ASYMPTOTICS FOR NONLINEAR TRANSFORMATIONS OF INTEGRATED TIME SERIES

BY

JOON Y. PARK
AND
PETER C.B. PHILLIPS

COWLES FOUNDATION PAPER NO. 980

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281
1999

Econometric Theory, **15**, 1999, 269–298. Printed in the United States of America.

ASYMPTOTICS FOR NONLINEAR TRANSFORMATIONS OF INTEGRATED TIME SERIES

JOON Y. PARK
Seoul National University

PETER C.B. PHILLIPS
Yale University

An asymptotic theory for stochastic processes generated from nonlinear transformations of nonstationary integrated time series is developed. Various nonlinear functions of integrated series such as ARIMA time series are studied, and the asymptotic distributions of sample moments of such functions are obtained and analyzed. The transformations considered in the paper include a variety of functions that are used in practical nonlinear statistical analysis. It is shown that their asymptotic theory is quite different from that of integrated processes and stationary time series. When the transformation function is exponentially explosive, for instance, the convergence rate of sample functions is path dependent. In particular, the convergence rate depends not only on the size of the sample but also on the realized sample path. Some brief applications of these asymptotics are given to illustrate the effects of nonlinearly transformed integrated processes on regression. The methods developed in the paper are useful in a project of greater scope concerned with the development of a general theory of nonlinear regression for nonstationary time series.

1. INTRODUCTION

Nonstationary time series arising from autoregressive models with roots on the unit circle have been an intensive subject of recent research. The asymptotic behavior of regression statistics based on integrated time series (those for which one or more of the autoregressive roots are unity) has received the most attention, and a fairly complete theory is now available for linear time series regressions. The resulting limit theory forms the basis of much ongoing empirical econometric work, especially on the subject of unit root testing and cointegration model-

This paper was planned, and the main results were obtained, during the fall of 1995, although we spent a longer time completing the paper. The paper was presented at the Winter Meetings of the Econometric Society in 1998. The authors thank a referee for comments on an earlier version. Park thanks the Department of Economics at Rice University, where he is an Adjunct Professor, for its continuing hospitality. Park acknowledges research support from the Korea Research Foundation and also thanks the Cowles Foundation at Yale University for hospitality during several visits over the period 1993–1998. Phillips thanks the NSF for research support under grant SBR 94-22922. Address correspondence to: Peter C.B. Phillips, Cowles Foundation for Research in Economics, Yale University, Box 208281, New Haven, CT 06520-8281, USA; e-mail: peter.phillips@yale.edu.

 0266-4666/99 $9.50

ing. The main elements of this limit theory as it is needed for linear regression were reviewed in Phillips (1988), and a recent overview of the asymptotic statistical theory on which some of the literature draws was given in Jeganathan (1995).

As in other regression contexts, linear models can be restrictive, and they eliminate many interesting cases of practical importance where there are nonlinear responses to covariates. However, extension of the existing limit theory for integrated processes to nonlinear models is not straightforward. This is because nonlinear functions of integrated processes often depend on fine-grain details of the underlying process, most especially the sojourn time that the process spends in the vicinity of certain points. These details need to be dealt with in the development of a limit theory for the same functions that arise in regression.

The present paper seeks to provide some tools that will be useful in the analysis of time series regressions that involve nonlinear functions of integrated processes. Various nonlinear functions that commonly arise in practical nonlinear statistical analysis are studied. The results show that the limit theory can be very different from that for simple linear and polynomial functions of integrated processes. The case of exponential functions is especially interesting, because here the sojourn time that the process spends in the neighborhood of its extrema determines the asymptotic behavior of the sample function. In consequence, the convergence rate of sample moments of exponential functions of the process is path dependent and relies on extreme sample path realizations of the time series.

2. ASSUMPTIONS AND PRELIMINARY RESULTS

We consider a time series $\{x_t\}$ generated by

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

where the error w_t follows the linear process

$$w_t = \varphi(L)\varepsilon_t = \sum_{k=0}^{\infty} \varphi_k \varepsilon_{t-k}, \tag{2}$$

in which $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (i.i.d.) random variables with mean zero and for which $\varphi(1) \neq 0$. The system (1) is initialized at $t = 0$ with $x_0 = O_p(1)$. One of the following two assumptions will be made throughout the paper.

Assumption 2.1. $\sum_{k=0}^{\infty} k^{1/2}|\varphi_k| < \infty$ and $\mathbf{E}\varepsilon_t^2 < \infty$.

Assumption 2.2.

(a) $\sum_{k=0}^{\infty} k|\varphi_k| < \infty$ and $\mathbf{E}|\varepsilon_t|^p < \infty$ for some $p > 2$.

(b) The distribution of ε_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 $r > 0$.

For simplicity, assume $\varphi(1) = 1$ and $\mathbf{E}\varepsilon_t^2 = 1$. Other values simply have a scaling effect in the subsequent analysis.

Construct the stochastic process

$$W_n^0(r) = \frac{1}{\sqrt{n}} \sum_{t=1}^{[nr]} w_t,$$

which takes values in $D[0,1]$, the set of cadlag functions on the interval $[0,1]$. Phillips and Solo (1992) showed that Assumption 2.1 is sufficient to ensure that W_n^0 converges weakly to a standard linear Brownian motion W on $[0,1]$. In our context, it is more convenient to endow $D[0,1]$ with the uniform topology rather than the usual Skorohod topology (see Billingsley, 1968, pp. 150–152). It then follows from the so-called Skorohod representation theorem (e.g., Pollard, 1984, pp. 71–72) that there exists W_n such that $W_n \overset{d}{=} W_n^0$ in $D[0,1]$, where $\overset{d}{=}$ signifies equivalence in distribution and for which $W_n \to_{\text{a.s.}} W$ uniformly on $[0,1]$. We shall use this representation repeatedly in the proofs in Section 8 of the paper. Using strong approximation methods, specific rates of convergence for $W_n \to W$ can be obtained under moment conditions like those of Assumption 2.2(a). The following result (Csörgő and Horváth, 1993, p. 4; Akonom, 1993) is especially convenient.

LEMMA 2.3. *Let* $n \to \infty$.

(a) *If Assumption* 2.1 *holds, then* $\sup_{0 \le r \le 1} |W_n(r) - W(r)| = o(1)$ a.s.
(b) *If Assumption* 2.2(a) *holds, then* $\sup_{0 \le r \le 1} |W_n(r) - W(r)| = o_p(n^{-(p-2)/2p})$.

Our development relies on the local time, $L(t,s)$, spent by the Brownian motion W at the spatial point s over the interval $[0,t]$. Here $L(t,s)$ is a jointly continuous stochastic process and satisfies the equation (e.g., Chung and Williams, 1990, ch. 7)

$$|W(t) - s| = |W(0) - s| + \int_0^r \operatorname{sgn}(W(t) - s)\, dW(t) + L(t,s),$$

where $\operatorname{sgn}(y) = 1, 0, -1$ as $y > 0, = 0, < 0$, respectively. The following important formula applies, relating temporal integrals of functions of Brownian motion to spatial integrals involving local time.

LEMMA 2.4. (Occupation Times Formula) *Let T be locally integrable. Then*

$$\int_0^t T(W(r))\, dr = \int_{-\infty}^{\infty} T(s) L(t,s)\, ds$$

for all $t \in \mathcal{R}$.

The local time $L(t,s)$ can be interpreted as a spatial occupation density in s for the Brownian motion W. From the continuity of $L(t,\cdot)$, Lemma 2.4 can be applied with $T(x) = 1\{|x - s| < \varepsilon\}$ to give

$$L(t,s) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t 1\{|W(r) - s| < \varepsilon\}\, dr, \tag{3}$$

the representation that explains why $L(\cdot, s)$ is called the local time of W at s.

We define

$$N_n(\nu_n; a, b) = \int_0^1 1\{a \le \nu_n W_n(r) \le b\}\, dr, \tag{4}$$

and similarly

$$N(\nu_n; a, b) = \int_0^1 1\{a \le \nu_n W(r) \le b\}\, dr, \tag{5}$$

where a and b are nonrandom constants and $\nu_n > 0$ for all n. The following useful result is due to Akonom (1993).

LEMMA 2.5. *Let Assumption 2.2 hold. Then as $n \to \infty$*

(a) $\mathbf{E}(N_n(\nu_n;0,\delta) - N_n(\nu_n;k\delta,(k+1)\delta))^2 \le c(\delta/n\nu_n)(1 + (k\delta^2 n \log n/\nu_n^2))$ *for some constant c, and*
(b) $N_n(\nu_n;0,\pi_n) = N(\nu_n;0,\pi_n) + o_p(n^{(2p-1)/3p+\varepsilon})$ *for* $\pi_n \ge \nu_n n^{-2(p+1)/3p}$ *and any* $\varepsilon > 0$.

It follows from (3) and (5) that $(\nu_n/\pi_n)N(\nu_n;0,\pi_n) \to_{\text{a.s.}} L(1,0)$ as $n \to \infty$. And from Lemma 2.5(b), $(\nu_n/\pi_n)N_n(\nu_n;0,\pi_n) = L(1,0) + o_p(1)$ for $\pi_n \ge \nu_n n^{-(2p-1)/3p+\varepsilon}$ with some $\varepsilon > 0$. In this sense, an appropriately defined N_n approximates L for large n. Also $nN_n(\nu_n; a, b)$ is the number of visits of the process $\nu_n W_n(r)$ to the interval $[a, b]$.

3. FUNCTIONS OF NORMALIZED INTEGRATED PROCESSES

We start by investigating the asymptotic behavior of functions of normalized integrated processes. Such functions sometimes arise in models formulated with nonlinear functions of standardized partial sums of stationary time series. Let T be a measurable transformation in $\mathcal{R}$. We will consider *regular* transformations T defined as follows.

DEFINITION 3.1. *A transformation T is said to be* regular *if and only if, on every compact set C, there exist* $\underline{T}_\varepsilon, \overline{T}_\varepsilon$ *and* $\delta_\varepsilon > 0$ *for each* $\varepsilon > 0$ *satisfying*

$$\underline{T}_\varepsilon(x) \le T(y) \le \overline{T}_\varepsilon(x) \tag{6}$$

for all $x, y \in C$ *such that* $|x - y| < \delta_\varepsilon$, *and*

$$\int_C (\overline{T}_\varepsilon - \underline{T}_\varepsilon)(x)\, dx \to 0, \tag{7}$$

as $\varepsilon \to 0$.

The class of regular transformations includes locally bounded monotone functions and continuous functions. For a locally bounded monotone increasing function, for instance, set $\underline{T}_\varepsilon(x) = \underline{T}(x-\varepsilon), \overline{T}_\varepsilon(x) = T(x+\varepsilon)$ and $\delta_\varepsilon = \varepsilon$. Likewise, we set $\underline{T}_\varepsilon(x) = T(x) - \varepsilon$ and $\overline{T}_\varepsilon(x) = T(x) + \varepsilon$ for a continuous function with the usual δ_ε for the ε, δ formulation of uniform continuity. It is easy to see that conditions (6) and (7) are satisfied for such choices. They work for any compact set. It is also clear that finite sums of locally bounded monotone functions (and hence functions that are locally of bounded variation) and piecewise continuous functions are regular.

THEOREM 3.2. *Let Assumption* 2.1 *hold. If T is regular, then*

$$\frac{1}{n}\sum_{t=1}^{n} T\left(\frac{x_t}{\sqrt{n}}\right) \xrightarrow{d} \int_0^1 T(W(r))\,dr,$$

as $n \to \infty$.

Remarks 3.3.

(a) Any regular transformation T is locally integrable. The local integrability of T guarantees that the limiting distribution is well defined. Indeed, T is locally integrable if and only if

$$\Pr\left\{\int_0^t T(W(r))\,dr \quad \text{exists for all } t\right\} = 1$$

(see, e.g., Karatzas and Shreve, 1988, Proposition 6.27, p. 216). We need a stronger condition to ensure that the limiting distribution is invariant across different data generating processes.

(b) Given a transformation T on $\mathcal{R}$, we define a functional Π_T on $D[0,1]$ given by

$$\Pi_T : f \mapsto \int_0^1 T(f(r))\,dr.$$

For T defining a continuous Π_T on $D[0,1]$, the result in Theorem 3.2 follows directly from the continuous mapping theorem (e.g., Billingsley, 1968, Theorem 5.1, p. 30). Uniformly continuous T generate such a functional. If T is continuous, but not uniformly continuous, the corresponding Π_T is assured of being continuous only on $C[0,1]$, a subset of $D[0,1]$. But the continuous mapping theorem still applies, because $C[0,1]$ is of Wiener measure one. Indeed, the proof of Theorem 3.2 shows that, for any regular T, Π_T is continuous on a subset of $D[0,1]$ with Wiener measure one.

(c) The functions

$$T(x) = \log|x| \quad \text{and} \quad T(x) = |x|^\kappa \quad \text{for } -1 < \kappa < 0 \tag{8}$$

are locally integrable and, therefore, $\int_0^1 T(W(r))\,dr$ is well defined for such functions. However, they are *not* regular and Theorem 3.2 does not apply.

To deal with such functions we may proceed as follows. Let T be locally integrable with a pole or logarithmic type of discontinuity at a certain point, say, zero. Define

$$T_n(x) = T(x)1\{|x| \geq c_n\} + T(c_n)1\{0 < x < c_n\} + T(-c_n)1\{-c_n < x < 0\}. \tag{9}$$

Similar modifications can be made for transformations with discontinuities at points other than zero.

THEOREM 3.4. *Let T be locally integrable. Suppose for a sequence $\{c_n\}$ such that $c_n \to 0$ and $c_n \geq n^{-2(p+1)/3p}$,*

$$|T(x) - T(y)| \leq \nu(c_n)|x - y|,$$

with $\nu(c_n) = O(n^{(p-2)/2p})$ for all $x, y \in \{z \| z| \geq c_n\}$, and $T(\pm c_n) = O(n^{(2p-1)/3p+\varepsilon})$ for some $\varepsilon > 0$. If Assumption 2.2 holds, then

$$\frac{1}{n}\sum_{t=1}^{n} T_n\left(\frac{x_t}{\sqrt{n}}\right) \xrightarrow{d} \int_0^1 T(W(r))\,dr,$$

as $n \to \infty$.

Remarks 3.5.

(a) The conditions in Theorem 3.4 require that the function T be Lipschitz continuous on $\{x : |x| \geq c_n\}$. Also, the value of the function $T(\pm c_n)$ around the discontinuity point and the Lipschitz constant $\nu(c_n)$ may not grow too quickly with n.
(b) For the logarithmic function $T(x) = \log|x|$, the conditions in Theorem 3.4 are satisfied with $c_n = n^{-\delta}$ for any δ such that $0 < \delta \leq (p-2)/2p$. For the reciprocal function $T(x) = |x|^{\kappa}$ with $-1 < \kappa < 0$, one may choose $c_n = n^{-\delta}$ for $0 < \delta < (p-2)/2p(1-\kappa)$ to show that the result in Theorem 3.4 is applicable.
(c) For any fixed n, T and T_n are identical over any finite set of nonzero points, if we take c_n to be smaller than the minimum of their moduli. Therefore, if $\{x_t\}$ is driven by an error process whose underlying distribution is of the continuous type specified in Assumption 2.2(b), then T and T_n are practically indistinguishable in finite samples.

4. ADDITIVE FUNCTIONALS OF BROWNIAN MOTION

The asymptotic behavior of functions of unnormalized integrated processes can be quite different from the results in the previous section. In particular, the asymptotics depend in a more critical way on the properties of the functions involved. To illustrate the dependencies that arise, we first investigate the asymptotic behavior of additive functionals of Brownian motion given by

$$\int_0^{\lambda t} T(W(r))\,dr,$$

as $\lambda \to \infty$. The results from this section will be applicable in the statistical analysis of the data that are continuously recorded from Brownian motion, or in the development of the asymptotics when the sampling frequency and the time span of the data increase. Applications of this type occur with financial data in econometrics (Phillips, 1987). More directly, the limit behavior of these functionals sheds light on the behavior of nonlinear functions of integrated processes and is thereby useful in the development of an asymptotic theory for regression that involves such nonlinear functions.

Three classes of transformation are explored here: integrable (I) functions, asymptotically homogeneous (H) functions, and explosive (E) functions. These will be referred to respectively as Classes (I), (H), and (E) in the paper and will be denoted by $\mathcal{T}(I)$, $\mathcal{T}(H)$, and $\mathcal{T}(E)$. More explicitly we define these classes as follows.

DEFINITION 4.1. *A transformation T is said to be in Class (I), denoted by $T \in \mathcal{T}(I)$, iff it is integrable.*

DEFINITION 4.2. *A transformation T is said to be in Class (H), denoted by $T \in \mathcal{T}(H)$, iff*

$$T(\lambda x) = \nu(\lambda)H(x) + R(x,\lambda),$$

where H is locally integrable and R is such that

(a) $|R(x,\lambda)| \le a(\lambda)P(x)$, *where* $\limsup_{\lambda\to\infty} a(\lambda)/\nu(\lambda) = 0$ *and P is locally integrable, or*
(b) $|R(x,\lambda)| \le b(\lambda)Q(\lambda x)$, *where* $\limsup_{\lambda\to\infty} b(\lambda)/\nu(\lambda) < \infty$ *and Q is locally integrable and vanishes at infinity, i.e., $Q(x) \to 0$ as $|x| \to \infty$.*

Transformations $T \in \mathcal{T}(H)$ with R satisfying conditions (a) and (b) will be said to belong to $\mathcal{T}(H_1)$ and $\mathcal{T}(H_2)$, respectively.

Remarks 4.3.

(a) If $T \in \mathcal{T}(H)$, T has an asymptotically dominating component that is homogenous. All homogenous functions are of this type and therefore belong to $\mathcal{T}(H)$ as long as they are locally integrable. If T is homogeneous of degree κ, then we have $H = T$ and $\nu(\lambda) = \lambda^\kappa$. Examples of such functions include $T(x) = x^\kappa$ for $\kappa > 0$ and $T(x) = \operatorname{sgn}(x)$.
(b) The finite order polynomial given by $T(x) = x^k + a_1x^{k-1} + \cdots + a_k$ for $k \ge 1$ is in $\mathcal{T}(H_1)$ with $\nu(\lambda) = \lambda^k$ and $H(x) = x^k$. For $a(\lambda) = \lambda^{k-1}|a_1 + a_2/\lambda + \cdots + a_k/\lambda^{k-1}|$ and $P(x) = 1 + |x|^{k-1}$, we may easily show that $|R(x,\lambda)| \le a(\lambda)P(x)$. Clearly, $a(\lambda)/\nu(\lambda) \to 0$ as $\lambda \to \infty$, and P is locally integrable for $k \ge 1$.
(c) The logarithmic function $T(x) = \log|x|$ belongs to $\mathcal{T}(H_1)$, with the homogenous component given by $\nu(\lambda) = \log\lambda$ and $H(x) = 1$. The residual function then becomes $R(x,\lambda) = \log|x|$. To see that it satisfies the preceding conditions, set $a(\lambda) = 1$ and $P(x) = \log|x|$. Iterated logarithmic functions and polynomials in logarithms are also in $\mathcal{T}(H)$, which can be shown similarly.
(d) The distribution function of any random variable belongs to class $\mathcal{T}(H_2)$, with the homogeneous component specified by $\nu(\lambda) = 1$ and $H(x) = 1\{x \ge 0\}$. Clearly, H is locally integrable. If T is such a function, the residual $R(x,\lambda)$ is bounded in modulus by $Q(\lambda x)$, where $Q(x) = T(x)1\{x<0\} + (1 - T(x))1\{x \ge 0\}$. It is easy to see that Q is locally integrable and vanishes at infinity. If, in particular, the underlying random variable has finite expectation, then $Q \in \mathcal{T}(I)$.

DEFINITION 4.4. *A transformation T is said to be in Class (E), denoted by $T \in \mathcal{T}(E)$, iff*

$$T(x) = E(x) + R(x),$$

with E and R satisfying the following conditions.

(a) *E is monotone. If E is increasing (decreasing), then it is positive and differentiable on $\mathcal{R}_+$ ($\mathcal{R}_-$). Furthermore, if we define $G(x) = \log E(x)$ on $\mathcal{R}_+$ ($\mathcal{R}_-$) with derivative $\dot{G}$, then as $\lambda \to \infty$, $\dot{G}(\lambda x) = \nu(\lambda)D(x) + o(\nu(\lambda))$ uniformly in a neighborhood of x, where D is positive (negative) and continuous, and $\lambda\nu(\lambda) \to \infty$.*
(b) *R is given such that for any x and y*

$$\frac{\lambda\nu(\lambda)\bar{R}(\lambda x)}{E(\lambda y)} \to 0$$

as $\lambda \to \infty$, where $\bar{R}(x) = \sup_{y \le |x|}|R(y)|$.

Remarks 4.5.

(a) For $T \in \mathcal{T}(E)$, E denotes the exponential component that is asymptotically dominating. The derivative of the exponent function of E is assumed to be asymptotically homogeneous with base function D and degree of homogeneity ν. If we write $E(x) = \exp(G(x))$, then the condition $\lambda\nu(\lambda) \to \infty$ ensures that G increases on $\mathcal{R}_+$ (or decreases on $\mathcal{R}_-$) faster than the logarithmic function. When there is such an exponential component, all other components with polynomial orders become negligible. They satisfy our conditions for R, as one may easily check.
(b) The conditions for the exponential component E of $T \in \mathcal{T}(E)$ obviously hold for functions like $E(x) = \exp(x^\kappa)$ for $\kappa > 0$, or $E(x) = x^\kappa e^x\{x > 0\}$ for any finite κ. In the former case, we have $\nu(\lambda) = \lambda^{\kappa-1}$ and $D(x) = \kappa x^{\kappa-1}$. For the latter, $\nu(\lambda) = 1$ and $D(x) = 1$.

THEOREM 4.6. *Let $T \in \mathcal{T}(I)$. Then*

$$\frac{1}{\lambda}\int_0^{\lambda^2 t} T(W(r))\,dr \xrightarrow{d} \left(\int_{-\infty}^{\infty} T(s)\,ds\right)L(t,0),$$

as $\lambda \to \infty$.

THEOREM 4.7. *Let $T \in \mathcal{T}(H)$ with $H(\cdot)$ as in Definition 4.2. Then*

$$\frac{1}{\lambda^2\nu(\lambda)}\int_0^{\lambda^2 t} T(W(r))\,dr \xrightarrow{d} \int_{-\infty}^{\infty} H(s)L(t,s)\,ds,$$

as $\lambda \to \infty$.

THEOREM 4.8. *Let $T \in \mathcal{T}(E)$ with ν and D as in Definition 4.4. Then as $\lambda \to \infty$*

$$\frac{\nu(\lambda)}{\lambda T(\sup_{0\le r\le\lambda^2 t} W(r))}\int_0^{\lambda^2 t} T(W(r))\,dr \xrightarrow{d} \frac{1}{D(s_{\max})}L(t,s_{\max}),$$

or

$$\frac{\nu(\lambda)}{\lambda T(\inf_{0\le r\le\lambda^2 t} W(r))}\int_0^{\lambda^2 t} T(W(r))\,dr \xrightarrow{d} \frac{1}{-D(s_{\min})}L(t,s_{\min}),$$

depending upon whether the exponential component E is increasing or decreasing, and where $s_{\max} = \sup_{0\le r\le 1} W(r)$ *and* $s_{\min} = \inf_{0\le r\le 1} W(r)$.

Remarks 4.9.

(a) Theorems 4.6–4.8 reveal that the asymptotic behavior of the three different types of additive functionals of Brownian motion differs in fundamental ways. For integrable functions, only the local time spent by W in the vicinity of the origin matters. This is not so for asymptotically homogenous functions, for which the local time of W at all points contributes to the limit distribution. Finally, the local time that W spends in the neighborhood of one of its extrema completely determines the asymptotic behavior of an explosive function.
(b) The convergence rates for explosive functions are path dependent, i.e., they depend not only on the size of the sample but also on the actual path of the sample by virtue of the fact that $\sup_r W(r)$ and $\inf_r W(r)$ influence the convergence rate.

5. FUNCTIONS OF INTEGRATED PROCESSES

Not surprisingly, the moments of functions of integrated processes asymptotically behave rather like the corresponding additive functionals of Brownian motion. We just need some extra conditions to make their limiting behavior invariant with respect to the underlying data generating processes.

THEOREM 5.1. *Suppose* $T \in \mathcal{T}(I)$ *and Assumption* 2.2 *holds with* $p > 4$. *If T is square integrable and satisfies the Lipschitz condition*

$$|T(x) - T(y)| \le c|x - y|^{\ell}$$

over its support for some constant c and $\ell > 6/(p-2)$, *then*

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} T(x_t) \xrightarrow{d} \left(\int_{-\infty}^{\infty} T(s)\, ds\right) L(1,0),$$

as $n \to \infty$.

Remarks 5.2.

(a) For an indicator function on a bounded set, the result in Theorem 5.1 is applicable as long as $p > 4$. The Lipschitz function with $\ell = 1$ requires, in particular, that $p > 8$.
(b) The collection of transformations for which Theorem 5.1 applies is closed under the operation of finite linear combinations. Thus, the result in Theorem 5.1 holds for any piecewise function for which each piece satisfies the given conditions.

THEOREM 5.3. *Let* $T \in \mathcal{T}(H)$ *with* $H(\cdot)$ *regular. Also, assume that T is either in* $\mathcal{T}(H_1)$ *with P locally bounded or in* $\mathcal{T}(H_2)$ *with Q bounded and vanishing at infinity. If Assumption* 2.1 *holds, then*

$$\frac{1}{n\nu(\sqrt{n})} \sum_{t=1}^{n} T(x_t) \xrightarrow{d} \int_{-\infty}^{\infty} H(s) L(1,s)\, ds,$$

as $n \to \infty$.

Remarks 5.4.

(a) For Theorem 5.3, we only need Assumption 2.1. This is in contrast to Theorems 5.1 and 5.5 for functions in $\mathcal{T}(I)$ and $\mathcal{T}(E)$, where the stronger Assumption 2.2 is invoked.
(b) The result in Theorem 5.3 is applicable to such functions as $T(x) = x^\kappa$ for $\kappa > 0$, $T(x) = \mathrm{sgn}(x)$, $T(x) = x^k + a_1 x^{k-1} + \cdots + a_k$ for $k \geq 1$, and to all "distribution function"-like transformations.

THEOREM 5.5. *Let $T \in \mathcal{T}(E)$ and $\nu(\lambda) = \lambda^m$ with $m < (p-8)/6p$. If Assumption* 2.2 *holds, then as $n \to \infty$*

$$\frac{\nu(\sqrt{n})}{\sqrt{n}T(\max_{1\le t\le n} x_t)} \sum_{t=1}^{n} T(x_t) \xrightarrow{d} \frac{1}{D(s_{\max})} L(1, s_{\max})$$

or

$$\frac{\nu(\sqrt{n})}{\sqrt{n}T(\min_{1\le t\le n} x_t)} \sum_{t=1}^{n} T(x_t) \xrightarrow{d} \frac{1}{-D(s_{\min})} L(1, s_{\min}),$$

depending upon whether the exponential component E is increasing or decreasing.

Remarks 5.6.

(a) The convergence rates are path dependent, as in Theorem 4.8, i.e., they depend upon $\max x_t$ or $\min x_t$, $t = 1, \dots, n$, respectively, for the increasing and decreasing exponential component of the transformation in $\mathcal{T}(E)$.
(b) The result in Theorem 5.5 is applicable for explosive functions such as $x^\kappa \exp(x)\{x > 0\}$, as long as $p > 8$. However, we only allow functions to be mildly explosive. Functions like $T(x) = \exp(x^2)$ are excluded. The asymptotic behaviors of such functions may not be invariant and can be more dependent upon the underlying data generating process.

6. NONLINEAR REGRESSION ILLUSTRATIONS WITH INTEGRATED PROCESSES

In this section, we briefly show how to apply the preceding theory to develop regression asymptotics for models with transformed integrated regressors. Let $\{x_t\}$ be generated by (1) and (2) and consider the regression model

$$y_t = \alpha f(x_t) + u_t \tag{10}$$

for $t = 1, \dots, n$, where α is the regression coefficient, f is a transformation in $\mathcal{R}$, and $\{u_t\}$ are stationary errors. The least squares estimator $\hat{\alpha}_n$ of α in regression (10) is given by

$$\hat{\alpha}_n = \frac{\sum_{t=1}^{n} f(x_t) y_t}{\sum_{t=1}^{n} f^2(x_t)} = \alpha + \frac{\sum_{t=1}^{n} f(x_t) u_t}{\sum_{t=1}^{n} f^2(x_t)}.$$

When f is the identity transform, regression (10) reduces to what is known as (a linear) cointegrating regression. Such regressions have become very popular in time series econometrics following the work of Engle and Granger (1987). However, it is not always clear that the relationship between y_t and x_t is linear, and such considerations lead naturally to models of the form (10) (just as in the case where y_t and x_t are stationary).

Let $\{\mathcal{F}_t\}$ be the natural filtration for $\{u_t\}$ and make the following assumption.

Assumption 6.1.

(a) $\{u_t\}$ is independent of $\{w_t\}$, and
(b) $(u_t, \mathcal{F}_t)$ is a martingale difference sequence with $\mathbf{E}(u_t^2 | \mathcal{F}_{t-1}) = \sigma^2$ for all t, and $\sup_t \mathbf{E}(|u_t|^q | \mathcal{F}_{t-1}) < \infty$ a.s. for some $q > 2$.

Assumption 6.1(a) is stronger than is needed but is made for simplicity to highlight the effect of the nonlinear transformation on the regression asymptotics. As before, we let $\sigma^2 = 1$, because it has only a scaling effect.

The lemma that follows gives the Skorohod embedding of a partial sum and a strong approximation to its quadratic variation as in Phillips and Ploberger (1996). It is useful in the derivation of the regression asymptotics in Theorem 6.3, which follows.

LEMMA 6.2. *Let Assumption* 6.1(b) *hold. Then there exists a probability space supporting a standard linear Brownian motion U and an increasing sequence of stopping times $\{\tau_t\}_{t \geq 0}$ with $\tau_0 = 0$ such that $1/\sqrt{n} \sum_{k=1}^{t} u_k \overset{d}{=} U(\tau_t / n)$ and*

$$\sup_{1 \leq t \leq n} \left| \frac{\tau_t - t}{n^{\delta}} \right| \xrightarrow{a.s.} 0,$$

as $n \to \infty$ for any $\delta > \max(1/2, 2/q)$.

In view of Assumption 6.1(a), we may assume that W and U are independent and defined on a common probability space.

THEOREM 6.3. *Let $T = f^2$ and denote by V a standard linear Brownian motion independent of W. Suppose Assumption* 6.1 *holds.*

(a) *If T satisfies the conditions in Theorem* 5.1, *then as $n \to \infty$*

$$\sqrt[4]{n}(\hat{\alpha}_n - \alpha) \xrightarrow{d} \frac{V(1)}{\left(\int_{-\infty}^{\infty} T(s)\, ds\, L(1,0) \right)^{1/2}}.$$

(b) *If T satisfies the conditions in Theorem 5.3, then as $n \to \infty$*

$$\sqrt{n\nu(\sqrt{n})}(\hat{\alpha}_n - \alpha) \xrightarrow{d} \frac{V(1)}{\left(\int_{-\infty}^{\infty} H(s)L(1,s)\,ds\right)^{1/2}}.$$

(c) *If T satisfies the conditions in Theorem 5.5, then as $n \to \infty$*

$$\left(\frac{\sqrt{n}T\left(\max_{1\le t\le n} x_t\right)}{\nu(\sqrt{n})}\right)^{1/2}(\hat{\alpha}_n - \alpha) \xrightarrow{d} \frac{D(s_{\max})^{1/2}V(1)}{L(1,s_{\max})^{1/2}},$$

or

$$\left(\frac{\sqrt{n}T\left(\min_{1\le t\le n} x_t\right)}{\nu(\sqrt{n})}\right)^{1/2}(\hat{\alpha}_n - \alpha) \xrightarrow{d} \frac{-D(s_{\min})^{1/2}V(1)}{L(1,s_{\min})^{1/2}},$$

depending upon whether the exponential component E is increasing or decreasing.

Theorem 6.3 shows that $\hat{\alpha}_n$ is consistent when the conditions in Theorems 5.1 and 5.5 are met for $T = f^2$. Also, it is consistent if $T = f^2$ satisfies the conditions in Theorem 5.3 with $\lambda^2\nu(\lambda) \to \infty$ as $\lambda \to \infty$. Thus, we may generally expect consistency, in the same way as in other time series regressions under persistent excitation. The limiting distributions are mixed normal, in the same way as for cointegrating regressions (Phillips, 1991). The rate of convergence, however, will vary depending on f. It can be faster than the convergence rate (n) for linear cointegrating regressions, but it can also be slower than the $\sqrt{n}$ rate for stationary regression. When f is explosive, as in the case of exponential functions, the convergence rate for $\hat{\alpha}_n$ is dependent upon the entire sample path of x_t and on the sample size.

Because the sample path of an integrated process typically shows trending behavior, it is interesting to compare (10) with nonlinear regressions on deterministically trending regressions. To be explicit, consider the following two regressions:

$$y_t = \frac{\alpha}{|x_t|^{\beta}} + u_t \tag{11}$$

and

$$y_t = \frac{\alpha}{t^{\beta}} + u_t, \tag{12}$$

where $\beta > 0$ is a known constant and the other notation is defined as in (10). The least squares estimators of α in (11) and (12) are denoted, respectively, by $\hat{\alpha}_n$ and $\tilde{\alpha}_n$. Unlike $\tilde{\alpha}_n$, $\hat{\alpha}_n$ is not properly defined without some modification, because x_t may take values in the neighborhood of zero (or could even be zero with positive probability in the case of discrete innovations w_t) in which case the regression function is singular. Therefore, we follow the convention introduced in (9) and assume that $\hat{\alpha}_n$ is computed from a regression on $x_{nt} = x_t\{|x_t| \geq c_n\} + c_n\{|x_t| < c_n\}$ (in lieu of x_t) with $c_n = n^{-\delta}$ for $0 < \delta < (p-2)/2p(1+2\beta)$. See Remark 3.5(b) for our choice of c_n here. We let Assumption 2.2 hold in the subsequent discussion.

The asymptotic behavior of both $\hat{\alpha}_n$ and $\tilde{\alpha}_n$ is critically dependent upon the value of β. For $0 < \beta < \frac{1}{2}$, both $\hat{\alpha}_n$ and $\tilde{\alpha}$ are consistent and have limiting distributions given, respectively, by

$$n^{(1-\beta)/2}(\hat{\alpha}_n - \alpha) \xrightarrow{d} \left(\int_0^1 \frac{1}{|W(r)|^{2\beta}}\, dr\right)^{-1/2} V(1)$$

and

$$n^{1/2-\beta}(\tilde{\alpha}_n - \alpha) \xrightarrow{d} \left(\int_0^1 \frac{1}{r^{2\beta}\, dr}\right)^{-1/2} V(1).$$

If $\beta > \frac{1}{2}$, however, the asymptotic behavior is very different.

When $\beta = \frac{1}{2}$, $(\log n)^{1/2}(\hat{\alpha}_n - \alpha) \to_d V(1)$ and $\tilde{\alpha}_n$ from regression (12) is therefore consistent. The estimator $\tilde{\alpha}_n$ becomes inconsistent if β exceeds the critical value $\frac{1}{2}$, because $\sum_{t=1}^{n} 1/t^{2\beta} < \infty$ for $\beta > \frac{1}{2}$, and the excitation condition fails to hold. Faulty intuition here might suggest that regression (12) with $\beta = \frac{1}{2}$ is analogous to regression (11) with $\beta = 1$, because $x_t = O_p(\sqrt{t})$. This might lead to the conjecture that $\hat{\alpha}_n$ from regression (11) becomes inconsistent when $\beta > 1$. Interestingly, however, $\hat{\alpha}_n$ from regression (11) is consistent for *all* values of β, including $\beta > 1$, as shown in the following proposition, which establishes the validity of the excitation condition for the regressor in (11) for all β.

PROPOSITION 6.4. *Let Assumption* 2.2 *hold. Then*

$$\sum_{t=1}^{n} |x_t|^{\kappa} \xrightarrow{p} \infty,$$

as $n \to \infty$, *for any* $\kappa \neq -\infty$.

7. CONCLUSION

The examples given in the previous section involve models that are linear in the parameters and nonlinear in the regressor. Such models are obviously very simple examples of regressions that involve nonlinear functions of integrated processes, and our theory therefore provides only a basic extension of cointegrating regression asymptotics even though its methods are quite novel. In spite of their simplicity, however, the models do illustrate some important features of more general nonlinear cointegrating regression problems.

First, it is apparent that the signal emanating from a nonstationary regressor can be substantially altered in strength by nonlinear transformations. Moreover, as the strength of the signal is modified, the corresponding rate of convergence of the regression coefficient is affected. Our simple examples show that nonlinear transformations can decrease the rate of convergence over that of a linear cointegrating regression and also increase this rate. Second, the rate of convergence may in some cases be path dependent, in the sense that the rate itself is stochastic and depends on properties of the process such as its maximum or minimum. Finally, the limit theory in all cases considered turns out to be mixed normal, as in linear cointegrating regressions. Indeed, if a Gaussian likelihood approach were adopted, the likelihood would turn out to be in the locally asymptotically mixed normal class, so that an optimal theory of inference can be developed, as in Jeganathan (1995) and Phillips (1991).

With these new methods in hand, we are ready to undertake the general task of developing a theory of regression for nonlinear functions of nonstationary regressors in which the parameters also enter in a nonlinear fashion. This task is inevitably more complex and of broader scope than what has been completed in this paper. Nevertheless, the results rely intimately on the methods we have introduced here. The results of the broader investigation will be reported by the authors in a later article (Park and Phillips, 1998).

8. PROOFS

Proof of Lemma 2.3. Parts (a) and (b) are, respectively, Theorem 3.4 of Phillips and Solo (1992) and Theorem 3 of Akonom (1993). ■

Proof of Lemma 2.4. See, e.g., Corollary 7.4 of Chung and Williams (1990). ■

Proof of Lemma 2.5. In what follows, let $N_n(a,b) = N_n(\nu_n;a,b)$ to simplify notation. For the proof of part (a), we first deduce from Lemma 4 of Akonom (1993) that

$$\mathbf{E}\left(N_n(0,\delta) - \frac{1}{k}N_n(\delta,(k+1)\delta)\right)^2 \leq c\,\frac{\delta}{n\nu_n}\left(1 + \frac{k\delta^2 n\log n}{\nu_n^2}\right),$$

and similarly

$$\mathbf{E}\left(N_n(k,(k+1)\delta) - \frac{1}{k}N_n(\delta,(k+1)\delta)\right)^2 \leq c\,\frac{\delta}{n\nu_n}\left(1 + \frac{k\delta^2 n\log n}{\nu_n^2}\right),$$

where c is some constant depending only upon the distribution of $\{\varepsilon_t\}$ and $\{\varphi_k\}$. The stated result now follows immediately because

$$\begin{aligned}\mathbf{E}&(N_n(0,\delta) - N_n(k\delta,(k+1)\delta))^2\\ &\leq 2\left(\mathbf{E}\left(N_n(0,\delta) - \frac{1}{k}N_n(\delta,(k+1)\delta)\right)^2\right.\\ &\qquad\left. + \mathbf{E}\left(N_n(k,(k+1)\delta) - \frac{1}{k}N_n(\delta,(k+1)\delta)\right)^2\right).\end{aligned}$$

Part (b) is due to Akonom (1993), Theorem 4. ■

Proof of Theorem 3.2. Assume temporarily that $x_0 = 0$, and, using the Skorohod representation, write

$$\frac{1}{n}\sum_{t=1}^{n} T\left(\frac{x_t}{\sqrt{n}}\right) \overset{d}{=} \int_0^1 T(W_n(r))\,dr.$$

Let $C = [s_{\min} - 1, s_{\max} + 1]$, where $s_{\min}$ and $s_{\max}$ are defined as in Theorem 4.8. As a result of Lemma 2.3(a), we may take n sufficiently large so that $\sup|W_n(r) - W(r)| < \delta_\varepsilon$ for any $\delta_\varepsilon > 0$ and so that both W_n and W are in C a.s. (Note that C is path dependent on W by construction.) Therefore,

$$\underline{T}_\varepsilon(W(r)) \le T(W_n(r)) \le \overline{T}_\varepsilon(W(r)) \tag{13}$$

for large n because of (6). However,

$$\begin{aligned}\int_0^1 (\overline{T}_\varepsilon - \underline{T}_\varepsilon)(W(r))\,dr &= \int_{-\infty}^{\infty} (\overline{T}_\varepsilon - \underline{T}_\varepsilon)(s)L(1,s)\,ds \\ &\le (\sup_s L(1,s)) \int_C (\overline{T}_\varepsilon - \underline{T}_\varepsilon)(x)\,dx \\ &\xrightarrow{\text{a.s.}} 0, \end{aligned} \tag{14}$$

as $\varepsilon \to 0$, due to (7). The stated result now easily follows from (13) and (14). For the case $x_0 \neq 0$, simply replace W_n with $x_0/\sqrt{n} + W_n$ in the preceding proof. ■

Proof of Theorem 3.4. Again, temporarily assume $x_0 = 0$, and write

$$\frac{1}{n}\sum_{t=1}^{n} T_n\left(\frac{x_t}{\sqrt{n}}\right) \overset{d}{=} \int_0^1 T_n(W_n(r))\,dr,$$

as in the proof of Theorem 3.2. We define

$$A_n = \left|\int_0^1 T_n(W_n(r))\,dr - \int_0^1 T_n(W(r))\,dr\right|,$$

$$B_n = \left|\int_0^1 T_n(W(r))\,dr - \int_0^1 T(W(r))\,dr\right|$$

and show

$$\left|\int_0^1 T_n(W_n(r))\,dr - \int_0^1 T(W(r))\,dr\right| \le A_n + B_n = o_p(1)$$

subsequently.

Given the conditions on the orders of $\nu(c_n)$ and $T(\pm c_n)$, we may easily deduce from Lemmas 2.3(b) and 2.5(b), setting $\pi_n/\nu_n = c_n$ in the latter, that

$$A_n \le \nu(c_n)\int_0^1 |W_n(r) - W(r)|\,dr + |T(\pm c_n)| \times \left|\int_0^1 1\{|W_n(r)| < c_n\}\,dr - \int_0^1 1\{|W(r)| < c_n\}\,dr\right| = o_p(1). \tag{15}$$

Therefore, it suffices to show that

$$B_n \le \left|\int_0^1 T(W(r))1\{|W(r)| \ge c_n\} - \int_0^1 T(W(r))\,dr\right| + |T(\pm c_n)|\int_0^1 1\{|W(r)| < c_n\}\,dr = o(1) \quad \text{a.s.} \tag{16}$$

It follows from (3) that

$$T(\pm c_n)\int_0^1 1\{|W(r)| \le c_n\}\,dr = c_n T(\pm c_n)(L(1,0) + o(1)) \xrightarrow{\text{a.s.}} 0,$$

because T is locally integrable and therefore $c_n T(\pm c_n) \to 0$ for $c_n \to 0$. Moreover,

$$\int_0^1 T(W(r))1\{|W(r)| \ge c_n\}\,dr = \int_{-\infty}^{\infty} T(s)1\{|s| \ge c_n\}\,L(1,s)\,ds \xrightarrow{\text{a.s.}} \int_{-\infty}^{\infty} T(s)L(1,s)\,ds = \int_0^1 T(W(r))\,dr,$$

by dominated convergence and repeated applications of Lemma 2.4. Notice that $T(\cdot)\{|\cdot| \ge c_n\} \to T(\cdot)$ pointwise except at zero, which is of Lebesgue measure zero. The stated result now follows from (15) and (16).

When $x_0 \ne 0$, we may define

$$A_n' = \int_0^1 T_n\left(\frac{x_0}{\sqrt{n}} + W_n(r)\right)dr - \int_0^1 T_n\left(\frac{x_0}{\sqrt{n}} + W(r)\right)dr,$$

$$B_n' = \int_0^1 T_n\left(\frac{x_0}{\sqrt{n}} + W(r)\right)dr - \int_0^1 T(W(r))\,dr,$$

instead of A_n and B_n, and the stated result holds in the same way. ∎

Proof of Theorem 4.6. See Proposition 2.2 in Chapter XIII of Revuz and Yor (1994). ∎

Proof of Theorem 4.7. We have

$$\begin{aligned}\frac{1}{\lambda^2\nu(\lambda)}\int_0^{\lambda^2 t} T(W(r))\,dr &= \frac{1}{\nu(\lambda)}\int_0^t T(W(\lambda^2 r))\,dr \\ &\overset{d}{=} \frac{1}{\nu(\lambda)}\int_0^t T(\lambda W(r))\,dr \\ &= \int_0^t H(W(r))\,dr + \frac{1}{\nu(\lambda)}\int_0^t R(W(r),\lambda)\,dr.\end{aligned}$$

Because H is assumed to be locally integrable,

$$\int_0^t H(W(r))\,dr = \int_{-\infty}^{\infty} H(s)L(t,s)\,ds,$$

by Lemma 2.4. Therefore, it suffices to show that

$$\frac{1}{\nu(\lambda)}\int_0^t R(W(r),\lambda)\,dr \xrightarrow{\text{a.s.}} 0$$

to finish the proof.

If $T \in \mathcal{T}(H_1)$, it is immediate that

$$\frac{1}{\nu(\lambda)}\int_0^t |R(W(r),\lambda)|\,dr \leq \frac{a(\lambda)}{\nu(\lambda)}\int_0^t P(W(r))\,dr \xrightarrow{\text{a.s.}} 0,$$

because P is locally integrable. For $T \in \mathcal{T}(H_2)$, we have from Lemma 2.4

$$\begin{aligned}\frac{1}{\nu(\lambda)}\int_0^t |R(W(r),\lambda)|\,dr &\leq \frac{b(\lambda)}{\nu(\lambda)}\int_0^t Q(\lambda W(r))\,dr \\ &= \frac{b(\lambda)}{\nu(\lambda)}\int_{-\infty}^{\infty} Q(\lambda s)L(t,s)\,ds.\end{aligned}$$

Because Q vanishes at infinity, $Q(\lambda s) \to 0$ for all s except $s = 0$, which is of Lebesgue measure zero. We may assume w.l.o.g. that Q is monotone decreasing (increasing) as $x \to \infty$ ($x \to -\infty$), by considering Q_*, $Q_*(x) = \sup_{y \geq |x|} Q(y)$, in place of Q, if necessary. Now, for all $\lambda \geq 1$, $Q(\lambda\,\cdot)$ is bounded by $Q(\cdot)$ which is locally integrable. Because $L(t,\cdot)$ has compact support for any fixed t, we have

$$\int_{-\infty}^{\infty} Q(\lambda s)L(t,s)\,ds \xrightarrow{\text{a.s.}} 0, \tag{17}$$

by dominated convergence. ■

Proof of Theorem 4.8. We let E be increasing. The proof for the decreasing E is quite similar and is omitted. In the proof, we let $s_{\max} = \bar{s}$ and $s_{\min} = \underline{s}$ for notational simplicity. Notice first that

$$\frac{\nu(\lambda)}{\lambda T\Big(\sup_{0\le r\le \lambda^2 t} W(r)\Big)}\int_0^{\lambda^2 t} T(W(r))\,dr = \frac{\lambda\nu(\lambda)}{T\Big(\sup_{0\le r\le t} W(\lambda^2 r)\Big)}\int_0^t T(W(\lambda^2 r))\,dr$$

$$\stackrel{d}{\equiv} \frac{\lambda\nu(\lambda)}{T(\lambda\bar{s})}\int_0^t T(\lambda W(r))\,dr,$$

for all λ. However, we have

$$\frac{\lambda\nu(\lambda)}{T(\lambda\bar{s})}\int_0^t T(\lambda W(r))\,dr = \frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})}\int_0^t T(\lambda W(r))\,dr(1+o(1)) \quad \text{a.s.}$$

$$= \frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})}\int_0^t E(\lambda W(r))\,dr(1+o(1)) \quad \text{a.s.} \tag{18}$$

because for $s_m = \max(\bar{s}, -\underline{s})$

$$\frac{|R(\lambda\bar{s})|}{E(\lambda\bar{s})} \le \frac{\bar{R}(\lambda s_m)}{E(\lambda\bar{s})} \xrightarrow{\text{a.s.}} 0$$

$$\frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})}\int_0^1 |R(\lambda W(r))|\,dr \le \frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})}\bar{R}(\lambda s_m) \xrightarrow{\text{a.s.}} 0,$$

by the condition on R.

It follows from Lemma 2.4 that

$$\frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})}\int_0^t E(\lambda W(r))\,dr = \frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})}\int_{-\infty}^{\infty} E(\lambda s)L(t,s)\,ds$$

$$= \frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})}\int_0^{\infty} E(\lambda(\bar{s}-s))L(t,\bar{s}-s)\,ds. \tag{19}$$

Now we choose a function $s(\lambda) \ge 0$ of λ such that

$$s(\lambda) \to 0 \quad \text{and} \quad \lambda\nu(\lambda)s(\lambda) \to \infty, \tag{20}$$

as $\lambda \to \infty$. It will be sufficient in what follows to set $s(\lambda) = (\lambda\nu(\lambda))^{-\eta}$ for some small $\eta > 0$. As a result of (18) and (19), it suffices to show that

$$\frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})}\int_{s(\lambda)}^{\infty} E(\lambda(\bar{s}-s))L(1,\bar{s}-s)\,ds \xrightarrow{\text{a.s.}} 0, \tag{21}$$

$$\frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})}\int_0^{s(\lambda)} E(\lambda(\bar{s}-s))L(t,\bar{s}-s)\,ds \xrightarrow{\text{a.s.}} \frac{1}{D(\bar{s})}L(t,\bar{s}) \tag{22}$$

to finish the proof. Note for $0 \le s \le s(\lambda)$ that

$$G(\lambda(\bar{s}-s)) - G(\lambda\bar{s}) = -\lambda s\dot{G}(\lambda(\bar{s}-s_0(\lambda)))$$

$$= -\lambda\nu(\lambda)s(D(\bar{s}) + o_{\text{a.s.}}(1)), \tag{23}$$

uniformly in s for large λ, where $0 \leq s_0(\lambda) \leq s(\lambda)$. By (20), $s(\lambda), s_0(\lambda) \to 0$ as $\lambda \to \infty$.

Subsequently using the fact that E is increasing and $\int_{-\infty}^{\infty} L(t,s)\, ds = t$, along with (23), we have

$$\begin{aligned}
&\frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})} \int_{s(\lambda)}^{\infty} E(\lambda(\bar{s}-s))L(t,\bar{s}-s)\, ds \\
&\quad \leq \lambda\nu(\lambda)t\, \frac{E(\lambda(\bar{s}-s(\lambda)))}{E(\lambda\bar{s})} \\
&\quad = \lambda\nu(\lambda)t \exp(G(\lambda(\bar{s}-s(\lambda))) - G(\lambda\bar{s})) \\
&\quad = \lambda\nu(\lambda)t \exp(-\lambda\nu(\lambda)s(\lambda)(D(\bar{s}) + o(1))) \xrightarrow{\text{a.s.}} 0,
\end{aligned}$$

as $\lambda \to \infty$, because $D(\bar{s}) > 0$ and

$$\lambda\nu(\lambda)\exp(-\lambda\nu(\lambda)s(\lambda)) = \lambda\nu(\lambda)\exp(-(\lambda\nu(\lambda))^{1-\eta}) \to 0,$$

by (20). This shows (21). Now, by (23) again,

$$\begin{aligned}
&\frac{\lambda\nu(\lambda)}{E(\lambda\bar{s})} \int_{0}^{s(\lambda)} E(\lambda(\bar{s}-s))L(t,\bar{s}-s)\, ds \\
&\quad = \lambda\nu(\lambda) \int_0^{s(\lambda)} \exp(-\lambda\nu(\lambda)sD(\bar{s})(1+o(1)))L(t,\bar{s}-s)\, ds \\
&\quad = L(t,\bar{s})\lambda\nu(\lambda) \int_0^{s(\lambda)} \exp(-\lambda\nu(\lambda)sD(\bar{s}))\, ds(1+o(1)) \\
&\quad = L(t,\bar{s}) \int_0^{\lambda\nu(\lambda)s(\lambda)} \exp(-sD(\bar{s}))\, ds(1+o(1)) \\
&\quad \xrightarrow{\text{a.s.}} \frac{1}{D(\bar{s})} L(t,\bar{s}),
\end{aligned}$$

and this proves (22). ∎

Proof of Theorem 5.1. Assume $x_0 = 0$ and write

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} T(x_t) =_d \sqrt{n} \int_0^1 T(\sqrt{n}W_n(r))\, dr.$$

If $x_0 \neq 0$, then we may consider the function $T(\cdot + x_0)$ in place of $T(\cdot)$. It is easy to see that all the proofs go through under this replacement.

Now let

$$\kappa_n = n^a \quad \text{and} \quad \delta_n = n^{-b} \tag{24}$$

for $a, b > 0$ satisfying

$$a - (1+\ell)b < 0, \tag{25}$$

$$(6b-1)p + 2 < 0, \tag{26}$$

$$2a - 1 < 0, \tag{27}$$

$$4a - 4b - 1 < 0, \tag{28}$$

$$(a-b)p - 1 > 0 \tag{29}$$

and define T_n, T_n', and T_n'' by

$$T_n(x) = T(x)1\{-\kappa_n\delta_n \le x < \kappa_n\delta_n\},$$

$$T_n'(x) = T(x)1\{x \ge \kappa_n\delta_n\},$$

$$T_n''(x) = T(x)1\{x < -\kappa_n\delta_n\},$$

so that $T = T_n + T_n' + T_n''$. We will show that

$$\sqrt{n}\int_0^1 T_n(\sqrt{n}W_n(r))\,dr = \left(\int_{-\infty}^{\infty} T(s)\,ds\right)L(1,0) + o_p(1), \tag{30}$$

and

$$\sqrt{n}\int_0^1 T_n'(\sqrt{n}W_n(r))\,dr = o_p(1), \tag{31}$$

$$\sqrt{n}\int_0^1 T_n''(\sqrt{n}W_n(r))\,dr = o_p(1), \tag{32}$$

from which the stated result follows directly. For notational brevity, set $\nu_n = \sqrt{n}$ and let $N_n(a,b) = N_n(\nu_n; a,b)$ and $N(a,b) = N(\nu_n; a,b)$ in what follows, for N_n and N defined in (4) and (5).

To show (30), we first define

$$T_{\delta_n}(x) = \sum_{k=-\kappa_n}^{\kappa_n - 1} T(k\delta_n)1\{k\delta_n \le x < (k+1)\delta_n\}.$$

It follows from the Lipschitz condition for T that $\sup|T_n(x) - T_{\delta_n}(x)| \le c\delta_n^{\ell}$, and therefore,

$$\left|\sqrt{n}\int_0^1 T_n(\sqrt{n}W_n(r))\,dr - \sqrt{n}\int_0^1 T_{\delta_n}(\sqrt{n}W_n(r))\,dr\right|$$
$$\le c\kappa_n\delta_n^{1+\ell}\left(\frac{\sqrt{n}}{\kappa_n\delta_n}N_n(-\kappa_n\delta_n, \kappa_n\delta_n)\right) = O_p(\kappa_n\delta_n^{1+\ell}) = o_p(1), \tag{33}$$

given the conditions for κ_n and δ_n in (24) and (25). Note that

$$\frac{\sqrt{n}}{\kappa_n\delta_n}N_n(-\kappa_n\delta_n, \kappa_n\delta_n) = 2L(1,0) + o_p(1),$$

under condition (26), as a result of Lemma 2.5(b).

Now,

$$\sqrt{n}\int_0^1 T_{\delta_n}(\sqrt{n}W_n(r))\,dr = \sqrt{n}\sum_{k=-\kappa_n}^{\kappa_n-1} T(k\delta_n)N_n(k\delta_n,(k+1)\delta_n)$$
$$= \sqrt{n}\left(\sum_{k=-\kappa_n}^{\kappa_n-1} T(k\delta_n)\right)N_n(0,\delta_n) + R_n, \quad (34)$$

where

$$R_n = \sqrt{n}\sum_{k=-\kappa_n}^{\kappa_n-1} T(k\delta_n)(N_n(k\delta_n,(k+1)\delta_n) - N_n(0,\delta_n)).$$

It follows from the Cauchy–Schwarz inequality and Lemma 2.5(a) that

$$\mathbf{E}(R_n^2) \le n\left(\sum_{k=-\kappa_n}^{\kappa_n-1} T(k\delta_n)^2\right)\sum_{k=-\kappa_n}^{\kappa_n-1}\mathbf{E}(N_n(0,\delta_n) - N_n(k\delta_n,(k+1)\delta_n))^2$$
$$\le \left(\int_{-\infty}^{\infty} T_{\delta_n}^2(s)\,ds\right)\left(c_1\frac{\kappa_n}{\sqrt{n}} + c_2\frac{\kappa_n^2\delta_n^2\log n}{\sqrt{n}}\right) = o(1),$$

because of the conditions for κ_n and δ_n in (24), (27), and (28), and where c_1 and c_2 are some constants.

However, we have

$$\sqrt{n}\left(\sum_{k=-\kappa_n}^{\kappa_n-1} T(k\delta_n)\right)N_n(0,\delta_n) = \left(\int_{-\infty}^{\infty} T_{\delta_n}(s)\,ds\right)\frac{\sqrt{n}}{\delta_n}N_n(0,\delta_n)$$
$$= \left(\int_{-\infty}^{\infty} T(s)\,ds\right)L(1,0) + o_p(1), \quad (35)$$

as a result of (26) for κ_n and δ_n in (24). Notice that

$$\frac{\sqrt{n}}{\delta_n}N_n(0,\delta_n) \xrightarrow{p} L(1,0),$$

under condition (26), by Lemma 2.5(b). Moreover,

$$\int_{-\infty}^{\infty} T_{\delta_n}(s)\,ds = \int_{-\infty}^{\infty} T_n(s)\,ds + o(\kappa_n\delta_n^2),$$
$$\int_{-\infty}^{\infty} T_n(s)\,ds = \int_{-\infty}^{\infty} T(s)\,ds + o(1).$$

We now have (30) from (33), (34), and (35).

Next we show (31) and (32). Let

$$\varepsilon_n = \sup_{0\le r\le 1}|W_n(r) - W(r)|. \quad (36)$$

By taking n sufficiently large, we may assume that T_n' and T_n'' are monotone (decreasing and increasing, respectively) on their supports. This causes no loss in generality, because we may always bound T_n' and T_n'' by such functions if T is integrable. Therefore,

$$T_n'(\sqrt{n}W_n(r)) \leq T(\sqrt{n}(W(r) - \varepsilon_n))1\{\sqrt{n}(W_n(r) + \varepsilon_n) > \kappa_n\delta_n\},$$

$$T_n''(\sqrt{n}W_n(r)) \leq T(\sqrt{n}(W(r) + \varepsilon_n))1\{\sqrt{n}(W_n(r) - \varepsilon_n) < -\kappa_n\delta_n\}.$$

It follows that

$$\begin{aligned}
&\sqrt{n}\int_0^1 T_n'(\sqrt{n}W_n(r))\,dr \\
&\quad \leq \sqrt{n}\int_0^1 T(\sqrt{n}(W(r) - \varepsilon_n))1\{\sqrt{n}(W(r) + \varepsilon_n) > \kappa_n\delta_n\}\,dr \\
&\quad = \sqrt{n}\int_{-\infty}^{\infty} T(\sqrt{n}(s - \varepsilon_n))1\{\sqrt{n}(s + \varepsilon_n) > \kappa_n\delta_n\}L(1,s)\,ds \\
&\quad = \int_{-\infty}^{\infty} T(s)1\{s > \kappa_n\delta_n - 2\sqrt{n}\varepsilon_n\}L\left(1, \frac{s}{\sqrt{n}} + \varepsilon_n\right)ds \xrightarrow{p} 0,
\end{aligned}$$

because $\kappa_n\delta_n - 2\sqrt{n}\varepsilon_n \to_p \infty$, as a result of (24) and (29). Similarly,

$$\begin{aligned}
&\sqrt{n}\int_0^1 T_n''(\sqrt{n}W_n(r))\,dr \\
&\quad \leq \sqrt{n}\int_0^1 T(\sqrt{n}(W(r) + \varepsilon_n))1\{\sqrt{n}(W(r) - \varepsilon_n) > \kappa_n\delta_n\}\,dr \\
&\quad = \sqrt{n}\int_{-\infty}^{\infty} T(\sqrt{n}(s + \varepsilon_n))1\{\sqrt{n}(s - \varepsilon_n) > \kappa_n\delta_n\}L(1,s)\,ds \\
&\quad = \int_{-\infty}^{\infty} T(s)1\{s < -\kappa_n\delta_n + 2\sqrt{n}\varepsilon_n\}L\left(1, \frac{s}{\sqrt{n}} - \varepsilon_n\right)ds \xrightarrow{p} 0,
\end{aligned}$$

because $-\kappa_n\delta_n + 2\sqrt{n}\varepsilon_n \to_p -\infty$, again because of (24) and (29). The proof is therefore complete. ∎

Proof of Theorem 5.3. Again let $x_0 = 0$ for simplicity. The proof for $x_0 \neq 0$ is the same with only $W_n(r)$ being replaced by $x_0/\sqrt{n} + W_n(r)$ in what follows. Write

$$\begin{aligned}
\frac{1}{n\nu(\sqrt{n})}\sum_{t=1}^{n} T(x_t) &\stackrel{d}{=} \frac{1}{\nu(\sqrt{n})}\int_0^1 T(\sqrt{n}W_n(r))\,dr \\
&= \int_0^1 H(W_n(r))\,dr + \frac{1}{\nu(\sqrt{n})}\int_0^1 R(W_n(r),\sqrt{n})\,dr.
\end{aligned}$$

Because H is regular, it follows that

$$\int_0^1 H(W_n(r))\,dr \xrightarrow{\text{a.s.}} \int_0^1 H(W(r))\,dr = \int_{-\infty}^{\infty} H(s)L(1,s)\,ds.$$

It therefore suffices to show

$$\frac{1}{\nu(\sqrt{n})}\int_0^1 R(W_n(r),\sqrt{n})\,dr \xrightarrow{\text{a.s.}} 0$$

to complete the proof.

If $T \in \mathcal{T}(H_1)$, it follows immediately that

$$\frac{1}{\nu(\sqrt{n})}\int_0^1 |R(W_n(r),\sqrt{n})|\,dr \le \frac{a(\sqrt{n})}{\nu(\sqrt{n})}\int_0^1 P(W_n(r))\,dr \xrightarrow{\text{a.s.}} 0,$$

because P is locally bounded. For $T \in \mathcal{T}(H_2)$, we need to show

$$\frac{1}{\nu(\sqrt{n})}\int_0^1 |R(W_n(r),\sqrt{n})|\,dr \le \frac{b(\sqrt{n})}{\nu(\sqrt{n})}\int_0^1 Q(\sqrt{n}W_n(r))\,dr \xrightarrow{\text{a.s.}} 0, \tag{37}$$

where Q is bounded and vanishes at infinity. We may assume w.l.o.g. that Q is monotone decreasing (increasing) for $x > 0$ ($x < 0$), as noted in the proof of Theorem 4.7. We may thus write $Q = Q_1 - Q_2$ with both Q_1 and Q_2 bounded and nondecreasing and let ε_n be defined as in (36). It follows that

$$\begin{aligned} &Q_1(\sqrt{n}(W(r)-\varepsilon_n)) - Q_2(\sqrt{n}(W(r)+\varepsilon_n)) \\ &\quad \le Q(\sqrt{n}W_n(r)) \\ &\quad \le Q_1(\sqrt{n}(W(r)+\varepsilon_n)) - Q_2(\sqrt{n}(W(r)-\varepsilon_n)). \end{aligned} \tag{38}$$

However,

$$\begin{aligned} \int_0^1 Q_i(\sqrt{n}(W(r)\pm\varepsilon_n))\,dr &= \int_{-\infty}^{\infty} Q_i(\sqrt{n}(s\pm\varepsilon_n))L(1,s)\,ds \\ &= \int_{-\infty}^{\infty} Q_i(\sqrt{n}s)L(1,s\mp\varepsilon_n)\,ds \\ &= \int_{-\infty}^{\infty} Q_i(\sqrt{n}s)L(1,s)\,ds(1+o(1)) \text{ a.s.}, \end{aligned}$$

because the Q_i's are bounded and $L(1,\cdot)$ is continuous. Therefore,

$$\begin{aligned} &\int_0^1 (Q_1(\sqrt{n}(W(r)\mp\varepsilon_n)) - Q_2(\sqrt{n}(W(r)\pm\varepsilon_n)))\,dr \\ &\quad = \int_{-\infty}^{\infty} (Q_1(\sqrt{n}(s\mp\varepsilon_n)) - Q_2(\sqrt{n}(s\pm\varepsilon_n)))L(1,s)\,ds \\ &\quad = \int_{-\infty}^{\infty} Q(\sqrt{n}s)L(1,s)\,ds(1+o(1)) \text{ a.s.} \end{aligned} \tag{39}$$

Now (37) follows easily from (38) and (39), as a result of (17). ∎

Proof of Theorem 5.5. Let E be increasing and let $\bar{s}_n = \sup W_n(r)$ and $\bar{s} = \sup W(r)$. For simplicity, assume $x_0 = 0$. For the case $x_0 \neq 0$, we replace $W_n(r)$ and $\bar{s}_n$, respectively, by $W_n(r) + x_0/\sqrt{n}$ and $\bar{s}_n + x_0/\sqrt{n}$ in what follows. All the proofs go through with this replacement. Write

$$\frac{\nu(\sqrt{n})}{\sqrt{n}T\left(\max\limits_{1\le t\le n} x_t\right)} \sum_{t=1}^{n} T(x_t) \stackrel{d}{=} \frac{\sqrt{n}\nu(\sqrt{n})}{T(\sqrt{n}\bar{s}_n)} \int_0^1 T(\sqrt{n}W_n(r))\, dr$$

and notice that

$$\frac{\sqrt{n}\nu(\sqrt{n})}{T(\sqrt{n}\bar{s}_n)} \int_0^1 T(\sqrt{n}W_n(r))\, dr = \frac{\sqrt{n}\nu(\sqrt{n})}{E(\sqrt{n}\bar{s}_n)} \int_0^1 E(\sqrt{n}W_n(r))\, dr(1 + o_p(1)), \tag{40}$$

which we can show in the same way as (18) in the proof of Theorem 4.8.

Let $\nu_n = \sqrt{n}\nu(\sqrt{n})$ and let s_n be a sequence of numbers such that $s_n \to 0$ and $\nu_n s_n \to \infty$. Because $\bar{s}_n \to_p \bar{s}$ and $s_n \to 0$, we have similar to (23) in the proof of Theorem 4.8

$$G(\sqrt{n}(\bar{s}_n - s)) - G(\sqrt{n}\bar{s}_n) = -\nu_n s(D(\bar{s}) + o_p(1)), \tag{41}$$

uniformly in $s \in [0, s_n]$, for sufficiently large n. Therefore, if we write

$$\frac{\nu_n}{E(\sqrt{n}\bar{s}_n)} \int_0^1 E(\sqrt{n}W_n(r))\, dr = A_n + B_n, \tag{42}$$

where

$$A_n = \frac{\nu_n}{E(\sqrt{n}\bar{s}_n)} \int_0^1 E(\sqrt{n}W_n(r))1\{W_n(r) \ge \bar{s}_n - s_n\}\, dr,$$

$$B_n = \frac{\nu_n}{E(\sqrt{n}\bar{s}_n)} \int_0^1 E(\sqrt{n}W_n(r))1\{W_n(r) < \bar{s}_n - s_n\}\, dr,$$

then it follows from (41) that

$$A_n = \nu_n \int_0^1 \exp(-\nu_n D(\bar{s})(\bar{s}_n - W_n(r)))\, dr(1 + o_p(1)),$$

$$B_n = o_p(1),$$

in parallel to (21) and (22) in the proof of Theorem 4.8.

Define W_n' and W' by

$$W_n'(r) = \bar{s}_n - W_n(r) \quad \text{and} \quad W'(r) = \bar{s} - W(r),$$

i.e., Brownian motion reflected at the supremum and its sample analogue. Denote by L' the local time of W'. Furthermore, we define N_n' and N' for W_n' and W' in the same way as N_n and N for W_n and W given in (4) and (5), respectively. Write $N_n'(a,b) = N_n'(\nu_n; a,b)$ and $N'(a,b) = N'(\nu_n; a,b)$ for short. Though we do not provide the details, it is obvious that all the results in Akonom (1993), and therefore our Lemmas 2.3 and 2.5, hold for W_n' and W', and also for W_n and W.

Now we write

$$A_n = \nu_n \int_0^1 F(\nu_n W_n'(r))\,dr(1 + o_p(1)),$$

with

$$F(x) = e^{-xD(\bar{s})}\{x \geq 0\}.$$

To analyze A_n, we define κ_n and δ_n as in (24) with a and b satisfying

$$a - 2b < 0, \tag{43}$$

$$2a + m - 1 < 0, \tag{44}$$

$$4a - 4b - m - 1 < 0, \tag{45}$$

$$(6b + 3m - 1)p + 2 < 0, \tag{46}$$

$$(2a - 2b - m)p - 2 > 0 \tag{47}$$

and let $\nu_n s_n = \kappa_n \delta_n$. It is tedious but straightforward to check that a and b satisfying all (43)–(47) exist, given our conditions on m and p.

We decompose F into F_n and F_n', where

$$F_n(x) = e^{-xD(\bar{s})}1\{0 \leq x < \kappa_n \delta_n\},$$

$$F_n'(x) = e^{-xD(\bar{s})}1\{x \geq \kappa_n \delta_n\}.$$

It will be shown that

$$\nu_n \int_0^1 F_n(\nu_n W_n'(r))\,dr = \left(\int_{-\infty}^{\infty} F(s)\,ds\right) L'(0,1) + o_p(1), \tag{48}$$

$$\nu_n \int_0^1 F_n'(\nu_n W_n'(r))\,dr = o_p(1), \tag{49}$$

from which we may easily deduce the stated result, upon noticing that

$$\int_{-\infty}^{\infty} F(s)\,ds = \frac{1}{D(\bar{s})} \quad \text{and} \quad L'(1,0) = L(1,\bar{s}),$$

together with (40) and (42).

To show (48), we first introduce

$$F_{\delta_n}(x) = \sum_{k=0}^{\kappa_n - 1} e^{-k\delta_n D(\bar{s})} 1\{k\delta_n \leq x < (k+1)\delta_n\}$$

and notice that

$$\left| \nu_n \int_0^1 F_n(\nu_n W_n'(r))\,dr - \nu_n \int_0^1 F_{\delta_n}(\nu_n W_n'(r))\,dr \right| \leq \kappa_n \delta_n^2 \left(\frac{\nu_n}{\kappa_n \delta_n} N_n'(0, \kappa_n \delta_n) \right)$$

$$= O_p(\kappa_n \delta_n^2) = o_p(1), \tag{50}$$

under conditions (43) and (46). Note that

$$\frac{\nu_n}{\kappa_n \delta_n} N_n'(0,\kappa_n \delta_n) = L'(0,1) + o_p(1),$$

under condition (46) by Lemma 2.5(b).

Second,

$$\nu_n \int_0^1 F_{\delta_n}(\nu_n W_n'(r))\, dr = \nu_n \sum_{k=0}^{\kappa_n - 1} e^{-k\delta_n D(\bar{s})} N_n'(k\delta_n,(k+1)\delta_n)$$
$$= \nu_n \left(\sum_{k=0}^{\kappa_n - 1} e^{-k\delta_n D(\bar{s})} \right) N_n'(0,\delta_n) + R_n', \quad (51)$$

where

$$R_n' = \nu_n \sum_{k=0}^{\kappa_n - 1} e^{-k\delta_n D(\bar{s})} (N_n'(k\delta_n,(k+1)\delta_n) - N_n'(0,\delta_n)),$$

and therefore,

$$\mathbf{E}(R_n'^2) \le \nu_n^2 \left(\sum_{k=0}^{\kappa_n - 1} e^{-2k\delta_n D(\bar{s})} \right) \sum_{k=0}^{\kappa_n - 1} \mathbf{E}(N_n'(k\delta_n,(k+1)\delta_n) - N_n'(0,\delta_n))^2$$
$$\le \left(\int_{-\infty}^{\infty} F_{\delta_n}^2(s)\, ds \right) \left(c_1 \frac{\nu_n \kappa_n}{n} + c_2 \frac{\delta_n^2 \kappa_n^2 \log n}{\nu_n} \right) \to 0,$$

by conditions (44) and (45), where c_1 and c_2 are some constants.

Third,

$$\nu_n \left(\sum_{k=0}^{\kappa_n - 1} e^{-k\delta_n D(\bar{s})} \right) N_n'(0,\delta_n) = \left(\int_{-\infty}^{\infty} F_{\delta_n}(s)\, ds \right) \left(\frac{\nu_n}{\delta_n} N_n'(0,\delta_n) \right)$$
$$= \left(\int_{-\infty}^{\infty} F(s)\, ds \right) L'(1,0) + o_p(1). \quad (52)$$

Notice that

$$\int_{-\infty}^{\infty} F_{\delta_n}(s)\, ds = \int_{-\infty}^{\infty} F_n(s)\, ds + O(\kappa_n \delta_n^2),$$
$$\int_{-\infty}^{\infty} F_n(s)\, ds = \int_{-\infty}^{\infty} F(s)\, ds + O(e^{-\kappa_n \delta_n}).$$

Also, by Lemma 2.5(b)

$$\frac{\nu_n}{\delta_n} N_n'(0,\delta_n) = L'(1,0) + o_p(1),$$

under condition (46). Then (48) follows from (50), (51), and (52).

Finally, for ε_n defined in (36)

$$\nu_n \int_0^1 F_n'(\nu_n W_n'(r))\,dr$$

$$\leq \nu_n \int_0^1 F(\nu_n(W'(r) - \varepsilon_n))1\{\nu_n(W'(r) + \varepsilon_n) > \kappa_n \delta_n\}\,dr$$

$$= \nu_n \int_{-\infty}^{\infty} F(\nu_n(s - \varepsilon_n))1\{\nu_n(s + \varepsilon_n) > \kappa_n \delta_n\} L'(1,s)\,ds$$

$$= \int_{-\infty}^{\infty} F(s)1\{s > \kappa_n \delta_n - \nu_n \varepsilon_n\} L'\left(1, \frac{s}{\sqrt{n}} + \varepsilon_n\right) ds \xrightarrow{p} 0,$$

because $\kappa_n \delta_n - \nu_n \varepsilon_n \to_p \infty$ under condition (47), which proves (49). The proof is therefore complete. ∎

Proof of Lemma 6.2. By Theorem A1, page 269 of Hall and Heyde (1980), there exist a probability space $(\Omega, \mathbf{P}, \mathcal{F})$ supporting $\{U_t\}$, $U_t = \sum_{k=1}^{t} u_k$, a Brownian motion U with variance σ^2, and a time change $\{\tau_t\}$ such that

(a) τ_t is $\mathcal{F}_t$-measurable,
(b) $\mathbf{E}((\Delta\tau_t)^r | \mathcal{F}_{t-1}) \leq \mathbf{E}(|u_t|^{2r} | \mathcal{F}_{t-1})$ a.s. for $r \geq 1$, and
(c) $\mathbf{E}(\Delta\tau_t | \mathcal{F}_{t-1}) = 1$,

where $\mathcal{F}_t$ is the σ-field generated by $(U_k)_{k=1}^{t}$ and $U(r)$ for $0 \leq r \leq \tau_t$.

Let $1 \leq r \leq \min(2, q/2)$. Then we have

$$\mathbf{E}(|\Delta\tau_t - 1|^r | \mathcal{F}_{t-1}) \leq c \sup_{t\geq 1} \mathbf{E}(|u_t|^q | \mathcal{F}_{t-1}) < \infty \quad \text{a.s.}$$

for some constant c. Therefore,

$$\sum_{t=1}^{\infty} t^{-r\delta} \mathbf{E}(|\Delta\tau_t - 1|^r | \mathcal{F}_{t-1}) < \infty \quad \text{a.s.}$$

because $r\delta > 1$, and we have from Theorem 2.18 of Hall and Heyde (1980) that

$$\frac{\tau_t - t}{t^\delta} \xrightarrow{\text{a.s.}} 0,$$

as $t \to \infty$ for $\delta > \max(\frac{1}{2}, 2/q)$. Therefore, for any $\varepsilon > 0$ given, there exists n' such that $|\tau_t - t|/t^\delta < \varepsilon$ for all $t > n'$. Choose $n \geq n'$ such that $n > (\max_{1\leq t\leq n'} |\tau_t - t|/\varepsilon)^{1/\delta}$. It is easy to check

$$\sup_{1\leq t\leq n} \left| \frac{\tau_t - t}{n^\delta} \right| < \varepsilon \quad \text{a.s.}$$

as was to be shown. ∎

Proof of Theorem 6.3. To prove part (a), construct the process

$$M_n(r) = \sqrt[4]{n}\sum_{t=1}^{k-1} f\left(\sqrt{n}W_n\left(\frac{t}{n}\right)\right)\left(U\left(\frac{\tau_t}{n}\right) - U\left(\frac{\tau_{t-1}}{n}\right)\right)$$
$$+ \sqrt[4]{n} f\left(\sqrt{n}W_n\left(\frac{k}{n}\right)\right)\left(U(r) - U\left(\frac{\tau_{k-1}}{n}\right)\right), \quad (53)$$

for $\tau_{k-1}/n < r \leq \tau_k/n$, $k = 1,\dots,n$. Note that M_n is a continuous martingale such that

$$\frac{1}{\sqrt[4]{n}}\sum_{t=1}^{n} f(x_t)u_t \stackrel{d}{=} M_n\left(\frac{\tau_n}{n}\right),$$

in view of the construction in Lemma 6.2. The quadratic variation process $[M_n]$ of M_n is given by

$$[M_n]_r = \sqrt{n}\sum_{t=1}^{k-1} f^2\left(\sqrt{n}W_n\left(\frac{t}{n}\right)\right)\left(\frac{\tau_t}{n} - \frac{\tau_{t-1}}{n}\right) + \sqrt{n} f^2\left(\sqrt{n}W_n\left(\frac{k}{n}\right)\right)\left(r - \frac{\tau_{k-1}}{n}\right)$$
$$= \sqrt{n}\int_0^r f^2(\sqrt{n}W_n(s))\,ds + o_p(1),$$

because

$$\sup_{1\leq t\leq n}\left|\left(\frac{\tau_t}{n} - \frac{\tau_{t-1}}{n}\right) - \frac{1}{n}\right| = o(1) \quad \text{a.s.},$$

as a result of Lemma 6.2. Therefore,

$$[M_n]_r \xrightarrow{p} \left(\int_{-\infty}^{\infty} T(s)\,ds\right)L(r,0), \quad (54)$$

as shown in the proof of Theorem 5.1. Moreover, if we denote by $[M_n, W]$ the covariation process of M_n and W, then

$$[M_n, W]_r = 0, \quad (55)$$

for all $r \in [0,1]$, because of the independence of U and W. The asymptotic distribution of the continuous martingale M_n in (53) is completely determined by (54) and (55), as shown in Revuz and Yor (1994, Theorem 2.3, p. 496).

Now define the sequence of time changes

$$\rho_n(r) = \inf\{s \mid [M_n]_s > r\}$$

and subsequently set

$$V_n(r) = M_n(\rho_n(r)).$$

The process V_n is the DDS (or Dambis, Dubins–Schwarz) Brownian motion of the continuous martingale M_n (see, e.g., Revuz and Yor, 1994, Theorem 1.6, p. 173). It follows that (V_n, W) converges jointly in distribution to two independent standard linear Brownian motions (V, W), say. Therefore,

$$M_n\left(\frac{\tau_n}{n}\right) = M_n(1) + o_p(1)$$
$$\xrightarrow{d} V\left(\int_{-\infty}^{\infty} T(s)\, ds\, L(1,0)\right),$$

which gives the result stated in (a). The proofs for (b) and (c) are similar and are therefore omitted. ■

Proof of Proposition 6.4. The case $\kappa \geq 0$ is straightforward because

$$n^{-1-\kappa/2} \sum_{t=1}^{n} |x_t|^\kappa = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{x_t}{\sqrt{n}} \right|^\kappa \xrightarrow{d} \int_0^1 |W(r)|^\kappa\, dr, \tag{56}$$

by Theorem 3.2, because $T(x) = |x|^\kappa$ is regular. In the case where $-1 < \kappa < 0$ Theorem 3.4 is applicable and (56) again yields the stated result. For the case $\kappa \leq -1$ we use a different argument. Bound $\sum_{t=1}^{n} |x_t|^\kappa$ below as

$$\sum_{t=1}^{n} |x_t|^\kappa \geq \sum_{t=1}^{n} |x_t|^\kappa 1\left\{ \frac{|x_t - x_0|}{\sqrt{n}} \leq \frac{1}{\sqrt{n}} \right\}$$
$$\geq (1 + |x_0|)^\kappa \sum_{t=1}^{n} 1\left\{ \frac{|x_t - x_0|}{\sqrt{n}} \leq \frac{1}{\sqrt{n}} \right\}$$
$$\stackrel{d}{=} n(1 + |x_0|)^\kappa \int_0^1 1\left\{ |W_n(r)| \leq \frac{1}{\sqrt{n}} \right\} dr.$$

Then, from Lemma 2.5(b)

$$\sqrt{n} \int_0^1 1\left\{ |W_n(r)| \leq \frac{1}{\sqrt{n}} \right\} dr = \sqrt{n} \int_0^1 1\left\{ |W(r)| \leq \frac{1}{\sqrt{n}} \right\} dr + o_p(1)$$
$$= 2L(1,0) + o_p(1),$$

and thus for any $\delta > 0$

$$n^{-1/2+\delta} \sum_{t=1}^{n} |x_t|^\kappa \xrightarrow{p} \infty,$$

thereby establishing the stated result. ■

REFERENCES

Akonom, J. (1993) Comportement asymptotique du temps d'occupation du processus des sommes partielles. *Annales de l'Institut Henri Poincaré* 29, 57–81.

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

Chung, K.L. & R.J. Williams (1990) *Introduction to Stochastic Integration*, 2nd ed. Boston: Birkhäuser.

Csörgő, M. & L. Horváth (1993) *Weighted Approximations in Probability and Statistics.* New York: Wiley.

Engle, R.F. & C.W.J. Granger (1987) Co-integration and error correction: Representation, estimation and testing. *Econometrica* 55, 251–276.
Hall, P. & C.C. Heyde (1980) *Martingale Limit Theory and Its Application*. New York: Academic Press.
Jeganathan, P. (1995) Some aspects of asymptotic theory with applications to time series models. *Econometric Theory* 11, 818–887.
Karatzas, I. & S.E. Shreve (1988) *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag.
Park, J.Y. & P.C.B. Phillips (1988) Statistical inference in regressions with integrated processes: Part 1. *Econometric Theory* 4, 468–497.
Park, J.Y. & P.C.B. Phillips (1998) Nonlinear Regressions with Integrated Time Series. Mimeo, Yale University.
Phillips, P.C.B. (1987) Time series regression with a unit root. *Econometrica* 55, 277–301.
Phillips, P.C.B. (1988) Multiple regression with integrated time series. *Contemporary Mathematics* 80, 79–105.
Phillips, P.C.B. (1991) Optimal inference in cointegrated systems. *Econometrica* 59, 283–306.
Phillips, P.C.B. & W. Ploberger (1996) An asymptotic theory of Bayesian inference for time series. *Econometrica* 64, 381–412.
Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.
Pollard, D. (1984) *Convergence of Stochastic Processes*. New York: Springer-Verlag.
Revuz, D. & M. Yor (1994) *Continuous Martingale and Brownian Motion*, 2nd ed. New York: Springer-Verlag.