to \infty$, $\max_{1\leq k\leq n} |\mathbb{E}(\epsilon_{k+1}^2 \mid \mathcal{F}_k) - 1| \to 0$, a.s. and for any constant sequence $0 < A_n \to \infty$, $\max_{1\leq k\leq n} \mathbb{E}\left(|\epsilon_{k+1}|^2 I(|\epsilon_{k+1}| \geq A_n) \mid \mathcal{F}_k\right) = o_P(1)$.

Assumption 6 is a particular version of Assumption 1 by making use of the independence between v_k. Typically, in applications, we may choose $\mathcal{F}_k = \sigma\{v_k, v_{k-1}, ...\}$ or $\mathcal{F}_k = \sigma\{\epsilon_k, ..., \epsilon_1; v_k, v_{k-1}, ...\}$ together with the condition that v_{k+1} is independent of $\mathcal{F}_k$ for all $k \geq 1$. To investigate the asymptotics of S_{1n}, we start with the following notation. A fractional Brownian motion with Hurst parameter $0 < \beta < 1$ on $D[0,1]$ is defined by

$$W_\beta(t) = \frac{1}{A(\beta)} \left\{ \int_{-\infty}^{0} \left[(t-s)^{\beta-1/2} - (-s)^{\beta-1/2}\right] dW^*(-s) + \int_0^t (t-s)^{\beta-1/2} dW(s) \right\},$$

where

$$A(\beta) = \left(\frac{1}{2\beta} + \int_0^\infty \left[(1+s)^{\beta-1/2} - s^{\beta-1/2}\right]^2 ds\right)^{1/2},$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

$W(s), 0 \le s < \infty$ is a standard Brownian motion, and $W^*(u), 0 \le u < \infty$ is an independent copy of $W(s), 0 \le s < \infty$. It is readily seen that $W_{1/2}(t) = W(t)$ and $W_\beta(t)$ has a continuous local time $L_{W_\beta}(t,s)$ with regard to (t,s) in $[0,\infty) \times \mathbf{R}$. See, e.g., Theorem 22.1 of Geman and Horowitz (1980). Here and below, the process $\{L_\zeta(t,s), t \ge 0, s \in R\}$ is said to be the local time of a measurable process $\{\zeta(t), t \ge 0\}$ if, for any locally integrable function $T(x)$,

$$\int_0^t T[\zeta(s)]ds = \int_{-\infty}^{\infty} T(s) L_\zeta(t,s) ds, \quad \text{all } t \in \mathbf{R},$$

with probability one.

Note that $L_{W_\beta}(1,0)$ is a functional of $W(s), 0 \le s \le 1$ and $W^*(u), 0 \le u < \infty$. Write $G_{1n}^2 = \frac{c_n}{n} \sum_{k=1}^n f^2(c_n y_{n,k})$. As a direct consequence of Theorem 2.1, we have the following theorem.

THEOREM 2.2. *Under Assumptions 5 and 6, when $c_n \to \infty$ and $c_n/n \to 0$, we have*

$$\left\{S_{1n},\ G_{1n}^2\right\} \to_D \left\{g(W,W^*)N,\ g^2(W,W^*)\right\}, \tag{2.16}$$

where N is a standard normal variate independent of $g^2(W,W^*) = \int_{-\infty}^{\infty} f^2(x)dx\, \mathcal{L}_W$ with $\mathcal{L}_W$ being defined by

$$\mathcal{L}_W = \begin{cases} L_{W_{3/2-u}}(1,0), & \textit{under } \textbf{Condition 1} \\ L_W(1,0), & \textit{under } \textbf{Condition 2} \end{cases}$$

Consequently,

$$S_{1n}/G_{1n} \to_D N(0,1). \tag{2.17}$$

Remark 7. The result (2.16) under **Condition 1** is new, in which the $y_{n,k}$ includes the fractionally integrated process as an example. Theorem 3.2 of Park and Phillips (2001) investigated the asymptotics of S_{1n} under **Condition 2**, but only allows the c_n to be equal $\left(\mathbb{E}y_n^2\right)^{1/2} = \phi n^{1/2}$ and imposes strong restrictions on f and v_0. Using similar ideas as in Park and Phillips (2001), Wang and Phillips (2009a) considered more general situations under **Condition 2** (the result is included in the proof of their Theorem 3.2), but still imposed some additional conditions. The present paper improves these existing results. Furthermore, unlike these existing results, our proof is quite neat and straightforward by Theorem 2.1, namely the extended martingale limit theorem.

Remark 8. Under the conditions of Theorem 2.2, one always has $cov\left(y_t, \epsilon_{t+1}\right) = \mathbb{E}\left[y_t \mathbb{E}(\epsilon_{t+1} \mid \mathcal{F}_t)\right] = 0$. Wang and Phillips (2009b) investigated the asymptotics of S_{1n} under **Condition 2**, imposing the ϵ_{t+1} to be serially dependent and cross correlated with y_s for $|t-s| < m_0$, where m_0 is a given constant. As a consequence, in Wang and Phillips (2009b), one may have $cov(y_t, \epsilon_{t+1}) \ne 0$. This makes the functionals in Wang and Phillips (2009b) essentially different from the one investigated in current Theorem 2.2. It should be mentioned that

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

(2.16) does not follow from Wang and Phillips (2009b), because in that paper, the regression disturbance distribution ϵ_t is required to be strictly stationary (their Assumption 2), whereas here ϵ_t is permitted to be a martingale difference, which may fail to be strictly stationary.

Remark 9. It is interesting to notice that the result would be quite different if $c_n = 1$. Indeed, under much weaker conditions on f, $y_{n,k}$ and ϵ_k which adds to

$$\left\{ y_{n,[nt]}, \frac{1}{\sqrt{n}} \sum_{j=1}^{[nt]} \epsilon_{j+1} \right\} \Rightarrow \{U(t), V(t)\},$$

one may prove $\frac{1}{\sqrt{n}} \sum_{k=1}^{n} f\left(y_{n,k}\right)\epsilon_{k+1} \to_D \int_0^1 f[U(s)]dV(s)$, a stochastic integral. There are many papers in which the convergence to stochastic integrals are investigated. We only cite Kurtz and Protter (1991) and Hansen (1992) for reference. For more current works, see Ibragimov and Phillips (2008) and the references therein.

Remark 10. The individual asymptotics of the functionals like G_{1n}^2 have been investigated in Wang and Phillips (2009a) (also see Wang and Phillips (2011) for the zero energy situation) under weaker conditions than Assumption 5. Explicitly, it follows from (2.5) of Wang and Phillips (2011) [1] that, for each $x \in \mathbf{R}$ and all h satisfying $h \to 0$ ($h^2 \log n \to 0$ under **Condition 2**) and $d_n/nh \to 0$,

$$\left\{ \left(\frac{d_n}{nh}\right)^{1/2} \sum_{k=1}^{n} g\left[(y_k - x)/h\right], \frac{d_n}{nh} \sum_{k=1}^{n} f_1\left[(y_k - x)/h\right], \frac{d_n}{nh} \sum_{k=1}^{n} f_2\left[(y_k - x)/h\right] \right\}$$
$$\to_D \left\{ \tau_0 N \mathcal{L}_W^{1/2}, \tau_1 \mathcal{L}_W, \tau_2 \mathcal{L}_W \right\}, \tag{2.18}$$

where $d_n^2 = \mathbb{E} y_n^2$ is given as in (2.14), $\tau_0^2 = \int_{-\infty}^{\infty} g^2(x)dx$, $\tau_1 = \int f_1(x)dx$, and $\tau_2 = \int f_2(x)dx$, provided that $\int |g(t)|dt < \infty$, $\int |\hat{g}(t)|dt < \infty$ and $|\hat{g}(t)| \le C \min\{|t|, 1\}$, where $\hat{g}(x) = \int e^{itx} g(t)dt$; $|f_j(x)|$ and $f_j^2(x)$, $j = 1, 2$, are Lebesgue integrable functions on $\mathbf{R}$ with $\tau_1 \neq 0$ and $\tau_2 \neq 0$. It is readily seen from (2.18) that

$$\frac{\left(\frac{d_n}{nh}\right)^{-1/2} \sum_{k=1}^{n} g\left[(y_k - x)/h\right]}{\sum_{k=1}^{n} f_2[(y_k - x)/h]} \to_D \frac{\tau_0}{\tau_2} N \mathcal{L}_W^{-1/2}, \tag{2.19}$$

$$\frac{\sum_{k=1}^{n} f_1\left[(y_k - x)/h\right]}{\sum_{k=1}^{n} f_2\left[(y_k - x)/h\right]} \to_P \frac{\tau_1}{\tau_2}. \tag{2.20}$$

These results will be used in the proofs of Theorems 2.2–2.4.

Remark 11. The proof of Theorem 2.2 also depends on the following joint weak convergence, which is stated as a proposition for convenience of reference.

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

PROPOSITION 2.1. *Under Assumption 5, for any $c_n/n \to 0$ and $c_n \to \infty$, we have*

$$\left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^{[nt]} v_j,\ \frac{1}{\sqrt{n}} \sum_{j=1}^{[nt]} v_{-j+1},\ y_{n,[nt]},\ \frac{c_n}{n} \sum_{k=1}^{n} f^2(c_n y_{n,k}) \right\}$$
$$\Rightarrow \left\{ W(t),\ W^*(t),\ Y(t),\ g^2(W, W^*) \right\}, \tag{2.21}$$

where

$$Y(t) = \begin{cases} W_{3/2-u}(t), & \textit{under } \textbf{Condition 1} \\ W(t), & \textit{under } \textbf{Condition 2} \end{cases}$$

The proof of Proposition 2.1 is simple. Indeed, the joint convergence of the first three components is well-known. See, e.g., Kasahara and Maejima (1988). On the other hand, it follows from the proof of Theorem 2.1 in Wang and Phillips (2009a) that, $\forall \epsilon > 0$,

$$\frac{c_n}{n} \sum_{k=1}^{n} f^2(c_n y_{n,k}) = \int_{-\infty}^{\infty} f^2(x)dx \int_0^1 \phi_\epsilon(y_{n,[nt]})dt + o_P(1), \tag{2.22}$$

where $\phi_\epsilon(x) = \frac{1}{\epsilon\sqrt{2\pi}} \exp\left\{ -\frac{x^2}{2\epsilon^2} \right\}$. Since the continuous mapping theorem implies that, $\forall \epsilon > 0$,

$$\left\{ \frac{1}{\sqrt{n}} \sum_{j=1}^{[nt]} v_j,\ \frac{1}{\sqrt{n}} \sum_{j=1}^{[nt]} v_{-j+1},\ y_{n,[nt]},\ \int_0^1 \phi_\epsilon(y_{n,[nt]})dt \right\}$$
$$\Rightarrow \left\{ W(t),\ W^*(t),\ Y(t),\ \int_0^1 \phi_\epsilon[Y(t)]dt \right\}, \tag{2.23}$$

the result (2.21) follows immediately from (2.22)–(2.23) and the fact that

$$\int_0^1 \phi_\epsilon[Y(t)]dt = \int_{-\infty}^{\infty} \phi(x) L_Y(1, \epsilon x)dx = \mathcal{L}_W + o_{a.s.}(1),$$

as $\epsilon \to 0$.

2.3. Nonlinear Cointegrating Regression

As stated in the Introduction, the results given in Theorem 2.2 play a key role in the investigation of nonstationary cointegration regression. To illustrate, consider the nonlinear cointegrating regression model

$$z_t = m(y_t) + \epsilon_{t+1}, \qquad t = 1, 2, ..., n, \tag{2.24}$$

where ϵ_{t+1} is a zero mean stationary equilibrium error and m is an unknown function to be estimated with the observed data $\{z_t, y_t\}_{t=1}^n$. The conventional Nadaraya-Watson estimator of $m(x)$ in model (2.24) is given by

$$\hat{m}(x) = \frac{\sum_{t=1}^n z_t K_h(y_t - x)}{\sum_{t=1}^n K_h(y_t - x)}, \tag{2.25}$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

where $K_h(s) = \frac{1}{h}K(s/h)$, $K(x)$ is a nonnegative real function, and the bandwidth parameter $h \equiv h_n \to 0$ as $n \to \infty$. Note that $\hat{m}(x)$ has the usual decomposition,

$$\hat{m}(x) - m(x) = \frac{\sum_{t=1}^n \epsilon_{t+1} K_h(y_t - x)}{\sum_{t=1}^n K_h(y_t - x)} + \frac{\sum_{t=1}^n \left[m(y_t) - m(x)\right] K_h(y_t - x)}{\sum_{t=1}^n K_h(y_t - x)}, \quad \textbf{(2.26)}$$

and it is the first term to determine the asymptotic distribution, depending on the asymptotics of

$$(S_{2n}, G_{2n}^2) := \left\{ \left(\frac{c_n}{n}\right)^{1/2} \sum_{t=1}^n \epsilon_{t+1} K(c_n\, y_{t,n}), \quad \frac{c_n}{n} \sum_{t=1}^n K^2(c_n\, y_{t,n}) \right\}, \quad \textbf{(2.27)}$$

where $y_{t,n} = (y_t - x)/d_n$, $d_n^2 = \mathbb{E}y_n^2$, and $c_n = d_n/h$. An application of Theorem 2.2 to $\left(S_{2n}, G_{2n}^2\right)$ improves the existing results on the asymptotics of the kernel estimator $\hat{m}(x)$ with nonstationary regressors, which was investigated in Wang and Phillips (2009a, 2011). To do this, we require the following additional assumptions.

Assumption 7. Assume $y_t = \sum_{j=1}^t \xi_j$, where ξ_j is defined as in (2.13) with ϕ_k satisfying **Condition 1** or **2**.

Assumption 8. The kernel K satisfies $\int_{-\infty}^{\infty} K(s)ds = 1$ and $\sup_s K(s) < \infty$.

Assumption 9. For given x, there exists a real positive function $m_1(s,x)$ and $0 < \gamma \leq 1$ such that, when δ is sufficiently small, $|m(\delta y + x) - m(x)| \leq \delta^\gamma m_1(y,x)$ for all $y \in \mathbf{R}$ and $\int_{-\infty}^{\infty} K(s)\,[m_1(s,x) + m_1^2(s,x)]ds < \infty$.

THEOREM 2.3. *Under Assumptions 6–9, for any h satisfying $nh/d_n \to \infty$ and $nh^{1+2\gamma}/d_n \to 0$, we have*

$$\left(\frac{d_n}{nh}\right)^{-1/2} (\hat{m}(x) - m(x)) \to_D \tau\, N\, \mathcal{L}_W^{-1/2} \quad \textbf{(2.28)}$$

where $d_n^2 = \mathbb{E}y_n^2$, given as in (2.14), $\tau^2 = \int_{-\infty}^{\infty} K^2(s)ds$, and N is a standard normal variate independent of $\mathcal{L}_W$, which is given as in Theorem 2.2. We also have

$$\left(h \sum_{t=1}^n K_h(y_t - x)\right)^{1/2} (\hat{m}(x) - m(x)) \to_D \tau\, N. \quad \textbf{(2.29)}$$

If Assumptions 8 and 9 are strengthened to the following Assumptions 8* and 9*, an explicit bias term may be incorporated into the limit theory (2.28) and (2.29).

Assumption 8*. (i) $K(x)$ satisfies that $\int K(y)dy = 1$ and for some $p \geq 2$,

$$\int y^p K(y)dy \neq 0, \qquad \int y^i K(y)dy = 0, \qquad i = 1, 2, ..., p-1.$$

(ii) $K(x)$ has compact support and is twice continuously differentiable on $\mathbf{R}$.

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

Assumption 9*. For given fixed x, $m(x)$ has a continuous $p+1$ derivative in a small neighborhood of x, where $p \geq 2$ is defined as in Assumption 8*.

THEOREM 2.4. *Under Assumptions 6–7 and* 8*–9*, *for any* h *satisfying* $nh/d_n \to \infty$ *and* $nh^{1+2(p+1)}/d_n \to 0$, *we have*

$$\left(\frac{d_n}{nh}\right)^{-1/2}\left[\hat{m}(x)-m(x)-\frac{h^p m^{(p)}(x)}{p!}\int_{-\infty}^{\infty} y^p K(y)dy\right] \to_D \tau N \mathcal{L}_W^{-1/2}, \tag{2.30}$$

and also,

$$\left(h\sum_{t=1}^{n} K_h(y_t-x)\right)^{1/2}\left[\hat{m}(x)-m(x)-\frac{h^p m^{(p)}(x)}{p!}\int_{-\infty}^{\infty} y^p K(y)dy\right] \to_D \tau N. \tag{2.31}$$

Remark 12. In the past decade there has been increasing interest in the investigation of asymptotics for the kernel estimator $\hat{m}(x)$ of $m(x)$ in the model (2.24), under different settings on the regressor y_t and the error process ϵ_t. Phillips and Park (1998) studied nonparametric autoregression in the context of a random walk. Karlsen and Thostheim (2001) and Guerre (2004) studied nonparametric estimation for certain nonstationary processes in the framework of recurrent Markov chains. Karlsen et al. (2007) developed an asymptotic theory for nonparametric estimation of a time series regression equation involving stochastically nonstationary time series. Under similar conditions and using related Markov chain methods, Schienle (2008) investigated additive nonlinear versions of (2.24) and obtained a limit theory for non-parametric regressions under smooth backfitting.

More recently, Wang and Phillips (2009a) and Cai, Li and Park (2009) considered an alternative treatment by making use of local time limit theory and, instead of recurrent Markov chains, worked with partial sum representations y_t of linear processes, as given in Assumption 7. Unlike these cited results where the independence between the regressor y_t and the error Process ϵ_t is essentially imposed, our results allow the endogenous regressor case and also allow the regressor y_t to be fractionally integrated processes. Furthermore Theorem 2.4 investigates the bias analysis, which has not been considered in the previous articles. In other papers, Wang and Phillips (2009b, 2011) considered the errors ϵ_t to be serially dependent and cross correlated with the regressor x_t for small lags. As stated in Remark 8, these works have a different error structure to that used in Theorems 2.3 and 2.4.

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

3. PROOFS OF MAIN RESULTS

3.1. Proof of Theorem 2.1.

We only prove (2.7), as (2.6) is similar but simpler. For convenience of notation, we omit the index n on $\eta_{n,k}$, $\epsilon_{n,k}$, and $\xi_{n,k}$. Suppose first that $\sigma^2 := g^2(W, \Xi)$ is a.s. bounded, so that, for some $\lambda > 1$,

$$P(\sigma^2 \le \lambda) = 1. \tag{3.1}$$

Let

$$x^*_{n,k} = x_{n,k} I\left(\sum_{j=1}^{k} x^2_{n,j} \le 2\lambda\right), \qquad S^*_n = \sum_{k=1}^{n} x^*_{n,k}\epsilon_{k+1}, \quad \text{and} \quad G^{*2}_n = \sum_{k=1}^{n} x^{*2}_{n,k}.$$

Note that $S^*_n = S_n$ and $G^{*2}_n = G^2_n$ on the set $G^2_n \le 2\lambda$ and $P(G^2_n > 2\lambda) \to 0$ due to (2.5) and (3.1). It is readily seen that, for all $\alpha, \beta \in \mathbf{R}$,

$$\mathbb{E}\left|e^{i\alpha S^*_n + i\beta G^{*2}_n} - e^{i\alpha S_n + i\beta G^2_n}\right| \le 2P(G^2_n > 2\lambda) \to 0.$$

Hence, $\mathbb{E}e^{i\alpha S_n + i\beta G^2_n} \to \mathbb{E}e^{(-\alpha^2/2 + i\beta)\sigma^2}$ if and only if

$$\mathbb{E}e^{i\alpha S^*_n + i\beta G^{*2}_n} \to \mathbb{E}e^{(-\alpha^2/2 + i\beta)\sigma^2}. \tag{3.2}$$

Consequently, to prove (2.7), it suffices to show (3.2).

The idea in the proof of (3.2) is partially encouraged by Jeganathan (1982). Also see Jeganathan (2006). The essential difference in the proof is that, rather than $G^{*2}_n \to_P \sigma^2$, we only have the condition

$$\left\{W_n(t),\ \Xi_n(t),\ G^{*2}_n\right\} \Rightarrow \left\{W(t),\ \Xi(t),\ \sigma^2\right\}, \tag{3.3}$$

which comes again from (2.5) and the fact that $G^{*2}_n = G^2_n$ on the set $G^2_n \le 2\lambda$ and $P\left(G^2_n > 2\lambda\right) \to 0$.

To prove (3.2), we start with some preliminaries. For $\beta_{1k}, \beta_{2k} \in R$, $0 = u_{10} < u_{11} < \cdots < u_{1N_1} = 1$ and $0 = u_{20} < u_{21} < \cdots < u_{2N_2} < \infty$, define

$$V_{1n} = \sum_{k=1}^{N_1} \beta_{1k}\left[W_n\left(u_{1k}\right) - W_n\left(u_{1,k-1}\right)\right],$$

$$V_{2n} = \sum_{k=1}^{N_2} \beta_{2k}\left[\Xi_n\left(u_{2k}\right) - \Xi_n\left(u_{2,k-1}\right)\right],$$

where $W_n(t)$ and $\Xi_n(t)$ appear as in Assumption 4, and

$$\Gamma_n = \sum_{k=1}^{n} \mathbb{E}\left\{\left(e^{i\beta^*_{1k}\eta_{k+1} + i\alpha x^*_{n,k}\epsilon_{k+1}} - 1\right) \mid \mathcal{F}_{n,k}\right\}, \tag{3.4}$$

where $\beta^*_{1k} = \beta_{1j}/\sqrt{n}$ when $\left[nu_{1,j-1}\right] < k \le [nu_{1j}]$, for $j = 1, ..., N_1$. We have the following two propositions. Their proofs will be given in Sections 3.5 and 3.6, respectively.

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

PROPOSITION 3.1. *For any* $\alpha, \beta_{1k} \in \mathbf{R}$ *and* $0 = u_{10} < u_{11} < \ldots < u_{1N_1} = 1$, *we have that* $e^{|\Gamma_n|}$ *is uniformly integrable and*

$$\left\{V_{1n},\ V_{2n},\ G_n^{*2},\ \Gamma_n\right\} \to_D \left\{V_1,\ V_2,\ \sigma^2,\ \Gamma\right\}, \tag{3.5}$$

where $\Gamma = -\frac{1}{2}\sum_{k=1}^{N_1} \beta_{1k}^2 (u_{1k} - u_{1,k-1}) - \frac{1}{2}\alpha^2\sigma^2$ *and*

$$V_1 = \sum_{k=1}^{N_1} \beta_{1k} \left[W\left(u_{1k}\right) - W\left(u_{1,k-1}\right)\right], \qquad V_2 = \sum_{k=1}^{N_2} \beta_{2k} \left[\Xi\left(u_{2k}\right) - \Xi\left(u_{2,k-1}\right)\right].$$

PROPOSITION 3.2. *For any* $\alpha, \beta_{1k}, \beta_{2k} \in \mathbf{R}$, $0 = u_{10} < u_{11} < \cdots < u_{1N_1} = 1$ *and* $0 = u_{20} < u_{21} < \cdots < u_{2N_2} < \infty$, *we have*

$$I_n := \mathbb{E}\left|\mathbb{E}\left(e^{i\alpha S_n^* + iV_{1n} - \Gamma_n} \mid \mathcal{F}_{n,1}\right) - 1\right| = o(1). \tag{3.6}$$

Furthermore, as $n \to \infty$, *we have*

$$\mathbb{E}e^{i\alpha S_n^* + iV_{1n} + iV_{2n} - \Gamma_n} \to \mathbb{E}e^{iV_2}. \tag{3.7}$$

We are now ready to prove (3.2). First note that $\left\{S_n^*\right\}_{n\geq 1}$ is tight, due to $\mathbb{E}S_n^{*2} \leq C\,\mathbb{E}G_n^{*2} \leq 2C\,\lambda$. This, together with (3.5), yields that

$$\left\{S_n^*,\ G_n^{*2},\ V_{1n},\ V_{2n},\ \Gamma_n\right\}_{n\geq 1}$$

is tight. Hence, for each $\{n'\} \subseteq \{n\}$, there exists a subsequence $\{n''\} \subseteq \{n'\}$ such that

$$\left\{S_{n''}^*,\ G_{n''}^{*2},\ V_{1n''},\ V_{2n''},\ \Gamma_{n''}\right\} \to_D \left\{S,\ \sigma^2,\ V_1,\ V_2,\ \Gamma\right\}, \tag{3.8}$$

where S is a limit random variable of $S_{n''}^*$. Since $\mathbb{E}e^{iS_{n''}^* + iV_{1n''} + iV_{2n''} - \Gamma_{n''}}$ is uniformly integrable due to Proposition 3.1, as $n'' \to \infty$, we have

$$\mathbb{E}e^{i\alpha S_{n''}^* + iV_{1n''} + iV_{2n''} - \Gamma_{n''}} \to \mathbb{E}e^{i\alpha S + iV_1 + iV_2 - \Gamma}.$$

This, together with (3.7), yields that $\mathbb{E}e^{i\alpha S + iV_1 + iV_2 - \Gamma} = \mathbb{E}e^{iV_2}$. Hence, by noting $\mathbb{E}e^{iV_1} = e^{-1/2\sum_{k=1}^{N_1} \beta_{1k}^2 (u_{1k} - u_{1,k-1})}$, it follows from the independence between $W(t)$ and $\Xi(t)$ that

$$\mathbb{E}e^{iV_1 + iV_2 + i\alpha S + 1/2\alpha^2\sigma^2} = e^{-1/2\sum_{k=1}^{N_1} \beta_{1k}^2 (u_{1k} - u_{1,k-1})}\,\mathbb{E}e^{iV_2} = \mathbb{E}e^{iV_1 + iV_2}. \tag{3.9}$$

Due to (3.9), using the Weierstrass theorem on approximation of continuous functions by trigonometric polynomials, we have

$$\mathbb{E}\left[e^{i\alpha S + 1/2\alpha^2\sigma^2} - 1\right] G\left[W\left(u_{10}\right), \ldots, W\left(u_{1N_1}\right); \Xi\left(u_{20}\right), \ldots, \Xi\left(u_{2N_2}\right)\right] = 0, \tag{3.10}$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

for any points $0 = u_{10} < u_{11} < \cdots < u_{1N_1} = 1$ and $0 = u_{20} < u_{21} < \cdots < u_{2N_2} < \infty$, where $G\left(y_0, ..., y_{N_1}; x_0, ..., x_{N_2}\right)$ is an arbitrary bounded continuous function. As $W(t)$ and $\Xi(t)$ are path continuous, the equality (3.10) remains true for an arbitrary bounded Borel measurable function, since each measurable function on $\mathbf{R}$ is the limit (almost everywhere) of a sequence of continuous functions. See Exercise 1.18 of Stein and Shakarchi (2005). As a consequence, we have

$$\mathbb{E}\left(\left[e^{i\alpha S + 1/2\alpha^2\sigma^2} - 1\right] \mid \mathcal{F}\right) = 0,$$

where $\mathcal{F} = \sigma\{W(s), 0 \le s \le 1; \Xi(t), 0 \le t < \infty\}$. Hence, for all $\alpha, \beta \in R$,

$$\mathbb{E}\left(e^{i\alpha S + i\beta\sigma^2} \mid \mathcal{F}\right) = e^{(-1/2\alpha^2 + i\beta)\sigma^2},$$

due to $\sigma^2 \in \mathcal{F}$. Now, by using (3.8) again, for each $\{n'\} \subseteq \{n\}$, there exists a subsequence $\{n''\} \subseteq \{n'\}$ such that

$$\lim_{n'' \to \infty} \mathbb{E}e^{i\alpha S^*_{n''} + i\beta G^{*2}_{n''}} = \mathbb{E}e^{i\alpha S + i\beta\sigma^2} = \mathbb{E}e^{\left(-\alpha^2/2 + i\beta\right)\sigma^2}.$$

This establishes (3.2), as the limitation does not depend on the choice of the subsequence, and also completes the proof in the special case where (3.1) holds.

It remains to remove the boundedness condition (3.1). To this end, for given $\epsilon > 0$, choose a continuous point λ of the distribution function of σ^2 such that $P(\sigma^2 \ge \lambda) \le \epsilon$. Let $\sigma^2_\lambda = \sigma^2 I\left(\sigma^2 \le \lambda\right) + \lambda I(\sigma^2 > \lambda)$, $x'_{n,k} = x_{n,k} I\left(\sum_{j=1}^k x^2_{n,j} \le \lambda\right)$, $S'_n = \sum_{k=1}^n x'_{n,k}\epsilon_{k+1}$,

$$G'^2_n = \sum_{k=1}^n x'^2_{n,k} \quad \text{and} \quad \widetilde{G}^2_n = G^2_n I\left(G^2_n \le \lambda\right) + \lambda\, I\left(G^2_n > \lambda\right).$$

It follows from Assumptions 2 and 3 that $\max_{1\le k\le n} |x'_{n,k}| \le \max_{1\le k\le n} |x_{n,k}| = o_P(1)$ and

$$\begin{aligned}\frac{1}{\sqrt{n}} \sum_{k=1}^n |x'_{n,k}|\, |\mathbb{E}(\eta_{n,k+1}\epsilon_{n,k+1} \mid \mathcal{F}_{n,k})| \\ \le \frac{1}{\sqrt{n}} \sum_{k=1}^n |x_{n,k}|\, |\mathbb{E}(\eta_{n,k+1}\epsilon_{n,k+1} \mid \mathcal{F}_{n,k})| = o_P(1).\end{aligned}$$

Furthermore, by noting $G'^2_n \le \widetilde{G}^2_n$ and

$$\begin{aligned}\widetilde{G}^2_n &= G'^2_n + I(G^2_n > \lambda)\left\{\lambda - \sum_{k=1}^n x^2_{n,k} I\left(G^2_{n,k} \le \lambda\right)\right\} \\ &\le G'^2_n + \max_{1\le k\le n} x^2_{n,k},\end{aligned}$$

where $G^2_{n,k} = \sum_{j=1}^k x^2_{n,j}$ and we have used the identity

$$\sum_{k=1}^n x^2_{n,k} I\left(G^2_{n,k} \le \lambda\right) = \sum_{k=1}^{n-1} x^2_{n,k} I\left(G^2_{n,k} \le \lambda < G^2_{n,k+1}\right) + G^2_n I(G^2_n \le \lambda),$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

we have $G_n'^2 \to_D \sigma_\lambda^2$, since $\widetilde{G}_n^2 \to_D \sigma_\lambda^2$ by the continuous mapping theorem. Now, by recalling σ_λ^2 is a.s. bounded, the first part of the proof shows that

$$\mathbb{E}e^{i\alpha S_n' + i\beta G_n'^2} \to \mathbb{E}e^{(-\alpha^2/2+i\beta)\sigma_\lambda^2}.$$

This, together with the facts that $S_n' = S_n$ and $G_n'^2 = G_n^2$ on the set $G_n^2 \le \lambda$ and $P\left(G_n^2 > \lambda\right) \to P\left(\sigma^2 > \lambda\right)$ due to the continuous mapping theorem, yields that, for all $\epsilon > 0$,

$$\begin{aligned}
&\limsup_{n\to\infty} \left|\mathbb{E}e^{i\alpha S_n + i\beta G_n^2} - \mathbb{E}e^{\left(-\alpha^2/2+i\beta\right)\sigma^2}\right| \\
&\le \limsup_{n\to\infty} \left|\mathbb{E}e^{i\alpha S_n' + i\beta G_n'^2} - \mathbb{E}e^{\left(-\alpha^2/2+i\beta\right)\sigma_\lambda^2}\right| + 2\limsup_{n\to\infty} P(G_n^2 > \lambda) + 2P(\sigma^2 > \lambda) \\
&\le 4\epsilon.
\end{aligned}$$

It follows that $\mathbb{E}e^{i\alpha S_n + i\beta G_n^2} \to \mathbb{E}e^{(-\alpha^2/2+i\beta)\sigma^2}$, which implies (2.7). This completes the proof of Theorem 2.1. ■

3.2. Proof of Theorem 2.2.

We may write $S_{1n} = \sum_{k=1}^n x_{n,k}\epsilon_{k+1}$ with

$$x_{n,k} = \left(\frac{c_n}{n}\right)^{1/2} f(c_n y_{n,k}) = f_n(\eta_1, ..., \eta_k; \xi_1, \xi_2, ...),$$

where $\eta_j = \nu_j, 1 \le j \le n$, and $\xi_j = \nu_{-j+1}, j \ge 1$. It suffices to check that η_j, ϵ_j, and $x_{n,j}$ satisfy Assumptions 1–2 and 4. In fact, by Assumption 6, it is readily seen that $\{\eta_j, \epsilon_j, \mathcal{F}_j\}_{j\ge1}$ forms a martingale difference satisfying Assumption 1. It follows from Corollary 2.2 of Wang and Phillips (2009a) that, for any $0 \le \gamma' \le 1+\gamma/2$,

$$\frac{c_n}{n}\sum_{k=1}^n |f(c_n y_{n,k})|^{1+\gamma'} \to_D \int_{-\infty}^{\infty} |f(x)|^{1+\gamma'} dx \begin{cases} L_{W_{3/2-\nu}}(1,0), & \text{under } \textbf{Condition 1} \\ L_W(1,0), & \text{under } \textbf{Condition 2} \end{cases} \tag{3.11}$$

whenever $c_n \to \infty$ and $c_n/n \to 0$. Due to (3.11), simple calculations show that

$$\begin{aligned}
&\max_{1\le k\le n} |x_{n,k}| \le \left[\left(\frac{c_n}{n}\right)^{1+\gamma/4} \sum_{k=1}^n |f(c_n y_{n,k})|^{2+\gamma/2}\right]^{2/(4+\gamma)} = o_P(1), \\
&\frac{1}{\sqrt{n}}\sum_{k=1}^n |x_{n,k}| \le c_n^{-1/2}\frac{c_n}{n}\sum_{k=1}^n |f(c_n y_{n,k})| = o_P(1),
\end{aligned}$$

which implies Assumption 2 by Remark 3. Finally, Assumption 4 follows from Proposition 2.1. Combining all these facts, we prove (2.16) by Theorem 2.1. ■

3.3. Proof of Theorem 2.3.

It follows from Theorem 2.2 with $f(x) = K(x)$ that

$$\left\{S_{2n}, G_{2n}^2\right\} \to_D \left\{\tau\, \mathcal{L}_W^{1/2} N, \tau^2 \mathcal{L}_W\right\}, \tag{3.12}$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

whenever $c_n/n \to 0$ and $c_n \to \infty$, where $c_n = d_n/h$. This, together with (2.20) with $f_1(y) = K^2(y)$ and $f_2(y) = K(y)$, implies that

$$\left(\frac{d_n}{nh}\right)^{-1/2} \frac{\sum_{t=1}^n \epsilon_{t+1} K_h(y_t - x)}{\sum_{t=1}^n K_h(y_t - x)} = \frac{S_{2n}}{G_{2,n}^2} \frac{\sum_{t=1}^n K_h^2(y_t - x)}{\sum_{t=1}^n K_h(y_t - x)} \to_D \tau N \mathcal{L}_W^{-1/2}, \tag{3.13}$$

where we have used Assumption 8. On the other hand, it is readily seen from Assumption 9 and (2.20) with $f_1(y) = K(y) m_1(y, x)$ and $f_2(y) = K(y)$ that

$$\left(\frac{d_n}{nh}\right)^{-1/2} \frac{\sum_{t=1}^n \left|m(y_t) - m(x)\right| K_h(y_t - x)}{\sum_{t=1}^n K_h(y_t - x)}$$
$$\leq \left(\frac{d_n}{nh}\right)^{-1/2} h^{\gamma} \frac{\sum_{t=1}^n m_1\left[(y_t - x)/h, x\right] K\left[(y_t - x)/h\right]}{\sum_{t=1}^n K\left[(y_t - x)/h\right]} = o_P(1),$$

since $nh^{2\gamma+1}/d_n \to \infty$. Taking these estimates into (2.26), we prove (2.28). The proof of (2.29) is similar, and hence the details are omitted. ■

3.4. Proof of Theorem 2.4.

By virtue of (2.26) and (3.13), to establish (2.30), it suffices to show that, under Assumptions 8* and 9*, for any h satisfying $nh^{1+2(p+1)}/d_n \to 0$,

$$\Lambda_{1n} := \frac{\sum_{t=1}^n \left[m(y_t) - m(x)\right] K_h(y_t - x)}{\sum_{t=1}^n K_h(y_t - x)}$$
$$= \frac{h^p m^{(p)}(x)}{p!} \int_{-\infty}^{\infty} y^p K(y) dy + o_P\left[\left(\frac{d_n}{nh}\right)^{1/2}\right]. \tag{3.14}$$

This is exactly the same as in the proof of Theorem 2.2 of Wang and Phillips (2011). We only provide a outline. The numerator of Λ_{1n} involves

$$\sum_{t=1}^n \{m(y_t) - m(x)\} K\left(\frac{y_t - x}{h}\right) = \sum_{j=1}^{p+1} I_j, \tag{3.15}$$

where

$$I_j = \frac{m^{(j)}(x)}{j!} \sum_{t=1}^n (y_t - x)^j K\left(\frac{y_t - x}{h}\right), \qquad j = 1, 2, ..., p,$$
$$I_{p+1} = \sum_{t=1}^n \left\{m(y_t) - \sum_{j=0}^p \frac{m^{(j)}(x)}{j!} (y_t - x)^j\right\} K\left(\frac{y_t - x}{h}\right).$$

Write $H_j(x) = x^j K(x)$, $j = 1, 2..., p$. Recall that $K(x)$ has compact support with twice continuous differentials. For $j = 1, ..., p-1$, $H_j(x)$ are twice continuously

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

differentiable satisfying $\int_{-\infty}^{\infty}\left[|H_j'(x)|+|H_j''(x)|\right]dx<\infty$. Hence, $\widehat{H}_j(t), j=1,...,p-1$, is integrable, where $\widehat{H}_j(t)=\int_{-\infty}^{\infty}e^{itx}H_j(x)dx$. See Proposition 18.1.2 of Gasquet and Witomski (1999), for instance. Furthermore, for $j=1,...,p-1$, $|\widehat{H}_j(t)|\le C\min\{|t|,1\}$ as $\int H_j(x)dx=0$. By virtue of these facts, together with $\int K(x)dx=1$ and $\int H_p(x)dx\neq 0$, it is readily seen from (2.18) that

$$\frac{h^{-j}\left(\frac{d_n}{nh}\right)^{-1/2}I_j}{\sum_{t=1}^n K\left(\frac{y_t-x}{h}\right)}\to_D \sigma_j N\mathcal{L}_W^{-1/2},\quad j=1,2,...,p-1, \tag{3.16}$$

where $\sigma_j^2=\left[\frac{m^{(j)}(x)}{j!}\right]^2\int H_j^2(x)dx$, and

$$\frac{h^{-p}I_p}{\sum_{t=1}^n K\left(\frac{y_t-x}{h}\right)}\to_P \frac{m^{(p)}(x)}{p!}\int H_p(x)dx. \tag{3.17}$$

On the other hand, by noting $\lim_{h\to 0}\sup_{y\in\Omega}|m^{(p+1)}(yh+x)|\le C$ by Assumption 9*, Taylor expansion yields

$$|I_{p+1}|\le C\sum_{t=1}^n|y_t-x|^{p+1}K\left(\frac{y_t-x}{h}\right),$$

and hence

$$\frac{h^{-(p+1)}|I_{p+1}|}{\sum_{t=1}^n K\left(\frac{y_t-x}{h}\right)}\le C\frac{\sum_{t=1}^n H_{p+1}\left(\frac{y_t-x}{h}\right)}{\sum_{t=1}^n K\left(\frac{y_t-x}{h}\right)}\to_P C\int H_{p+1}(x)dx, \tag{3.18}$$

where $H_{p+1}(x)=|x|^{p+1}K(x)$. Combining (3.15)–(3.18), simple calculation shows that

$$\begin{aligned}\left(\frac{d_n}{nh}\right)^{-1/2}&\left[\Lambda_{1n}-\frac{h^p m^{(p)}(x)}{p!}\int_{-\infty}^{\infty}y^pK(y)dy\right]\\ &\le\frac{\left(\frac{d_n}{nh}\right)^{-1/2}}{\sum_{t=1}^n K\left(\frac{y_t-x}{h}\right)}\sum_{j=1}^{p-1}|I_j|+\frac{\left(\frac{d_n}{nh}\right)^{-1/2}|I_{p+1}|}{\sum_{t=1}^n K\left(\frac{y_t-x}{h}\right)}\\ &=O_P\left[h+\left(\frac{d_n}{nh}\right)^{-1/2}h^{p+1}\right]=o_P(1),\end{aligned}$$

whenever $nh^{1+2(p+1)}/d_n\to 0$. This proves (3.14). The proof of (2.31) is similar, and hence the details are omitted. The proof of Theorem 2.3 is now complete. ■

3.5. Proof of Proposition 3.1.

We start with some preliminaries. Recall that $\max_{1\le k\le n}|x_{n,k}|=o_P(1)$. There exists a sequence of constants $A_n>0$ such that

$$A_n\to\infty,\quad A_n/\sqrt{n}\to 0\quad\text{and}\quad A_n\max_{1\le k\le n}|x_{n,k}|=o_P(1). \tag{3.19}$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

Write $Y_{nm} = \beta^*_{1m}\eta_{m+1} + \alpha\, x^*_{n,m}\epsilon_{m+1}$ and $\widehat{\Omega}_{nm} = R - \Omega_{nm}$, where

$$\Omega_{nm} = \{|\eta_{m+1}| \le A_n \quad \text{and} \quad |\epsilon_{m+1}| \le A_n\}.$$

As $|\beta^*_{1m}| \le C/\sqrt{n}$, $Y^2_{nm} \le C\left(n^{-1} + x^{*2}_{n,m}\right)(\eta^2_{m+1} + \epsilon^2_{m+1})$ and, on Ω_{nm}, $|Y_{nm}| \le C A_n(n^{-1/2} + \max_{1\le k\le n}|x_{n,k}|)$, it is readily seen that

$$\mathbb{E}\left[|Y_{nm}|^3 I(\Omega_{nm}) \mid \mathcal{F}_{n,m}\right] \le C(n^{-1} + x^{*2}_{n,m})\, R_{1nm}, \tag{3.20}$$

$$\mathbb{E}\left[Y^2_{nm} I(\widehat{\Omega}_{nm}) \mid \mathcal{F}_{n,m}\right] \le C(n^{-1} + x^{*2}_{n,m})\, R_{2nm}, \tag{3.21}$$

where

$$R_{1nm} = \left(A_n n^{-1/2} + A_n \max_{1\le k\le n}|x_{n,k}|\right)\left(\mathbb{E}\left[\eta^2_{m+1} \mid \mathcal{F}_{n,m}\right] + \mathbb{E}\left[\epsilon^2_{m+1} \mid \mathcal{F}_{n,m}\right]\right),$$
$$R_{2nm} = \mathbb{E}\left[\left(\eta^2_{m+1} + \epsilon^2_{m+1}\right) I\left(\widehat{\Omega}_{nm}\right) \mid \mathcal{F}_{n,m}\right].$$

It follows from (3.20)–(3.21) and the Taylor's expansion of e^{ix}: $|e^{ix} - 1 - ix| \le x^2/2$ on $\widehat{\Omega}_{nm}$ and $|e^{ix} - 1 - ix + x^2/2| \le |x|^3/6$ on Ω_{nm}, that

$$\begin{aligned}
&\mathbb{E}\left\{(e^{iY_{nm}} - 1 - iY_{nm}) \mid \mathcal{F}_{n,m}\right\} \\
&\quad = -\frac{1}{2}\mathbb{E}\left[Y^2_{nm} I(\Omega_{nm}) \mid \mathcal{F}_{n,m}\right] + \mathbb{E}\left\{(e^{iY_{nm}} - 1 - iY_{nm}) I(\widehat{\Omega}_{nm}) \mid \mathcal{F}_{n,m}\right\} \\
&\qquad + \mathbb{E}\left\{\left(e^{iY_{nm}} - 1 - iY_{nm} + \frac{1}{2}Y^2_{nm}\right) I(\Omega_{nm}) \mid \mathcal{F}_{n,m}\right\} \\
&\quad = -\frac{1}{2}\mathbb{E}\left[Y^2_{nm} \mid \mathcal{F}_{n,m}\right] + R_{3nm},
\end{aligned} \tag{3.22}$$

where

$$\begin{aligned}
|R_{3nm}| &\le \mathbb{E}\left[Y^2_{nm} I\left(\widehat{\Omega}_{nm}\right) \mid \mathcal{F}_{n,m}\right] + \mathbb{E}\left[|Y_{nm}|^3 I(\Omega_{nm}) \mid \mathcal{F}_{n,m}\right] \\
&\le C\left(n^{-1} + x^{*2}_{n,m}\right)(R_{1nm} + R_{2nm}).
\end{aligned}$$

Due to (2.2) and (3.19), $\max_{1\le m\le n}(R_{1nm} + R_{2nm}) = o_P(1)$. Consequently, we have

$$\sum_{m=1}^{n} |R_{3nm}| \le C \max_{1\le m\le n}(R_{1nm} + R_{2nm}) \sum_{m=1}^{n}\left(n^{-1} + x^{*2}_{n,m}\right) = o_P(1). \tag{3.23}$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

We are now ready to prove Proposition 3.1. Recalling $\mathbb{E}\big(Y_{nm} \mid \mathcal{F}_{n,m}\big) = 0$ by Assumption 1, it follows from the definition of β^*_{1m}, (3.22) and (3.23) that

$$\begin{aligned}
\Gamma_n &= \sum_{m=1}^{n} \mathbb{E}\Big\{\Big(e^{iY_{nm}} - 1 - iY_{nm}\Big) \mid \mathcal{F}_{n,m}\Big\} \\
&= -\frac{1}{2}\sum_{m=1}^{n} \mathbb{E}\Big\{\big[\beta^*_{1m}\eta_{m+1} + \alpha x^*_{n,m}\epsilon_{m+1}\big]^2 \mid \mathcal{F}_{n,m}\Big\} + o_P(1) \\
&= -\frac{1}{2}\sum_{m=1}^{n} \beta^{*2}_{1m}\mathbb{E}\Big(\eta^2_{m+1} \mid \mathcal{F}_{n,m}\Big) - \frac{\alpha^2}{2}\sum_{m=1}^{n} x^{*2}_{n,m}\mathbb{E}\Big(\epsilon^2_{m+1} \mid \mathcal{F}_{n,m}\Big) \\
&\quad - \sum_{m=1}^{n} \alpha\,\beta^*_{1m}\,x^*_{n,m}\,\mathbb{E}\big\{(\eta_{m+1}\epsilon_{m+1}) \mid \mathcal{F}_{n,m}\big\} + o_P(1) \\
&= -\frac{1}{2n}\sum_{m=1}^{N_1} \beta^2_{1m}([nu_{1m}] - [nu_{1,m-1}]) - \frac{\alpha^2}{2}\sum_{m=1}^{n} x^{*2}_{n,m} + o_P(1) \\
&= -\frac{1}{2}\sum_{m=1}^{N_1} \beta^2_{1m}(u_{1m} - u_{1,m-1}) - \frac{\alpha^2}{2}\sum_{m=1}^{n} x^{*2}_{n,m} + o_P(1), \qquad \textbf{(3.24)}
\end{aligned}$$

where we used the fact that, by (2.3),

$$\begin{aligned}
&\left|\sum_{m=1}^{n} \alpha\,\beta^*_{1m}\,x^*_{n,m}\,\mathbb{E}\big\{(\eta_{m+1}\epsilon_{m+1}) \mid \mathcal{F}_{n,m}\big\}\right| \\
&\quad \le \frac{C}{\sqrt{n}}\sum_{m=1}^{n} \big|x_{n,m}\big|\big|\mathbb{E}\big\{(\eta_{m+1}\epsilon_{m+1}) \mid \mathcal{F}_{n,m}\big\}\big| = o_P(1).
\end{aligned}$$

Due to (3.3) and (3.24), for any $\alpha_j \in R$, $j = 1, 2, 3, 4$, we have

$$\begin{aligned}
&\alpha_1 V_{1n} + \alpha_2 V_{2n} + \alpha_3 G^{*2}_n + \alpha_4\Gamma_n \\
&\quad = \alpha_1 V_{1n} + \alpha_2 V_{2n} + \left(\alpha_3 - \frac{\alpha^2}{2}\right) G^{*2}_n - \frac{1}{2}\sum_{m=1}^{N_1} \beta^2_{1m}(u_{1m} - u_{1,m-1}) + o_P(1) \\
&\quad \to_D \alpha_1 V_1 + \alpha_2 V_2 + \left(\alpha_3 - \frac{\alpha^2}{2}\right)\sigma^2 - \frac{1}{2}\sum_{m=1}^{N_1} \beta^2_{1m}(u_{1m} - u_{1,m-1}) \\
&\quad = \alpha_1 V_1 + \alpha_2 V_2 + \alpha_3\sigma^2 + \alpha_4\,\Gamma.
\end{aligned}$$

This proves (3.5). Recall that $\mathbb{E}\big(\epsilon^2_{k+1} \mid \mathcal{F}_{n,k}\big) + \mathbb{E}\big(\eta^2_{k+1} \mid \mathcal{F}_{n,k}\big) \le C$, a.s., and note that

$$\Big|\mathbb{E}\Big\{(e^{iY_{nm}} - 1) \mid \mathcal{F}_{n,m}\Big\}\Big| \le \frac{1}{2}\mathbb{E}\Big(Y^2_{nm} \mid \mathcal{F}_{n,m}\Big) \le C\,\Big(n^{-1} + x^{*2}_{n,m}\Big), \quad \text{a.s.} \qquad \textbf{(3.25)}$$

We have

$$|\Gamma_n| \le C\left(1 + \sum_{m=1}^{n} x^{*2}_{n,m}\right) \le C(1 + 2\lambda), \quad \text{a.s.}$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

The uniform integrability of $e^{|\Gamma_n|}$ is obvious. The proof of Proposition 3.1 is now complete. ■

3.6. Proof of Proposition 3.2.

Write $Y_{nk} = \beta^*_{1k}\eta_{k+1} + \alpha\, x^*_{n,m}\epsilon_{m+1}$ as in the proof of Proposition 3.1, $Z_{nk} = \mathbb{E}\left[\left(e^{i\,Y_{nk}} - 1\right) \mid \mathcal{F}_{n,k}\right]$ and let $\mathbb{E}_1 X = \mathbb{E}(X \mid \mathcal{F}_{n,1})$. Recalling

$$V_{1n} = \sum_{k=1}^{N_1} \beta_{1k}\left[W_n(u_{1k}) - W_n(u_{1,k-1})\right] = \sum_{k=1}^{n} \beta^*_{1k}\eta_{k+1}$$

due to the definitions of $W_n(t)$ and β^*_{1k}, we have that

$$\begin{aligned}
I_n &= \mathbb{E}\left|\mathbb{E}_1 e^{i\sum_{k=1}^n Y_{nk} - \sum_{k=1}^n Z_{nk}} - 1\right| \\
&\le \mathbb{E}\left|\mathbb{E}_1\left[e^{i\sum_{k=1}^{n-1} Y_{nk} - \sum_{k=1}^n Z_{nk}}\left(e^{i\,Y_{nn}} - e^{Z_{nn}}\right)\right]\right| + \mathbb{E}\left|\mathbb{E}_1 e^{i\sum_{k=1}^{n-1} Y_{nk} - \sum_{k=1}^{n-1} Z_{nk}} - 1\right| \\
&\le \cdots \le \\
&\le \sum_{m=2}^{n} \mathbb{E}\left|\mathbb{E}_1\left[e^{i\sum_{k=1}^{m-1} Y_{nk} - \sum_{k=1}^m Z_{nk}}\left(e^{i\,Y_{nm}} - e^{Z_{nm}}\right)\right]\right| + \mathbb{E}\left|\mathbb{E}_1 e^{i\,Y_{n1} - Z_{n1}} - 1\right| \\
&:= I_{1n} + I_{2n}. \qquad \textbf{(3.26)}
\end{aligned}$$

Recall (3.25). It follows from $\max_{1\le m\le n} x^{*2}_{n,m} = o_P(1)$ and $\max_{1\le m\le n} x^{*2}_{n,m} \le 2\lambda$ that

$$\begin{aligned}
\Delta_{1n} &:= \left|\mathbb{E}_1 e^{i\,Y_{n1} - Z_{n1}} - 1\right| \\
&= \left|\mathbb{E}_1 \exp\left\{i\,Y_{n1} - \mathbb{E}\left[\left(e^{i\,Y_{n1}} - 1\right) \mid \mathcal{F}_{n,1}\right]\right\} - 1\right| \\
&\le \left|e^{-\mathbb{E}\left[(e^{i\,Y_{n1}} - 1) \mid \mathcal{F}_{n,1}\right]} - 1\right| + \left|\mathbb{E}\left[\left(e^{i\,Y_{n1}} - 1\right) \mid \mathcal{F}_{n,1}\right]\right| \\
&\le C\left(n^{-1} + x^{*2}_{n,1}\right) e^{C(n^{-1} + x^{*2}_{n,1})} = o_P(1),
\end{aligned}$$

and Δ_{1n} is uniformly integrable as $\Delta_{1n} \le C(1+\lambda)e^{\lambda}$. This implies that $I_{2n} = \mathbb{E}\Delta_{1n} \to 0$, as $n \to \infty$.

To consider I_{1n}, write

$$u_{n,m} = e^{i\sum_{k=1}^{m-1} Y_{nk} - \sum_{k=1}^m \mathbb{E}\left[(e^{i\,Y_{nk}} - 1) \mid \mathcal{F}_{n,k}\right]},$$

$$v_{n,m} = e^{i\,Y_{nm}} - e^{\mathbb{E}\left[(e^{i\,Y_{nm}} - 1) \mid \mathcal{F}_{n,m}\right]}.$$

Using (3.25) and $\sum_{k=1}^n x^{*2}_{n,k} \le 2\lambda$ again, it follows that

$$|u_{n,m}| \le e^{\sum_{k=1}^m \left|\mathbb{E}\left[(e^{i\,Y_{nk}} - 1) \mid \mathcal{F}_{n,k}\right]\right|} \le e^{C\sum_{k=1}^m (n^{-1} + x^{*2}_{n,k})} \le e^{C(1+2\lambda)}. \qquad \textbf{(3.27)}$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

As for $\mathbb{E}(v_{n,m} \mid \mathcal{F}_{n,m})$, by using $|e^x - 1 - x| \le |x|^{(2+\delta)/2} e^{|x|}$ for some $\delta > 0$, we have

$$\begin{aligned}|\mathbb{E}(v_{n,m} \mid \mathcal{F}_{n,m})| &= \left|\mathbb{E}\left\{(e^{i\,Y_{nm}} - 1) \mid \mathcal{F}_{n,m}\right\} + 1 - e^{\mathbb{E}\left[(e^{i\,Y_{nm}} - 1)|\mathcal{F}_{n,m}\right]}\right| \\ &\le \left|\mathbb{E}\left\{(e^{i\,Y_{nm}} - 1) \mid \mathcal{F}_{n,m}\right\}\right|^{(2+\delta)/2} e^{|\mathbb{E}\left[(e^{i\,Y_{nm}} - 1)|\mathcal{F}_{n,m}\right]|} \\ &\le C\, e^{C(1+2\lambda)} \left(n^{-1} + x^{*2}_{n,m}\right)^{(2+\delta)/2}.\end{aligned}$$

Now, by recalling $\mathcal{F}_{n,m} \subseteq \mathcal{F}_{n,m+1}$ for any $n \ge m \ge 1$ and $n \ge 1$, it is readily seen that

$$\begin{aligned}I_{1n} &\le \sum_{m=2}^{n} \mathbb{E}\left[|u_{n,m}|\,|\mathbb{E}(v_{n,m} \mid \mathcal{F}_{n,m})|\right] \\ &\le C\, e^{2C(1+2\lambda)}\, \mathbb{E} \sum_{m=2}^{n} \left(n^{-1} + x^{*2}_{n,m}\right)^{(2+\delta)/2} \to 0,\end{aligned}$$

since

$$\begin{aligned}\Delta_{2n} &:= \sum_{m=2}^{n} \left(n^{-1} + x^{*2}_{n,m}\right)^{(2+\delta)/2} \\ &\le \left(n^{-\delta/2} + \max_{1\le m\le n} |x^{*}_{n,m}|^{\delta}\right)\left(1 + \sum_{m=1}^{n} x^{*2}_{n,m}\right) = o_P(1),\end{aligned}$$

and Δ_{2n} is uniformly integrable by noting $\Delta_{2n} \le C(1 + \lambda^{1+\delta/2})$. Taking these facts into (3.26), we prove (3.6).

The proof of (3.7) is simple. Indeed, it follows from (3.6), V_{2n} is $\mathcal{F}_{n,1}$-measurable and and $\mathbb{E}e^{i\,V_{2n}} \to \mathbb{E}e^{i\,V_2}$ due to (3.5) that

$$\begin{aligned}&\left|\mathbb{E}e^{i\,\alpha S^{*}_{n} + i\,V_{1n} + i\,V_{2n} - \Gamma_n} - \mathbb{E}e^{i\,V_2}\right| \\ &\quad\le \mathbb{E}\left|\mathbb{E}\left(e^{i\,\alpha S^{*}_{n} + i\,V_{1n} - \Gamma_n} \mid \mathcal{F}_{n,1}\right) - 1\right| + \left|\mathbb{E}e^{i\,V_{2n}} - \mathbb{E}e^{i\,V_2}\right| \to 0.\end{aligned}$$

The proof of Proposition 3.2 is now complete. ■

3.7. Proof of (2.10).

Since $X_t = B[g^2_t(W, \Xi)]$ is continuous in t, using Theorem 1.2 of Aldous (1989), it suffices to show that, for any $\alpha_1, ..., \alpha_m \in \mathbf{R}$,

$$\begin{aligned}\alpha_1\, S_{[nt_1]} + \cdots + \alpha_m\, S_{[nt_m]} &\to_D \alpha_1\, B\left[g^2_{t_1}(W, \Xi)\right] + \cdots + \alpha_m B\left[g^2_{t_m}(W, \Xi)\right] \\ &=_D N\,\sigma, \qquad \textbf{(3.28)}\end{aligned}$$

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

where N is a standard normal variate independent of σ^2, given by

$$\sigma^2 = (\alpha_1 + \cdots + \alpha_m)^2 g_{t_1}^2(W, \Xi) + (\alpha_2 + \cdots + \alpha_m)^2 \left[g_{t_2}^2(W, \Xi) - g_{t_1}^2(W, \Xi) \right] + \cdots + \alpha_m^2 \left[g_{t_m}^2(W, \Xi) - g_{t_{m-1}}^2(W, \Xi) \right].$$

The proof of (3.28) follows from an application of Theorem 2.1. Indeed, by noting

$$\alpha_1 S_{[nt_1]} + \cdots + \alpha_m S_{[nt_m]} = \sum_{k=1}^{[nt_m]} \beta_k x_{n,k} \epsilon_{n,k+1},$$

where

$$\beta_k = \begin{cases} \alpha_1 + \cdots + \alpha_m, \text{ if } 1 \le k \le [nt_1] \\ \alpha_2 + \cdots + \alpha_m, \text{ if } [nt_1] + 1 \le k \le [nt_2] \\ \cdots\cdots \\ \alpha_m, \qquad\qquad \text{if } [nt_{m-1}] + 1 \le k \le [nt_m], \end{cases}$$

routine calculations show Assumptions 1 and 2 hold with $x_{n,k}$ being replaced by $\beta_k x_{n,k}$. Furthermore, by virtue of (2.9), we have

$$\left\{ W_n(t),\ \Xi_n(t),\ \sum_{k=1}^{[nt_m]} \beta_k^2 x_{n,k}^2 \right\} \Rightarrow \left\{ W(t),\ \Xi(t),\ \sigma^2 \right\},$$

due to

$$\sum_{k=1}^{[nt_m]} \beta_k^2 x_{n,k}^2 = (\alpha_1 + \cdots + \alpha_m)^2 G_{[nt_1]}^2 + (\alpha_2 + \cdots + \alpha_m)^2 \left(G_{[nt_2]}^2 - G_{[nt_1]}^2 \right) + \cdots + \alpha_m^2 \left(G_{[nt_m]}^2 - G_{[nt_{m-1}]}^2 \right),$$

i.e., Assumption 4 holds with G_n^2 being replaced by $\sum_{k=1}^{[nt_m]} \beta_k^2 x_{n,k}^2$. This proves (3.28) by Theorem 2.1. ■

NOTE

1. The result is only for $x = 0$, but it is not difficult to see that the result holds true for any fixed $x \in \mathbf{R}$.

REFERENCES

Aldous, D. (1989) Stopping times and tightness. II. *Annals of Probability* 17, 586–595.

Billingsley, P. (1968) *Convergence of Probability Measures.* Wiley.

Borovskikh, Y.V. & V.S. Korolyuk (1997) *Martingale Approximation.* VSP.

Cai, Z., Q. Li, & J.Y. Park (2009) Functional-coefficient models for non-stationary time series data. *Journal of Econometrics* 148, 101–113.

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

Choi, I. & P. Saikkonen (2010) Tests for nonlinear cointegration. *Econometric Theory* 26, 682–709.
Csörgö, M. & P. Révész (1981) *Strong Approximations in Probability and Statistics.* Probability and Mathematical Statistics. Academic Press.
Gao, J.K. King, Z. Lu, & D. Tjøstheim (2009a) Non-parametric specification testing for nonlinear time series with nonstationarity. *Econometric Theory* 25, 1869–1892.
Gao, J., K. King, Z. Lu, & Tjøstheim D. (2009b) Specification testing in nonlinear and non-stationary time series autoregression. *Annals of Statistics.* 37, 3893–3928.
Gasquet, C. & P. Witomski (1999) *Fourier Analysis and Applications*. Springer.
Geman, D. & J. Horowitz (1980) Occupation densities. *Annals of Probability* 8, 1–67.
Guerre, E. (2004) Design-adaptive pointwise non-parametric regression estimation for recurrent Markov time series. Unpublished manuscript.
Hall, P. (1977) Martingale invariance principles. *Annals of Probability* 5, 875–887.
Hall, P. & C.C. Heyde (1980) *Martingale limit theory and its application*. Probability and Mathematical Statistics. Academic Press.
Hansen, B.E. (1992) Convergence to stochastic integrals for dependent heterogeneous processes. *Econometric Theory* 8, 489–500.
Ibragimov, R. & P.C.B. Phillips (2008) Regression asymptotics using martingale convergence methods. *Econometric Theory* 24(4) 888–947.
Jacod, J. & A.N. Shiryaev (2003) *Limit theorems for stochastic processes*, 2nd ed. Springer-Verlag.
Jeganathan, P. (1982) A solution of the martingale central limit problem, Part I. *Sankhya* 44, 299–318.
Jeganathan, P. (2006) Convergence in distribution of row sum processes to mixtures of additive processes. Unpublished manuscript.
Karlsen, H.A., T. Myklebust, & D. Tjøstheim (2007) Non-parametric estimation in a nonlinear cointegration model. *Annals Statistics* 35, 252–299.
Karlsen, H.A. & D. Tjøstheim (2001) Non-parametric estimation in null recurrent time series. *Annals Statistics* 29, 372–416.
Kasahara, Y. & M. Maejima (1988) Weighted sums of i.i.d. random variables attracted to integrals of stable processes. *Probability Theory and Related Fields* 78, 75–96.
Kasparis, I. (2008) Detection on functional form Misspecification in cointegrating relations. *Econometric Theory* 24, 1373–1403.
Kasparis, I. (2011) Functional form misspecification in regressions with a unit root. *Econometric Theory* 27, 285–311.
Kasparis, I. & P.C.B. Phillips (2009) Dynamic Misspecification in Non-parametric Cointegrating Regression. Working paper, Yale University.
Kurtz, T.G. & P. Protter (1991) Weak limit theorems for stochastic integrals and stochastic differential equations. *Annals of Probability* 19, 1035–1070.
Marmer, V. (2008) Nonlinearity, nonstationarity, and spurious forecasts. *Journal of Econometrics* 142, 1–27.
Park, J.Y. & P.C.B. Phillips (1999) Asymptotics for nonlinear transformation of integrated time series. *Econometric Theory* 15, 269–298.
Park, J.Y. & P.C.B. Phillips (2001) Nonlinear regressions with integrated time series. *Econometrica* 69, 117–161.
Phillips, P.C.B. & J.Y. Park (1998) Non-stationary Density Estimation and Kernel Autoregression. Cowles Foundation Discussion Paper No. 1181.
Prigent, J-L. (2003) *Weak Convergence of Financial Markets.* Springer Finance. Springer-Verlag.
Schienle, M. (2008) Non-parametric Non-stationary Regression. Ph.D. Thesis, University of Mannheim.
Stein, E.M. & R. Shakarchi (2005) *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton University Press.
Wang, Q., Y.-X. Lin, & C.M. Gulati (2003) Asymptotics for general fractionally integrated processes with applications to unit root tests. *Econometric Theory* 19, 143–164.

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press

Wang, Q. & P.C.B. Phillips (2009a) Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory* 25, 710–738.
Wang, Q. & P.C.B. Phillips (2009b) Structural non-parametric cointegrating regression. *Econometrica*. 77, 1901–1948.
Wang, Q. & P.C.B. Phillips (2011) Asymptotic theory for zero energy functionals with non-parametric regression applications. *Econometric Theory* 2, 235–359.
Wang, Q. & P.C.B. Phillips (2012) A specification test for nonlinear non-stationary models. *The Annals of Statistics* 40, 727–758.
Wu, W.B. & M. Woodroofe (2004) Martingale approximations for sums of stationary processes. *Annals of Probability* 32, 1674–1690.

https://doi.org/10.1017/S026646661300039X Published online by Cambridge University Press