REGRESSION WITH SLOWLY VARYING REGRESSORS AND NONLINEAR TRENDS

BY

PETER C. B. PHILLIPS

COWLES FOUNDATION PAPER NO. 1206

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

2007

http://cowles.econ.yale.edu/

Econometric Theory, **23**, 2007, 557–614. Printed in the United States of America.
DOI: 10.1017/S0266466607070260

REGRESSION WITH SLOWLY VARYING REGRESSORS AND NONLINEAR TRENDS

PETER C.B. PHILLIPS
Cowles Foundation, Yale University
University of Auckland
and
University of York

Slowly varying (SV) regressors arise commonly in empirical econometric work, particularly in the form of semilogarithmic regression and log periodogram regression. These regressors are asymptotically collinear. Usual regression formulas for asymptotic standard errors are shown to remain valid, but rates of convergence are affected and the limit distribution of the regression coefficients is shown to be one dimensional. Some asymptotic representations of partial sums of SV functions and central limit theorems with SV weights are given that assist in the development of a regression theory. Multivariate regression and polynomial regression with SV functions are considered and shown to be equivalent, up to standardization, to regression on a polynomial in a logarithmic trend. The theory involves second-, third-, and higher-order forms of slow variation. Some applications to the asymptotic theory of nonlinear trend regression are explored.

1. INTRODUCTION

Empirical models of economic time series often involve deterministic trend functions. Time polynomials and sinusoidal polynomials are the most common functions to appear in such models, and the properties of regressions of time series on these trend functions have been extensively explored in the literature, an early and definitive contribution being Grenander and Rosenblatt (1957, Ch. 7). A common element in much of the asymptotic theory that has been developed is a requirement of the type that ensures the existence of a positive definite limit to a suitably normalized sample second moment matrix of the regressors. Frequently, this requirement appears as one of a general set of conditions on the sample variances and autocovariances of the regressors, such as those that

The author thanks two referees and Pentti Saikkonen for comments and suggestions, Sidney Resnick for references on second-order regular variation, and a Kelly Fellowship and the NSF for partial research support under grants SBR 97-30295 and SES 04-142254. An original draft of the paper was written in June 2000 and circulated under the title "Regression with Slowly Varying Regressors." Address correspondence to Peter Phillips, Cowles Foundation for Research in Economics, Yale University, Box 208281, New Haven, CT 06520-8281, USA; e-mail: peter.phillips@yale.edu

are often characterized (e.g., Hannan 1970, p. 215) as "Grenander's conditions" (see Grenander and Rosenblatt, 1957, pp. 233–234).

Not all deterministic functions of interest are covered by these requirements, and when the conditions fail some adjustments to the asymptotic theory are usually needed. One example that is important in certain empirical applications is the semilogarithmic growth model

$$y_s = \alpha + \beta \log s + u_s \qquad s = 1, \ldots, n, \tag{1}$$

where u_s is an unobserved error process. This type of formulation arises naturally in the study of growth convergence problems and economic transition. For instance, in the study of growth convergence (Barro and Sala-i-Martin, 2004), if σ_t^2 denotes the variation across economies of the logarithm of per capita real output at time t, then under σ convergence, this dispersion declines over time. Accordingly, if we model dispersion in a power law form as $\sigma_t^2 = at^{\beta}e^{u_t}$ for some $a > 0, \beta \le 0$ and with accompanying disturbances u_t, then $y_t = \log(\sigma_t^2)$ follows (1), and σ convergence may be tested using a robust sign test on the slope coefficient β.

In quite a different context, an analogous formulation arises in the log periodogram analysis of long memory, a subject on which there is now a large literature (see Robinson, 1995; Hurvich, Deo, and Brodsky, 1998; Phillips, 1999; and the references therein). In that case (discussed in Example (a) in Section 3), y_s is the periodogram of the data measured at the Fourier frequencies $\lambda_s = 2\pi s/n$, $s = 1, \ldots, m \le n$, and the slope coefficient $\beta = -2d$, where d is the memory parameter.

The reason model (1) fails to fit within the usual framework is that the sample moment matrix of the regressors is asymptotically singular. Indeed, setting $D_n = diag(\sqrt{n}, \sqrt{n}\log n)$, and $F_n^{-1} = diag(\sqrt{n}/(\log n), \sqrt{n})$, we have (cf. eqns. (26) and (27) in Section 3)

$$D_n^{-1}\begin{bmatrix} n & \sum_{s=1}^{n} \log s \\ \sum_{s=1}^{n} \log s & \sum_{s=1}^{n} \log^2 s \end{bmatrix} D_n^{-1} \to \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

and

$$F_n^{-1}\begin{bmatrix} n & \sum_{s=1}^{n} \log s \\ \sum_{s=1}^{n} \log s & \sum_{s=1}^{n} \log^2 s \end{bmatrix}^{-1} F_n^{-1} \to \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

So, both the sample second moment of the regressors and its inverse have singular limits after standardization, thereby failing Grenander's conditions.

The same problem arises when the logarithmic function in (1) is replaced by any slowly varying (SV) function $L(s)$. In effect, the intercept and any SV function are asymptotically collinear after appropriate standardization. The phenomenon is manifest in a more serious way when one considers polynomial versions of (1) such as

$$y_s = \sum_{j=0}^{p} \beta_j \log^j s + u_s \qquad s \geq 1$$

or similar regressions involving polynomials in an SV function. In such cases, one finds that the sample moment matrix of the regressors, whereas of rank $p + 1$ for all $n > p$, is singular and of rank unity in the limit after suitable normalization. More generally still, the singularity persists when the regressors constitute a vector of different SV functions, such as $\{\log s, 1/\log s\}$ involving a logarithmic and inverse logarithmic trend.

In practical statistical work the phenomenon arises in nonlinear regressions of the type

$$y_s = \beta s^{\gamma} + u_s \qquad s = 1, \ldots, n, \tag{2}$$

where the trend exponent $\gamma > -\frac{1}{2}$ is to be estimated along with the regression coefficient β. The affine linear form of (2), taken about the true values of the parameters (denoted by β_0 and γ_0), involves the regressors s^{γ_0} and $s^{\gamma_0}(\log s)$, which are regularly varying and whose second moment matrix is asymptotically singular upon appropriate (multivariate) normalization (cf. eqn. (55) in Section 6). It follows that statistical models like (2) manifest asymptotic collinearity analogous to that of the linear regression (1). Wu (1981, p. 509) noted that model (2) failed his conditions (which require a single normalizing quantity and a positive definite limit matrix for the second moment matrix of the affine model) for asymptotic normality and consequently did not provide a limit distribution theory for this model.

The present paper provides a detailed treatment of regressions of this type. The discussion is conducted in terms of SV regressors, and some results on polynomial and multivariate functions of slow variation are obtained that may be of interest outside the present study. The paper is organized as follows. Section 2 lays out some assumptions and preliminary theory. Results for simple regression are given with some common examples in Section 3. Polynomial regressions in SV regressors are covered in Section 4. Some general multivariate extensions are reported in Section 5. Section 6 applies the theory to the nonlinear trend model (2). Sections 7 and 8 contain supplementary technical results and proofs. Notation is tabulated as follows:

$\to_{a.s.}$	almost sure convergence	SV	slowly varying
$\to_p$	convergence in probability	SSV	smoothly slowly varying
$=_d$	distributional equivalence	$\Rightarrow$, $\to_d$	weak convergence
$:=$	definitional equality	$[\cdot]$	integer part of
$(a)_k$	$a(a+1)\dots(a+k-1)$	$r \wedge s$	$\min(r,s)$
$B(r)$	standard Brownian motion	$\sim$	asymptotic equivalence
C^k	class of continuously differentiable functions to order k	$o_p(1)$	tends to zero in probability
		$o_{a.s.}(1)$	tends to zero almost surely

2. ASSUMPTIONS AND PRELIMINARY RESULTS

It will be convenient to use some standard theory of SV functions, and, in so doing, we shall repeatedly reference Bingham, Goldie, and Teugels (1987), hereafter designated as BGT. From the Karamata representation (e.g., BGT, Thm. 1.3.1, p. 12), any SV function $L(x)$ has the representation

$$L(x) = c(x)\exp\left(\int_a^x \frac{\varepsilon(t)}{t}\,dt\right) \quad \text{for } x > a \tag{3}$$

for some $a > 0$ and where $c(\cdot)$ is measurable with $c(x) \to c \in (0,\infty)$ and $\varepsilon(x) \to 0$ as $x \to \infty$. The function ε in (3) is referred to as the ε-function corresponding to L.

The present paper works with the subclass of (so-called) normalized SV functions for which $c(x)$ is a constant. In the development of an asymptotic theory of regression, little seems to be lost in making the restriction to constant c functions because the asymptotic behavior of $L(x)$ is equivalent to that of (3) with $c(x) = c$. It is also known that for every SV function L there is an asymptotically equivalent SV function that is arbitrarily smooth (e.g., BGT, Thm. 1.3.3, p. 14). This property is especially helpful in developing asymptotic representations and working with transforms that arise from the process of integration and differentiation. The limit behavior studied subsequently is determined by L and ε, and some properties, as we shall see, are invariant to the particular SV function.

To validate the expansions needed in our development of an asymptotic theory of regression, we shall make the following assumption.

Assumption SSV.

(a) $L(x)$ is a smoothly slowly varying (SSV) function with Karamata representation

$$L(x) = c\exp\left(\int_a^x \frac{\varepsilon(t)}{t}\,dt\right) \quad \text{for } x \geq a \tag{4}$$

for some $a > 0$ and where $c > 0$ is a constant, $\varepsilon \in C^\infty$, and $\varepsilon(x) \to 0$ as $x \to \infty$.

(b) $|\varepsilon(x)|$ is SSV, and ε has Karamata representation

$$\varepsilon(x) = c_\varepsilon \exp\left(\int_a^x \frac{\eta(t)}{t}\,dt\right) \quad \text{for } x \geq a \tag{5}$$

for some (possibly negative) constant c_ε and where $\eta \in C^\infty$, $|\eta|$ is SSV, and $\eta(x)^2 = o(\varepsilon(x)) \to 0$ as $x \to \infty$.

We call $\varepsilon(x)$ and $\eta(x)$ the ε- and η-functions of $L(x)$. Under Assumption SSV we have

$$\varepsilon(x) = \frac{xL'(x)}{L(x)} \to 0 \quad \text{and} \quad \eta(x) = \frac{x\varepsilon'(x)}{\varepsilon(x)} \to 0 \quad \text{as } x \to \infty$$

and, more generally,

$$\frac{x^m L^{(m)}(x)}{L(x)}, \frac{x^m \varepsilon^{(m)}(x)}{\varepsilon(x)}, \frac{x^m \eta^{(m)}(x)}{\eta(x)} \to 0 \quad \text{for all } m = 1, 2, \ldots \quad \text{as } x \to \infty$$

(BGT, p. 44). The class for $L(x)$ covered in Assumption SSV includes all of the common SV deterministic functions such as (for $\gamma > 0$) $\log^\gamma x$, $1/\log^\gamma x$, $\log\log x$, and $1/\log\log x$ that might appear directly in simple regression formulations or indirectly in nonlinear regression through the corresponding affine linear models. The ε- and η-functions of these functions are given in Examples (a)–(e) in Section 3.

Because we contemplate the use of L as a time series regressor, the value of the initialization a in (4) is not important. In fact, we may reset $a = 0$ by taking $\varepsilon(t) = 0$ over $t \in [0, \delta]$ for some small $\delta > 0$ and by interpolating ε over $[\delta, a]$ so that $\varepsilon \in C^\infty[0,\infty]$, thereby assuring existence, integrability, and smooth behavior for L over $[0, a]$. We shall henceforth presume that this change has been made and that we can majorize $L(rn)/L(n) - 1$ as follows:

$$\left|\frac{L(rn)}{L(n)} - 1\right| \leq K(n)g(r),$$

where $K(n)$ is SSV and $g(r) \in C[0,1]$. In consequence, and using the fact that for any SV function K, $K(n)/n^\eta \to 0$ for arbitrary $\eta > 0$, we have, given some $\alpha > 0$ and any positive integer k,

$$\int_0^{1/n^\alpha} \left(\frac{L(rn)}{L(n)} - 1\right)^k dr = o\left(\frac{1}{n^\delta}\right) \quad \text{as } n \to \infty, \tag{6}$$

where $\delta = \alpha - \eta > 0$.

To deliver an asymptotic theory of regression we need to appeal to a central limit result. For this purpose, it is convenient to assume that the regression errors u_s satisfy the following linear process condition.

Assumption LP. For all $t > 0$, u_t has Wold representation

$$u_t = C(L)e_t = \sum_{j=0}^{\infty} c_j e_{t-j}, \qquad \sum_{j=0}^{\infty} j|c_j| < \infty, \qquad C(1) \neq 0, \tag{7}$$

with $e_t = iid(0, \sigma_e^2)$ and $\mu_{2p} = E|u_t|^{2p} < \infty$ for some $p > 2$.

It is well known (e.g., Phillips and Solo, 1992, Thm. 3.4) that Assumption LP is sufficient for the partial sums $S_t = \sum_{s=1}^{t} u_s$ to satisfy the functional law $(1/\sqrt{n})S_{[n\cdot]} \to_{\text{d}} B(\cdot)$, where $B(\cdot)$ is Brownian motion with variance $\sigma^2 = \sigma_e^2 C(1)^2$. Further, extending the probability space as needed, the partial sum process S_t may be uniformly strongly approximated by a Brownian motion such as B, in the sense that

$$\sup_{1 \le t \le n} \left| \frac{S_{t-1}}{\sqrt{n}} - B\left(\frac{t-1}{n}\right) \right| = o_{a.s.}\left(\frac{1}{n^{(1/2)-(1/p)}}\right), \tag{8}$$

for some integer $p > 2$. Strong approximations such as (8) have been proved by many authors and are reviewed in Shorack and Wellner (1986) and Csörgö and Horváth (1993). A strong approximation justifying (8) in the case where u_t is a linear process is given in Phillips (2006) for time series data under Assumption LP. Akonom (1993) gave (8) with an $o_p(n^{-(1/2)+(1/p)})$ error under Assumption LP using the weaker moment requirement that $\mu_p = E|u_t|^p < \infty$ for some $p > 2$.

It seems likely that the results of the present paper may be extended under suitable conditions to allow for long memory in u_t. It has been shown by Wang, Lin, and Gulati (2003), for instance, that partial sums of long memory processes satisfy a strong approximation analogous to (8) in which the limit process B is replaced by a fractional Brownian motion. After suitable adjustments in convergence rates, that result may be used to extend some of the limit theory of the present paper, including (9) and Lemma 2.1. These extensions are not pursued here and are left for subsequent work.

Under Assumption LP, it follows by partial summation and by taking weak limits that for any $f \in C^1$

$$\frac{1}{\sqrt{n}} \sum_{s=1}^{n} f\left(\frac{s}{n}\right) u_s \to_d \int_0^1 f(r)\, dB(r) = N\left(0, \sigma^2 \int_0^1 f(r)^2\, dr\right). \tag{9}$$

Some related results hold when f is SV. In particular, we have the following lemma.

LEMMA 2.1. *If* $L(t)$ *satisfies Assumption SSV,* $\bar{L} = n^{-1} \sum_{t=1}^{n} L(t)$, *and* u_t *satisfies Assumption LP, then* *(i)* $(1/\sqrt{n}L(n))\sum_{t=1}^{n} L(t)u_t \to_d B(1) =_d N(0,\sigma^2)$ *as* $n \to \infty$.

(ii) $(1/\sqrt{n}L(n)\varepsilon(n))\sum_{t=1}^{n}(L(t) - \bar{L})u_t \to_d \int_0^1 (1 + \log r)\, dB(r) =_d N(0,\sigma^2)$ *as* $n \to \infty$.

(iii) $(1/\sqrt{n}\varepsilon(n)^j)\sum_{t=1}^{n}[(L(t)/L(n)) - 1]^j u_t \to_d \int_0^1 \log^j r dB(r) =_d N(0,\sigma^2(2j)!)$ *as* $n \to \infty$.

2.1. Heuristics

As shown in (60) in Section 7 and Lemma 7.2, one of the implications of Assumption SSV is that we have the following asymptotic representation of $L(t)$ for $t = nr$ with $r > 0$:

$$\frac{L(rn)}{L(n)} - 1 = \exp\{\varepsilon(n)\log r[1 + o(1)]\} - 1 = \varepsilon(n)\log r[1 + o(1)]. \tag{10}$$

Such a function may be called second-order SV (cf. de Haan and Resnick, 1996, who discuss second-order regular variation). For the sample mean $\bar{L}$, we have

$$\bar{L} = L(n) - L(n)\varepsilon(n) + o(L(n)\varepsilon(n)).$$

In consequence, the standardized sums that appear in (i)–(iii) of Lemma 2.1 have the approximate asymptotic forms

$$\frac{1}{\sqrt{n}L(n)}\sum_{t=1}^{n} L(t)u_t \sim \frac{1}{\sqrt{n}}\sum_{t=1}^{n} u_t,$$

$$\frac{1}{\sqrt{n}L(n)\varepsilon(n)}\sum_{t=1}^{n}(L(t) - \bar{L})u_t \sim \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(1 + \log\left(\frac{t}{n}\right)\right)u_t,$$

$$\frac{1}{\sqrt{n}\varepsilon(n)^j}\sum_{t=1}^{n}\left[\frac{L(t)}{L(n)} - 1\right]^j u_t \sim \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\log^j\left(\frac{t}{n}\right)u_t,$$

to which we may apply a standard central limit argument like that of (9). These cases indicate that, as far as first-order asymptotic theory is concerned, weighted means of u_t with arbitrary SV weights behave in a common way, at least up to a normalization factor that depends on the asymptotic form of the SV function and its corresponding ε-function. The "common" form that appears in these expressions is that of a logarithmic trend function $\log t$, whereas the influence of the particular SV function affects the normalization by way of $L(n)$ and $\varepsilon(n)$. This characteristic will be seen to apply more generally in regression asymptotics.

3. SIMPLE REGRESSION

We start with the simple regression model

$$y_s = \alpha + \beta L(s) + u_s \qquad s = a,\dots,n \quad \text{for some } a \geq 1, \tag{11}$$

where u_s satisfies Assumption LP. Initialization of (11) at some suitable finite integer a ensures that $L(s)$ is well defined for $s \geq a$ in cases such as $L(s) = 1/\log s$ where $L(1)$ is undefined. In such cases, $L(s)$ may be redefined as $L(s) = L(a)$ for $1 \leq s \leq a$ and some suitable finite a with no effect on subsequent results. Henceforth, it will be assumed that such adjustments have been made.

Let $\hat{\alpha}$ and $\hat{\beta}$ be the least squares regression coefficients. The limit behavior of these regression coefficients depends on that of the first and second sample moments

$$\bar{L} = \frac{1}{n}\sum_{s=1}^{n} L(s), \qquad \frac{1}{n}\sum_{s=1}^{n}(L(s)-\bar{L})^2 = \frac{1}{n}\sum_{s=1}^{n} L(s)^2 - \left(\frac{1}{n}\sum_{s=1}^{n} L(s)\right)^2. \tag{12}$$

The natural approach is to approximate these sample sums by an integral using Euler summation and then determine the asymptotic form of the resulting integrals as $n \to \infty$. Lemma 7.1 gives for the kth moment

$$\sum_{t=1}^{n} L(t)^k = \int_1^n L(t)^k\,dt + O(n^\delta), \tag{13}$$

where $\delta > 0$ is arbitrarily small, and Lemma 7.2 gives the following explicit asymptotic expansion:

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n} L(t)^k &= L(n)^k - kL(n)^k\varepsilon(n) + k^2L(n)^k\varepsilon(n)^2 + kL(n)^k\varepsilon(n)\eta(n)\\ &\quad - k^3[L(n)^k\varepsilon(n)^3 + 3L(n)^k\varepsilon(n)^2\eta(n) + L(n)^k\varepsilon(n)\eta(n)^2]\\ &\quad + o(L(n)^k[\varepsilon^3(n) + \varepsilon(n)^2\eta(n) + \varepsilon(n)\eta(n)^2]).\end{aligned} \tag{14}$$

The first two moments are then

$$\begin{aligned}\bar{L} &= L(n) - L(n)\varepsilon(n) + L(n)\varepsilon(n)^2 + L(n)\varepsilon(n)\eta(n)\\ &\quad - L(n)\varepsilon(n)^3 + 3L(n)\varepsilon(n)^2\eta(n) + L(n)\varepsilon(n)\eta(n)^2\\ &\quad + o(L(n)[\varepsilon^3(n) + \varepsilon(n)^2\eta(n) + \varepsilon(n)\eta(n)^2])\end{aligned} \tag{15}$$

and

$$\begin{aligned}\frac{1}{n}\sum_{t=1}^{n} L(t)^2 &= L(n)^2 - 2L(n)^2\varepsilon(n) + 4L(n)^2\varepsilon(n)^2 + 2L(n)^2\varepsilon(n)\eta(n)\\ &\quad - 8[L(n)^2\varepsilon(n)^3 + 3L(n)^2\varepsilon(n)^2\eta(n) + L(n)^2\varepsilon(n)\eta(n)^2]\\ &\quad + o(L(n)^2[\varepsilon^3(n) + \varepsilon(n)^2\eta(n) + \varepsilon(n)\eta(n)^2]),\end{aligned} \tag{16}$$

from which we deduce that

$$\frac{1}{n}\sum_{t=1}^{n}(L(t)-\bar{L})^2 = L(n)^2\varepsilon(n)^2\{1+o(1)\},$$

as in Lemma 7.3.

Then

$$\sqrt{n}L(n)\varepsilon(n)(\hat{\beta}-\beta) = \left[\frac{1}{nL(n)^2\varepsilon(n)^2}\sum_{t=1}^{n}(L(t)-\bar{L})^2\right]^{-1} \times \frac{1}{\sqrt{n}L(n)\varepsilon(n)}\sum_{t=1}^{n}(L(t)-\bar{L})u_t,$$

and

$$\begin{aligned}\sqrt{n}\varepsilon(n)(\hat{\alpha}-\alpha) &= \varepsilon(n)\frac{1}{\sqrt{n}}\sum_{t=1}^{n}u_t - \sqrt{n}L(n)\varepsilon(n)(\hat{\beta}-\beta)[1+(\varepsilon(n))] \\ &= -\sqrt{n}L(n)\varepsilon(n)(\hat{\beta}-\beta)+o_p(1). \qquad \textbf{(17)}\end{aligned}$$

The limit theory for the regression coefficients now follows directly from (17) and Lemma 2.1.

THEOREM 3.1. *If $L(t)$ satisfies Assumption SSV and u_t satisfies Assumption LP, then*

$$\begin{bmatrix}\sqrt{n}\varepsilon(n)(\hat{\alpha}-\alpha)\\ \sqrt{n}L(n)\varepsilon(n)(\hat{\beta}-\beta)\end{bmatrix} \to_d N\left(0,\sigma^2\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\right). \qquad \textbf{(18)}$$

Examples

(a) $L(s) = \log s$. This gives the semilogarithmic model. Here, $\varepsilon(n) = [1/(\log n)]$, $L(n)\varepsilon(n) = 1$, and (18) is

$$\begin{bmatrix}\dfrac{\sqrt{n}}{\log n}(\hat{\alpha}-\alpha)\\ \sqrt{n}(\hat{\beta}-\beta)\end{bmatrix} \to_d N\left(0,\sigma^2\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\right). \qquad \textbf{(19)}$$

This example also covers log periodogram analysis of long memory. In this case we have the regression

$$\log(I_X(\lambda_s)) = \hat{c} - 2\hat{d}\log\lambda_s + \text{residual}, \qquad s = 1,\ldots,m, \qquad \textbf{(20)}$$

where $I_X(\lambda_s)$ is the periodogram of a time series $(X_t)_{t=1}^n$ and $\lambda_s = 2\pi s/n$ are fundamental frequencies. The spectrum of X_t is assumed to have the local form $f_x(\lambda) \sim C\lambda^{-2d}$ for $\lambda \to 0+$, and, correspondingly, the regression (20) is taken over a band of frequencies that shrink to the origin, so that $(1/m) + (m/n) \to 0$. Then (20) has the alternate form

$$\begin{aligned}\log(I_X(\lambda_s)) &= \left(\hat{c} - 2\hat{d}\log\frac{2\pi}{n}\right) - 2\hat{d}\log s + \text{residual} \\ &= \hat{c}_n - 2\hat{d}\log s + \text{residual}, \qquad \textbf{(21)}\end{aligned}$$

where $\hat{c}_n = \hat{c} - 2\hat{d}\log(2\pi/n)$. Set $c = \log C$, $c_n = c - 2d\log(2\pi/n)$. The moment matrix of the regressors in (21) is asymptotically singular, just as in (1). Although the details of the central limit theory differ from Lemma 2.1 because of the properties of the residual terms in (20) (cf. Robinson, 1995; Hurvich et al., 1998), we nevertheless end up with a result analogous to (19) but with sample size m, namely,

$$\begin{bmatrix} \dfrac{\sqrt{m}}{\log m}(\hat{c}_n - c_n) \\ -2\sqrt{m}(\hat{d} - d) \end{bmatrix} \to_d N\left(0, \frac{\pi^2}{6}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right).$$

Because $\hat{c}_n - c_n = (\hat{c} - c) - 2(\hat{d} - d)\log(2\pi/n)$, we have

$$\frac{\sqrt{m}}{\log n}(\hat{c}_n - c_n) = \frac{\sqrt{m}}{\log n}(\hat{c} - c) + 2\sqrt{m}(\hat{d} - d) + O_p\left(\frac{1}{\log n}\right) = o_p(1),$$

from which we deduce that

$$\begin{bmatrix} \dfrac{\sqrt{m}}{\log n}(\hat{c} - c) \\ 2\sqrt{m}(\hat{d} - d) \end{bmatrix} \to_d N\left(0, \frac{\pi^2}{6}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right),$$

a result obtained by Robinson (1995, Thm. 3). The perfect negative asymptotic correlation between the estimates $\hat{c}$ and $\hat{d}$ induces a corresponding property between the estimates $\hat{C}$ and $\hat{d}$ of the original parameters appearing locally in $f_x(\lambda) \sim C\lambda^{-2d}$.

(b) $L(s) = 1/(\log s)$. This example arises when the regressor decays slowly. Here $\varepsilon(n) = -[1/(\log n)]$, $L(n)\varepsilon(n) = -[1/(\log^2 n)]$, and (18) is

$$\begin{bmatrix} \dfrac{\sqrt{n}}{\log n}(\hat{\alpha} - \alpha) \\ \dfrac{\sqrt{n}}{\log^2 n}(\hat{\beta} - \beta) \end{bmatrix} \to_d N\left(0, \sigma^2\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right). \tag{22}$$

(c) $L(s) = \log\log s$. Here, $\varepsilon(n) = [1/(\log\log n)][1/(\log n)]$, $L(n)\varepsilon(n) = [1/(\log n)]$, and (18) is

$$\begin{bmatrix} \dfrac{\sqrt{n}}{\log\log n \log n}(\hat{\alpha} - \alpha) \\ \dfrac{\sqrt{n}}{\log n}(\hat{\beta} - \beta) \end{bmatrix} \to_d N\left(0, \sigma^2\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right). \tag{23}$$

(d) $L(s) = [1/(\log\log s)]$. Here, $\varepsilon(n) = -[1/(\log\log n)][1/(\log n)]$, $L(n)\varepsilon(n) = -[1/(\log^2\log n)][1/(\log n)]$, and (18) is

$$\begin{bmatrix} \dfrac{\sqrt{n}}{\log\log n \log n}(\hat{\alpha}-\alpha) \\ \dfrac{\sqrt{n}}{\log^2\log n\log n}(\hat{\beta}-\beta) \end{bmatrix} \rightarrow_d N\left(0, \sigma^2\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right).$$

(e) $L(s) = \log^{\gamma} s$, $\gamma > 0$. In this case, $\varepsilon(n) = [\gamma/(\log n)]$, $L(n)\varepsilon(n) = \gamma\log^{\gamma-1} n$, and (18) is

$$\begin{bmatrix} \dfrac{\gamma\sqrt{n}}{\log n}(\hat{\alpha}-\alpha) \\ \gamma\sqrt{n}\log^{\gamma-1} n(\hat{\beta}-\beta) \end{bmatrix} \rightarrow_d N\left(0, \sigma^2\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right).$$

In all these cases the limit behavior is identical up to appropriate normalization of the coefficients, which is determined solely by L and its ε-function. Some intuition explaining the results is as follows. When $L(n) \to \infty$ as $n \to \infty$, the convergence rate of the slope coefficient, $\hat{\beta}$, exceeds that of the intercept, $\hat{\alpha}$, because the signal from the regressor $L(s)$ in (11) is stronger than that of a constant regressor. When $L(n) \to 0$ as $n \to \infty$, the convergence rate of $\hat{\beta}$ is less than that of $\hat{\alpha}$, because the signal from the regressor $L(s)$ is weaker than that of a constant regressor. The singularity in the asymptotic distribution arises because the regressor and the intercept are asymptotically collinear.

Some of the first-order asymptotic approximations implied by the preceding theory may not be good in finite samples. Refinements of the limit theory can be developed using asymptotic expansions in suitable powers of the SV functions. When $L(s) = 1/\log s$, for example, example (b) given previously is analyzed in Phillips and Sun (2003), where the following asymptotic expansion of the moment matrix of the regressors is developed in powers of $L(n)$:

$$\begin{aligned} &\frac{1}{n}\sum_{t=2}^{n}\{L(t)-\bar{L}\}^2 \\ &\quad = \frac{1}{\log^4 n}\left(1 + \frac{8}{\log n} + \frac{56}{\log^2 n} + \frac{408}{\log^3 n} + \frac{3,228}{\log^4 n} + \frac{28,032}{\log^5 n} + \frac{267,264}{\log^6 n}\right) \\ &\quad + O\left(\frac{1}{\log^{11} n}\right). \end{aligned} \tag{24}$$

This series (24) is an asymptotic refinement of (58) in Section 7 for the case where $L(s) = 1/\log s$ and shows that the error in such approximations is in powers of $1/\log n$ and therefore goes to zero rather slowly as $n \to \infty$. However, such expansions are not needed in practical work because expressions such as $\sum_{t=2}^{n}\{L(t)-\bar{L}\}^2$ may be calculated directly from the data. Phillips and Sun (2003) provide some alternate approximations based on the logarithmic integral in this case.

3.1. Standard Errors

These are computed by scaling the square root of the diagonal elements of the inverse of the second moment matrix with an estimate of σ^2 obtained from the regression residuals (either the sample variance, in the case where u_t is $iid(0,\sigma^2)$, or an estimate of $\sigma^2 = \sigma_e^2 C(1)^2$ obtained by kernel methods in the stationary time series case (7)). Using (13) and (14), we have

$$\sum_{s=1}^{n} \begin{bmatrix} 1 & L(s) \\ L(s) & L(s)^2 \end{bmatrix} = n \begin{bmatrix} 1 & L_{12}(n) \\ L_{12}(n) & L_{22}(n) \end{bmatrix} + O(n^{\eta}), \tag{25}$$

where

$$L_{12}(n) = L(n) - L(n)\varepsilon(n) + L(n)\varepsilon(n)^2 + - L(n)\varepsilon(n)\eta(n) + o(L(n)\varepsilon(n)[\eta(n) + \varepsilon(n)])$$

and

$$L_{22}(n) = L(n)^2 - 2L(n)^2\varepsilon(n) + 4L(n)^2\varepsilon(n)^2 - 2L(n)^2\varepsilon(n)\eta(n) + o(L(n)^2\varepsilon(n)[\eta(n) + \varepsilon(n)]).$$

Upon standardization with the diagonal matrix $D_n = diag(\sqrt{n}, \sqrt{n}L(n))$, (25) becomes

$$D_n^{-1} \sum_{s=1}^{n} \begin{bmatrix} 1 & L(s) \\ L(s) & L(s)^2 \end{bmatrix} D_n^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -\varepsilon(n) \\ -\varepsilon(n) & -2\varepsilon(n) \end{bmatrix} + o(\varepsilon(n)) \to \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \tag{26}$$

Similarly, upon inversion, we have

$$\left(\sum_{s=1}^{n} \begin{bmatrix} 1 & L(s) \\ L(s) & L(s)^2 \end{bmatrix} \right)^{-1} = \frac{1}{n \sum_{s=1}^{n} (L(s) - \bar{L})^2} \sum_{s=1}^{n} \begin{bmatrix} L(s)^2 & -L(s) \\ -L(s) & 1 \end{bmatrix}$$

$$= \frac{1}{nL(n)^2\varepsilon(n)^2 + o(nL(n)^2\varepsilon(n)[\eta(n) + \varepsilon(n)])} \begin{bmatrix} L_{22}(n) & L_{12}(n) \\ L_{12}(n) & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \dfrac{1}{n\varepsilon(n)^2} & -\dfrac{1}{nL(n)\varepsilon(n)^2} \\ -\dfrac{1}{nL(n)\varepsilon(n)^2} & \dfrac{1}{nL(n)^2\varepsilon(n)^2} \end{bmatrix} [1 + o(nL(n)\varepsilon(n)[\eta(n) + \varepsilon(n)])], \tag{27}$$

which, upon standardization by $F_n^{-1} = diag(\sqrt{n}\varepsilon(n), \sqrt{n}\varepsilon(n)L(n))$, gives

$$F_n^{-1}\left(\begin{bmatrix} n & \sum_{s=1}^{n} L(s) \\ \sum_{s=1}^{n} L(s) & \sum_{s=1}^{n} L(s)^2 \end{bmatrix}\right)^{-1} F_n^{-1} \to \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

It follows from these formulas that, in spite of the singularity in the limit matrix, the covariance matrix of the regression coefficients is consistently estimated as in conventional regression when an appropriate estimate s^2 of σ^2 is employed.

4. POLYNOMIAL REGRESSION IN $L(x)$

In this model the regressors are polynomials in the SSV function $L(s)$, and the data are generated by

$$y_s = \sum_{j=0}^{p} \beta_j L(s)^j + u_s = \beta' L_s + u_s, \tag{28}$$

where the regression error u_s satisfies Assumption LP. This model may be analyzed using the approach of the previous section. But, as the degree p increases in (28), the analysis becomes complicated because higher order expansions than (14) of the sample moments of $L(s)$ are needed to develop a complete asymptotic theory. An alternate approach is to rewrite the model (28) in a form wherein the moment matrix of the regressors has a full rank limit. The degeneracy in the new model, which has an array format, then passes from the data matrix to the coefficients and is simpler to analyze.

The process is first illustrated with model (1), which we can write in the form

$$\begin{aligned} y_s &= \alpha + \beta \log n + \beta \log \frac{s}{n} + u_s \\ &= \alpha_n + \beta \log \frac{s}{n} + u_s, \quad \text{say.} \end{aligned} \tag{29}$$

The regressors $\{1, \log(s/n)\}$ in (29) are not collinear. Writing $k(r) = [1, \log r]'$ and using standard manipulations, we obtain

$$\sqrt{n}\begin{bmatrix} \hat{\alpha}_n - \alpha_n \\ \hat{\beta} - \beta \end{bmatrix} \to_d N\left(0, \sigma^2 \left(\int_0^1 k(r)k(r)'\,dr\right)^{-1}\right) = N\left(0, \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}^{-1}\right).$$

Because $\hat{\alpha}_n - \alpha_n = \hat{\alpha} - \alpha + (\hat{\beta} - \beta)\log n$, we deduce that

$$\frac{\sqrt{n}}{\log n}(\hat{\alpha} - \alpha) = -\sqrt{n}(\hat{\beta} - \beta) + O_p\left(\frac{1}{\log n}\right),$$

which leads directly to the earlier result (19).

Extending this process to the model (28) gives the representation

$$y_s = \sum_{j=0}^{p} \beta_j \left\{ L(n) \left[\frac{L(s)}{L(n)} - 1 \right] + L(n) \right\}^j + u_s,$$

$$= \sum_{j=0}^{p} \beta_j L(n)^j \sum_{i=0}^{j} \binom{j}{i} \left[\frac{L(s)}{L(n)} - 1 \right]^i + u_s,$$

$$= \sum_{j=0}^{p} \alpha_{nj} \left[\frac{L(s)}{L(n)} - 1 \right]^j + u_s,$$

where

$$\alpha_{n0} = \sum_{j=0}^{p} \beta_j L(n)^j, \tag{30}$$

$$\alpha_{nk} = \sum_{j=k}^{p} \beta_j L(n)^j \binom{j}{k} \qquad k = 1, \ldots, p-1, \tag{31}$$

$$\alpha_{np} = \beta_p L(n)^p. \tag{32}$$

Define

$$K_{nj}\left(\frac{s}{n}\right) = \left[\frac{L(s)}{L(n)} - 1 \right]^j = \left[\frac{L\left(\frac{s}{n} n\right)}{L(n)} - 1 \right]^j, \qquad j = 0, 1, \ldots, p,$$

and the model (28) becomes

$$y_s = \sum_{j=0}^{p} \alpha_{nj} K_{nj}\left(\frac{s}{n}\right) + u_s := \alpha_n' K_n\left(\frac{s}{n}\right) + u_s. \tag{33}$$

Least squares estimation gives

$$\hat{\alpha}_n - \alpha_n = \left[\sum_{t=1}^{n} K_n\left(\frac{t}{n}\right) K_n\left(\frac{t}{n}\right)' \right]^{-1} \left[\sum_{t=1}^{n} K_n\left(\frac{t}{n}\right) u_t \right]. \tag{34}$$

The limit behavior of these coefficient estimates depends on that of the regressors $K_{nj}(t/n)$, and sample moment asymptotics for K_{nj} follow from that of its sample mean. Define the vector $K_n(t/n) = (K_{nj}(t/n))$ and the normalization matrix $D_{n\varepsilon} = diag[1, \varepsilon(n), \varepsilon(n)^2, \ldots, \varepsilon(n)^p]$.

THEOREM 4.1.

(i) If $L(t)$ satisfies Assumption SSV, then

$$\frac{1}{n} D_{n\varepsilon}^{-1} \sum_{t=1}^{n} K_n\left(\frac{t}{n}\right) \to \int_0^1 \ell_p(r)\, dr$$

$$= [1, \quad -1, \quad 2!, \quad -3!, \quad \ldots, \quad (-1)^p p!]',$$

where $\ell_p(r) = [1, \log r, \ldots, \log^p r]'$ and

$$\frac{1}{n} D_{n\varepsilon}^{-1} \sum_{t=1}^{n} K_n\left(\frac{t}{n}\right) K_n\left(\frac{t}{n}\right)' D_{n\varepsilon}^{-1}$$

$$\to \int_0^1 \ell_p(r)\ell_p(r)'\, dr$$

$$= \begin{bmatrix} 1 & -1 & 2! & -3! & \ldots & (-1)^p p! \\ -1 & 2! & -3! & 4! & \ldots & (-1)^{p+1}(p+1)! \\ 2! & -3! & 4! & -5! & \ldots & (-1)^{p+2}(p+2)! \\ -3! & 4! & -5! & 6! & \ldots & (-1)^{p+3}(p+3)! \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (-1)^p p! & (-1)^{p+1}(p+1)! & (-1)^{p+2}(p+2)! & (-1)^{p+3}(p+3)! & \ldots & (2p)! \end{bmatrix}, \quad \mathbf{(35)}$$

which is positive definite.

(ii) If $L(t)$ satisfies Assumption SSV and u_t satisfies Assumption LP, then

$$\sqrt{n} D_{n\varepsilon}[\hat{\alpha}_n - \alpha_n] \to_d N\left(0, \sigma^2 \left[\int_0^1 \ell_p(r)\ell_p(r)'\, dr\right]^{-1}\right). \quad \mathbf{(36)}$$

Next, we rewrite this limit distribution in terms of the original coefficients using relations (30)–(32). It transpires that only the final component, $\hat{\alpha}_{np}$, in $\hat{\alpha}_n$ (which translates to the component $\hat{\beta}_p$ in the original coordinates) determines the nondegenerate part of the limit theory for the full set of coefficients.

THEOREM 4.2. *If $L(t)$ satisfies Assumption SSV and u_t satisfies Assumption LP, then*

$$\sqrt{n}\varepsilon(n)^p D_{nL}(\hat{\beta} - \beta) = \mu_{p+1}\sqrt{n} L(n)^p \varepsilon(n)^p (\hat{\beta}_p - \beta_p)$$

$$+ o_p(1) \to_d N(0, v^{p+1,p+1} \mu_{p+1}\mu_{p+1}'),$$

where $D_{nL} = diag(1, L(n), \ldots, L(n)^p)$, $\mu_{p+1}' = [(-1)^p, (-1)^{p-1}\binom{p}{1}, \ldots, (-1)\binom{p}{p-1}, 1]$, and $v^{p+1,p+1} = (p!)^{-2}$ is the $p+1$th diagonal element of $[\int_0^1 \ell_p(r)\ell_p(r)'\, dr]^{-1}$.

The limit distribution of $\sqrt{n}\varepsilon(n)^p D_{nL}(\hat{\beta} - \beta)$ has a support given by the range of the vector μ_{p+1} and is therefore of dimension one. The variance matrix of $\hat{\beta}$ is given by

$$\frac{v^{p+1,p+1}}{n\varepsilon(n)^{2p}} D_{nL}^{-1} \mu_{p+1} \mu_{p+1}' D_{nL}^{-1}, \tag{37}$$

which, as we now show, is consistently estimated by the usual regression formula. The following result gives expressions for the asymptotic form of $L'L = \sum_{s=1}^{n} L_s L_s'$ and $(L'L)^{-1}$, showing that, indeed, (37) is the asymptotic form of $(L'L)^{-1}$.

THEOREM 4.3. *If $L(t)$ satisfies Assumption SSV, then*

(i)

$$L'L = nD_{nL} i_{p+1} i_{p+1}' D_{nL}[1 + o(1)], \tag{38}$$

where i_{p+1} is a $p+1$ vector with unity in each element.

(ii)

$$(L'L)^{-1} = \frac{e_{p+1}' \left[\int_0^1 \ell_p(r)\ell_p(r)'\, dr \right]^{-1} e_{p+1}}{n\varepsilon(n)^{2p}} \times D_{nL}^{-1} \mu_{p+1} \mu_{p+1}' D_{nL}^{-1}[1 + o(1)], \tag{39}$$

where $\ell_p(r)$ and μ_{p+1} are given in Theorems 4.1 and 4.2.

It follows from (39) that, in spite of the singularity in the limit matrix, the covariance matrix of the regression coefficients is consistently estimated as in conventional regression by $s^2(L'L)^{-1}$ whenever s^2 is a consistent estimate of σ^2.

5. REGRESSION WITH MULTIPLE SSV REGRESSORS

Multiple regression with different SV functions as regressors is also of some interest in applications. One such formulation is given in the example that appears later in this section and involves an SV growth component in conjunction with a trend decay component that slowly adjusts the intercept in the regression to a lower level. Such a model is relevant in empirical research where one wants to capture simultaneously two different opposing trends in the data. Such models can be analyzed by the methods of the previous section, with the SV regressors replacing the polynomials in a given function $L(s)$. We shall provide results for a model with two different regressors, which is the case of principal interest in practice and where our assumptions allow for a full treatment. We also briefly discuss the general case, where more structure is needed for a complete treatment.

Let $L_j(s)$ $(j = 1, 2)$ be SSV functions with corresponding ε- and η-functions ε_j and η_j $(j = 1, 2)$. We consider the two-variable regression model

$$y_s = \beta_0 + \beta_1 L_1(s) + \beta_2 L_2(s) + u_s = \beta' L_s + u_s, \quad \text{say,} \tag{40}$$

where the regression error u_s satisfies Assumption LP. An asymptotic theory of regression in this model is obtained by showing that (40) has an alternate, asymptotically equivalent, form involving a quadratic function of the simpler regressor $\log(s/n)$. Analysis similar to the previous section then applies.

Rewrite (40) as follows:

$$y_s = \beta_0 + \beta_1 L_1(n) + \beta_2 L_2(n) + \beta_1 L_1(n)\left[\frac{L_1\left(n\dfrac{s}{n}\right)}{L_1(n)} - 1\right] + \beta_2 L_2(n)\left[\frac{L_2\left(n\dfrac{s}{n}\right)}{L_2(n)} - 1\right] + u_s.$$

To transform the regressors in this version of the model, we note from Lemma 7.5 that L_j has a higher order representation in terms of its ε- and η-functions that has the asymptotic form

$$\frac{L_j(rn)}{L_j(n)} - 1 = \varepsilon_j(n)\log r + \frac{1}{2}\varepsilon_j(n)[\varepsilon_j(n) + \eta_j(n)]\log^2 r[1 + o(1)], \qquad r > 0. \tag{41}$$

Equation (41) shows L_j to be third-order SV in the sense that

$$\lim_{n\to\infty} \frac{\dfrac{\dfrac{L_j(rn)}{L_j(n)} - 1}{\varepsilon_j(n)} - \log r}{\dfrac{1}{2}[\varepsilon_j(n) + \eta_j(n)]} = \log^2 r, \qquad r > 0,$$

thereby extending the concept of second-order slow variation that appears in the earlier expression (10). Using the expansion (41) we write

$$\begin{aligned} y_s &= \beta_0 + \beta_1 L_1(n) + \beta_2 L_2(n) \\ &\quad + \beta_1 L_1(n)\varepsilon_1(n)\log\frac{s}{n} + \frac{1}{2}\beta_1\varepsilon_1(n)[\varepsilon_1(n) + \eta_1(n)]\log^2\frac{s}{n}[1 + o(1)] \\ &\quad + \beta_2 L_2(n)\varepsilon_2(n)\log\frac{s}{n} + \frac{1}{2}\beta_2\varepsilon_2(n)[\varepsilon_2(n) + \eta_2(n)]\log^2\frac{s}{n}[1 + o(1)] + u_s \\ &= \alpha_{n0} + \alpha_{n1}\log\left(\frac{s}{n}\right) + \alpha_{n2}\log^2\left(\frac{s}{n}\right)[1 + o(1)] + u_s, \quad \text{say,} \end{aligned} \tag{42}$$

giving a new form of the model with regressors that comprise a quadratic function in $\log(s/n)$. The new coefficients satisfy the system

$$\begin{bmatrix}\alpha_{n0}\\ \alpha_{n1}\\ \alpha_{n2}\end{bmatrix}=\begin{bmatrix}1 & L_1(n) & L_2(n)\\ 0 & L_1(n)\varepsilon_1(n) & L_2(n)\varepsilon_2(n)\\ 0 & \frac{1}{2}L_1(n)\varepsilon_1(n)[\varepsilon_1(n)+\eta_1(n)] & \frac{1}{2}L_2(n)\varepsilon_2(n)[\varepsilon_2(n)+\eta_2(n)]\end{bmatrix}\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix}$$

$$=\begin{bmatrix}1 & 1 & 1\\ 0 & \varepsilon_1(n) & \varepsilon_2(n)\\ 0 & \frac{1}{2}\varepsilon_1(n)[\varepsilon_1(n)+\eta_1(n)] & \frac{1}{2}\varepsilon_2(n)[\varepsilon_2(n)+\eta_2(n)]\end{bmatrix}$$

$$\times\begin{bmatrix}1 & 0 & 0\\ 0 & L_1(n) & 0\\ 0 & 0 & L_2(n)\end{bmatrix}\begin{bmatrix}\beta_0\\ \beta_1\\ \beta_2\end{bmatrix}. \tag{43}$$

For further asymptotic analysis, we impose the condition

$$\delta(n)=[\varepsilon_2(n)+\eta_2(n)]-[\varepsilon_1(n)+\eta_1(n)]\neq 0, \tag{44}$$

which is necessary if we are to solve (43) for the original coefficients in (40). If (44) does not hold, then the regressors L_1 and L_2 are collinear to the second order in (41). In that case, the situation is more complex—higher order representations are needed to develop an asymptotic theory, and rates of convergence need to be adjusted. The following result holds under (44), uses only the second-order form (41), and gives the limit theory for the original coefficients in (40).

THEOREM 5.1. *If $L(t)$ satisfies Assumption SSV, u_t satisfies Assumption LP, and $\delta(n)\neq 0$, then*

$$\sqrt{n}\,\delta(n)\begin{bmatrix}\varepsilon_{\min}(n)(\hat\beta_0-\beta_0)\\ \varepsilon_1(n)L_1(n)[\hat\beta_1-\beta_1]\\ \varepsilon_2(n)L_2(n)[\hat\beta_2-\beta_2]\end{bmatrix}\sim\begin{bmatrix}1_\varepsilon\\ -1\\ 1\end{bmatrix}\sqrt{n}[\hat\alpha_{n2}-\alpha_{n2}]$$

$$\to_d N\left(0,\frac{\sigma^2}{2!}\begin{bmatrix}1 & -1_\varepsilon & 1_\varepsilon\\ -1_\varepsilon & 1 & -1\\ 1_\varepsilon & -1 & 1\end{bmatrix}\right), \tag{45}$$

where

$$\varepsilon_{\min}(n) = \begin{cases} \varepsilon_2(n) & \textit{if } \varepsilon_2(n) = o(\varepsilon_1(n)) \\ \varepsilon_1(n) & \textit{if } \varepsilon_1(n) = o(\varepsilon_2(n)) \end{cases}$$

and

$$1_\varepsilon = \begin{cases} -1 & \textit{if } \varepsilon_2(n) = o(\varepsilon_1(n)) \\ 1 & \textit{if } \varepsilon_1(n) = o(\varepsilon_2(n)). \end{cases}$$

5.1. Discussion

1. Equation (42) indicates that multiple regression with different SV functions is asymptotically equivalent to polynomial regression on a logarithmic function. Theorem 5.1 shows that the outcome is analogous to that of a polynomial regression, but the rates of convergence are affected by the respective natures of the SV functions. The actual rate of convergence of the estimates depends not just on the asymptotic behavior of the functions $L_j(n)$ and their ε-functions but also on the divergence, $\delta(n)$, between the sum of the ε- and η-functions of the two regressors L_1 and L_2. In effect, the more divergent are the L_j asymptotically, then the faster the rate of convergence of the regression estimates.
2. The scaling factor $\varepsilon_{\min}(n)$ in (45) relates to the constant in the regression and determines that its rate of convergence is affected by that of the more slowly converging regression coefficient.
3. If $L_i(x) = \log x$ for some i then there is no second-order term in (41) and $\varepsilon_i(n) + \eta_i(n) = 0$ in that case. The first matrix in (43) is simpler in this case and can be made upper triangular by permuting coefficients if necessary.
4. Just as in the polynomial regression case, the limit distribution (45) is singular and has rank unity.

Example

The following example has iterated logarithmic growth, a trend decay component, and a constant regressor:

$$y_s = \beta_0 + \beta_1 \frac{1}{\log s} + \beta_2 \log\log s + u_s.$$

The secondary functions are $\varepsilon_1(n) = -[1/(\log n)]$, $\eta_1(n) = -[1/(\log n)]$, $\varepsilon_2(n) = [1/(\log\log n)][1/(\log n)]$, and $\eta_2(n) = -[1/(\log\log n)] - [1/(\log n)]$. Then

$$\varepsilon_1(n) + \eta_1(n) = -\frac{2}{\log n},$$

$$\varepsilon_2(n) + \eta_2(n) = -\frac{1}{\log\log n} + o\left(\frac{1}{\log\log n}\right),$$

$$\delta(n) = -\frac{1}{\log\log n} + o\left(\frac{1}{\log\log n}\right),$$

$$\varepsilon_{\min} = \varepsilon_2(n) = \frac{1}{\log\log n}\frac{1}{\log n}.$$

We deduce that

$$\frac{\sqrt{n}}{\log\log n}\begin{bmatrix} \dfrac{1}{\log\log n}\dfrac{1}{\log n}(\hat{\beta}_0 - \beta_0) \\ \dfrac{(-1)}{\log^2 n}[\hat{\beta}_1 - \beta_1] \\ \dfrac{1}{\log n}[\hat{\beta}_2 - \beta_2] \end{bmatrix} \sim \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \frac{\sqrt{n}}{\log n \log\log n}[\hat{\beta}_2 - \beta_2]$$

$$\to_d N\left(0, \frac{\sigma^2}{2!}\begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}\right).$$

The coefficient of the growth term converges fastest but at less than an $\sqrt{n}$ rate. The intercept converges next fastest, and finally the coefficient of the evaporating trend. All of these outcomes relate to the strength of the signal from the respective regressor.

5.2. The General Case

Consider the model

$$y_s = \sum_{j=0}^{p} \beta_j L_j(s) + u_s = \beta' L_s + u_s, \quad \text{say,} \tag{46}$$

where $L_0(s) = 1$. As in the preceding two-variable case, this model can be rewritten as

$$y_s = \sum_{j=0}^{p} \beta_j L_j(n) + \sum_{j=1}^{p} \beta_j L_j(n)\left[\frac{L_j\left(n\dfrac{s}{n}\right)}{L_j(n)} - 1\right] + u_s. \tag{47}$$

Assume that each L_j has a higher order representation extending (41) in terms of the following asymptotic expansion:

$$\frac{L_j(rn)}{L_j(n)} - 1 = \sum_{i=1}^{p-1} \varepsilon_{ji}(n) \log^i r + \varepsilon_{jp}(n) \log^p r[1 + o(1)], \qquad r > 0, \tag{48}$$

where $\varepsilon_{j1}(n) = \varepsilon_j(n)$ and

$$\varepsilon_{ji}(n) = o(\varepsilon_{ji-1}(n)), \tag{49}$$

for each j and each $i > 1$, so the coefficients, $\varepsilon_{ji}(n)$, in (48) decrease in order of magnitude as i increases. Such a higher order expansion can be developed under conditions analogous to Assumption SSV in which each function in the sequence $L, \varepsilon, \eta, \dots$ itself has a Karamata representation with an ε-function that is SSV. Applying (48) in (47), we obtain the transformed model

$$\begin{aligned} y_s &= \sum_{j=0}^{p} \beta_j L_j(n) + \sum_{j=1}^{p} \beta_j L_j(n) \\ &\quad \times \left\{ \sum_{i=1}^{p-1} \varepsilon_{ji}(n) \log^i \left(\frac{s}{n} \right) + \varepsilon_{jp}(n) \log^p \left(\frac{s}{n} \right) [1 + o(1)] \right\} + u_s \\ &= \alpha_{n0} + \sum_{i=1}^{p-1} \sum_{j=1}^{p} \beta_j L_j(n) \varepsilon_{ji}(n) \log^i \left(\frac{s}{n} \right) \\ &\quad + \sum_{j=1}^{p} \beta_j L_j(n) \varepsilon_{jp}(n) \log^p \left(\frac{s}{n} \right) [1 + o(1)] + u_s \\ &= \alpha_{n0} + \sum_{i=1}^{p-1} \alpha_{ni} \log^i \left(\frac{s}{n} \right) + \alpha_{np} \log^p \left(\frac{s}{n} \right) [1 + o(1)] + u_s. \end{aligned}$$

The coefficients in this system satisfy

$$\begin{aligned} \begin{bmatrix} \alpha_{n0} \\ \alpha_{n1} \\ \alpha_{n2} \\ \vdots \\ \alpha_{np} \end{bmatrix} &= \begin{bmatrix} 1 & 1 & \dots & 1 \\ 0 & \varepsilon_1(n) & \dots & \varepsilon_p(n) \\ 0 & \varepsilon_{12}(n) & \dots & \varepsilon_{p2}(n) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \varepsilon_{1p}(n) & \dots & \varepsilon_{pp}(n) \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & L_1(n) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & L_p(n) \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & \dots & 1 \\ 0 & 1 & \dots & 1 \\ 0 & \eta_{12}(n) & \dots & \eta_{p2}(n) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \eta_{1p}(n) & \dots & \eta_{pp}(n) \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & L_1(n)\varepsilon_1(n) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & L_p(n)\varepsilon_p(n) \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}, \end{aligned}$$

where

$$\eta_{ji}(n) = \frac{\varepsilon_{ji}(n)}{\varepsilon_j(n)} = o(1) \quad \text{as } n \to \infty.$$

Define

$$\Xi_n = \begin{bmatrix} 1 & \dots & 1 \\ \eta_{12}(n) & \dots & \eta_{p2}(n) \\ \vdots & \ddots & \vdots \\ \eta_{1p}(n) & \dots & \eta_{pp}(n) \end{bmatrix}$$

and note that, in view of (49), we have

$$\eta_{ji}(n) = o(\eta_{ji-1}(n)),$$

so that the final row ($i = p$) of Ξ_n has elements of the smallest order and the other rows decrease in magnitude as i increases. Then,

$$\Xi_n^{-1} = \frac{1}{\det \Xi_n} \begin{bmatrix} \eta^{11}(n) & \dots & \eta^{1p}(n) \\ \eta^{21}(n) & \dots & \eta^{2p}(n) \\ \vdots & \ddots & \vdots \\ \eta^{p1}(n) & \dots & \eta^{pp}(n) \end{bmatrix} = \frac{1}{\det \Xi_n} M_n, \quad \text{say,}$$

and, in view of the property of Ξ_n just mentioned, the first $p - 1$ columns of $M_n = \det(\Xi_n)\Xi_n^{-1}$ are of smaller order as $n \to \infty$ than the final column of M_n. (Indeed, the columns of M_n progressively increase in order of magnitude from left to right.) We therefore have

$$\det(\Xi_n) \begin{bmatrix} L_1(n)\varepsilon_1(n) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & L_p(n)\varepsilon_p(n) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}$$

$$= \begin{bmatrix} \eta^{11}(n) & \dots & \eta^{1p}(n) \\ \eta^{21}(n) & \dots & \eta^{2p}(n) \\ \vdots & \ddots & \vdots \\ \eta^{p1}(n) & \dots & \eta^{pp}(n) \end{bmatrix} \begin{bmatrix} \alpha_{n1} \\ \alpha_{n2} \\ \vdots \\ \alpha_{n2} \end{bmatrix} = \begin{bmatrix} \eta^{1p}(n) \\ \eta^{2p}(n) \\ \vdots \\ \eta^{pp}(n) \end{bmatrix} \alpha_{np}[1 + o_p(1)],$$

so that

$$\sqrt{n}\,\det(\Xi_n)\begin{bmatrix} \frac{L_1(n)\varepsilon_1(n)}{\eta^{1p}(n)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{L_p(n)\varepsilon_p(n)}{\eta^{pp}(n)} \end{bmatrix}\begin{bmatrix} \sqrt{n}(\hat{\beta}_1-\beta_1) \\ \vdots \\ \sqrt{n}(\hat{\beta}_p-\beta_p) \end{bmatrix}$$

$$\sim \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}[\sqrt{n}(\hat{\alpha}_{np}-\alpha_{np})] \to_d \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} N\left(0, \frac{\sigma^2}{(p!)^2}\right).$$

Turning to the intercept, we have $\alpha_{n0} = [1, L_1(n), \ldots, L_p(n)]\beta$. Define

$$\varepsilon_{\min} = \min_{j \le p} \frac{\varepsilon_j(n)}{\eta^{jj}(n)}$$

to be the ratio with the smallest order of magnitude as $n \to \infty$. Then, we have

$$\sqrt{n}(\hat{\beta}_0 - \beta_0) = \sqrt{n}(\hat{\alpha}_{n0} - \alpha_{n0}) - \sum_{j=1}^{p} L_j(n)\sqrt{n}(\hat{\beta}_j - \beta_j),$$

and scaling by $\det(\Xi_n)\varepsilon_{\min}$ and noting that $\det(\Xi_n)\varepsilon_{\min} = o(1)$ as $n \to \infty$, we deduce that

$$\begin{aligned} &\sqrt{n}\,\det(\Xi_n)\varepsilon_{\min}(\hat{\beta}_0 - \beta_0) \\ &\quad = \sqrt{n}\,\det(\Xi_n)\varepsilon_{\min}(\hat{\alpha}_{n0} - \alpha_{n0}) - \sum_{j=1}^{p} L_j(n)\sqrt{n}\,\det(\Xi_n)\varepsilon_{\min}(\hat{\beta}_j - \beta_j) \\ &\quad = o_p(1) - \sqrt{n}\,\frac{L_j(n)\varepsilon_j(n)}{\eta^{jj}(n)}\det(\Xi_n)(\hat{\beta}_j - \beta_j) \to_d N\left(0, \frac{\sigma^2}{(p!)^2}\right), \end{aligned}$$

giving the following result.

THEOREM 5.2. *If $L(t)$ satisfies Assumption SSV, u_t satisfies Assumption LP, and* $\det \Xi_n \neq 0$, *then*

$$\sqrt{n}\,\det(\Xi_n)\begin{bmatrix} \varepsilon_{\min} & 0 & \cdots & 0 \\ 0 & \frac{L_1(n)\varepsilon_1(n)}{\eta^{1p}(n)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{L_p(n)\varepsilon_p(n)}{\eta^{pp}(n)} \end{bmatrix}\begin{bmatrix} \sqrt{n}(\hat{\beta}_0-\beta_0) \\ \sqrt{n}(\hat{\beta}_1-\beta_1) \\ \vdots \\ \sqrt{n}(\hat{\beta}_p-\beta_p) \end{bmatrix}$$

$$\sim i_{p+1}\left[\sqrt{n}(\hat{\alpha}_{np}-\alpha_{np})\right] \to_d N\left(0, \frac{\sigma^2}{(p!)^2}\, i_{p+1} i_{p+1}'\right), \tag{50}$$

where

$$\varepsilon_{\min} = \min_{j \le p} \frac{\varepsilon_j(n)}{\eta^{jj}(n)}$$

is the ratio with the smallest order of magnitude as $n \to \infty$.

In (50) the scale coefficients $L_j(n)\varepsilon_j(n)/\eta^{jp}(n)$ and also $\varepsilon_{\min}$ are implicitly signed. That is, the elements $\varepsilon_j(n)/\eta^{jj}(n)$ may have positive or negative signs. In consequence, because the signs are built into the normalization factor, the covariance matrix of the limit distribution,

$$\sigma^2 v^{p+1,p+1} i_{p+1} i'_{p+1} = \frac{\sigma^2}{(p!)^2} i_{p+1} i'_{p+1},$$

displays perfect positive correlation among the elements of the standardized vector in the limit.

6. NONLINEAR TREND REGRESSION

In the nonlinear trend model (2), let u_s satisfy Assumption LP, let $\theta_0 = (\beta_0, \gamma_0)$ be the true values of the parameters, and assume that (β_0, γ_0) lies in the interior of the parameter space $\Theta = [0, b] \times [-\frac{1}{2}, c]$ where $0 < b, c < \infty$. Wu (1981, Exmp. 4, pp. 507, 509) considered the case where u_s is $iid(0, \sigma^2 > 0)$ and noted that the model satisfies his conditions for strong consistency of the least squares estimator $\hat{\theta} = (\hat{\beta}, \hat{\gamma})$ but not his conditions for asymptotic normality. There are two reasons for the failure: (i) the Hessian requires different standardizations for the parameters β and γ (whereas Wu's approach uses a common standardization); and (ii) the Hessian is asymptotically singular because of the asymptotic collinearity of the functions s^{γ_0} and $s^{\gamma_0} \log s$ that appear in the score (whereas Wu's theory requires the variance matrix to have a positive definite limit). Both issues are addressed by a version of the methods given earlier in the paper designed to deal with extremum estimation problems.

Setting $Q_n(\beta, \gamma) = \sum_{s=1}^n (y_s - \beta s^\gamma)^2$, the estimates $(\hat{\beta}, \hat{\gamma})$ solve the extremum problem

$$(\hat{\beta}, \hat{\gamma}) = \arg\min_{\beta, \gamma} Q_n(\beta, \gamma)$$

and satisfy the first-order conditions

$$S_n(\hat{\beta}, \hat{\gamma}) = 0, \tag{51}$$

where

$$S_n(\theta) = -\sum_{s=1}^n \begin{bmatrix} s^\gamma \\ \beta s^\gamma \log s \end{bmatrix} (y_s - \beta s^\gamma).$$

Expanding $S_n(\theta)$ about $S_n(\theta_0)$, we have

$$0 = S_n(\theta_0) + H_n(\theta_0)(\hat{\theta} - \theta_0) + [H_n^* - H_n(\theta_0)](\hat{\theta} - \theta_0), \tag{52}$$

where the Hessian H_n^* is evaluated at mean values between θ_0 and $\hat{\theta}$ and

$$H_n(\theta) = \sum_{s=1}^{n} \begin{bmatrix} s^{2\gamma} & \beta s^{2\gamma} \log s - u_s s^{\gamma} \log s \\ & \quad - (\beta_0 s^{\gamma_0} - \beta s^{\gamma}) s^{\gamma} \log s \\ \beta s^{2\gamma} \log s - u_s s^{\gamma} \log s & \beta^2 s^{2\gamma} \log^2 s - u_s \beta s^{\gamma} \log^2 s \\ \quad - (\beta_0 s^{\gamma_0} - \beta s^{\gamma}) s^{\gamma} \log s & \quad - (\beta_0 s^{\gamma_0} - \beta s^{\gamma}) \beta s^{\gamma} \log^2 s \end{bmatrix}.$$

The following lemmas assist in characterizing the asymptotic behavior of these quantities.

LEMMA 6.1 *Let L be an SV function satisfying Assumption SSV and suppose u_s satisfies Assumption LP. Let C_n be a diagonal matrix all of whose elements diverge to ∞ as $n \to \infty$. Define $N_n^0 = \{\theta \in \Theta : \|C_n(\theta - \theta_0)\| \le 1\}$ to be a shrinking neighborhood of θ_0 for any point θ_0 in the interior of a compact parameter space Θ. Let $f(r;\theta) \in C^2$ over $(r,\theta) \in [0,1] \times \Theta$ and let the derivatives $f_\theta = \partial f/\partial\theta$, $f_r = \partial f/\partial r$, $f_{r\theta} = \partial^2 f/\partial\theta\partial r$ be dominated as follows:*

$$\sup_{\theta \in N_n^0} |f_\theta(r;\theta)| \le g_1(r), \qquad \sup_{\theta \in N_n^0} |f_r(r;\theta)| \le g_2(r), \qquad \sup_{\theta \in N_n^0} |f_{r\theta}(r;\theta)| \le g_3(r)$$

by functions $\{g_i : i = 1,2,3\}$, which are absolutely integrable over $[0,1]$. Then

$$\frac{1}{\sqrt{n}L(n)} \sum_{s=1}^{n} f\left(\frac{s}{n};\theta\right) L(s) u_s \to_d \int_0^1 f(r;\theta_0)\, dB(r) = N\left(0, \sigma^2 \int_0^1 f(r;\theta_0)^2\, dr\right) \tag{53}$$

uniformly over $\theta \in N_n^0$.

LEMMA 6.2. *Suppose u_s satisfies Assumption LP and let the true parameter vector $\theta_0 = (\beta_0, \gamma_0)$ lie in the interior of $\Theta = [0,b] \times [-\frac{1}{2}, c]$ where $0 < b$, $c < \infty$. Define the normalization matrices*

$$D_n = diag[n^{\gamma_0 + (1/2)}, n^{\gamma_0 + (1/2)} \log n],$$

$$F_n = \frac{1}{\log n} D_n = diag\left[\frac{n^{\gamma_0 + (1/2)}}{\log n}, n^{\gamma_0 + (1/2)}\right].$$

Define $C_n = D_n / n^{\delta}$ for some small positive $\delta \in (0, \gamma_0 + \frac{1}{2})$ and the following shrinking neighborhood of θ_0:

$$N_n^0 = \{\theta \in \Theta : \|C_n(\theta - \theta_0)\| \le 1\}.$$

Then

(i)

$$D_n^{-1} S_n(\theta_0) \to_d -\int_0^1 \begin{bmatrix} r^{\gamma_0} \\ \beta_0 r^{\gamma_0} \end{bmatrix} dB(r) = N\left(0, \frac{\sigma^2}{2\gamma_0+1} \begin{bmatrix} 1 & \beta_0 \\ \beta_0 & \beta_0^2 \end{bmatrix}\right), \tag{54}$$

(ii)

$$D_n^{-1} H_n(\theta_0) D_n^{-1} \to_p \frac{1}{2\gamma_0+1} \begin{bmatrix} 1 & \beta_0 \\ \beta_0 & \beta_0^2 \end{bmatrix}, \tag{55}$$

(iii)

$$\lambda_{\min}(F_n^{-1} H_n(\theta_0) F_n^{-1}) = O(\log n) \to \infty,$$

(iv)

$$\begin{aligned} &F_n H_n(\theta_0)^{-1} F_n \\ &= \frac{(2\gamma^0+1)^3}{2\beta_0^2} \begin{bmatrix} \beta_0^2 - \frac{2\beta_0^2}{\log n}\frac{1}{(2\gamma^0+1)} + \frac{\beta_0^2}{\log^2 n}\frac{2}{(2\gamma^0+1)^2} & -\beta_0 + \frac{\beta_0}{\log n}\frac{1}{(2\gamma^0+1)} \\ -\beta_0 + \frac{\beta_0}{\log n}\frac{1}{(2\gamma^0+1)} & 1 \end{bmatrix} \\ &\quad + o_p\left(\frac{1}{n^\delta}\right) \\ &\to_p \frac{(2\gamma^0+1)^3}{\beta_0^2} \begin{bmatrix} \beta_0^2 & -\beta_0 \\ -\beta_0 & 1 \end{bmatrix}, \end{aligned}$$

(v)

$$\sup_{\theta \in N_n^0} \| C_n^{-1} [H_n(\theta) - H_n(\theta_0)] C_n^{-1} \| = o_p(1).$$

Remarks.

(a) Part (i) of Lemma 6.2 reveals that the order of convergence of the first member of (52), the score $S_n(\theta_0)$, is determined by the scaling factor D_n^{-1}. However, from part (ii), the Hessian matrix under the same standardization by D_n^{-1} evidently has a singular limit as $n \to \infty$, which prevents the application of the usual approach of solving (52) to find a limit theory for a standardized form of $(\hat{\theta} - \theta_0)$. Part (iv) shows that upon standardization by F_n, rather than D_n, the inverse Hessian matrix converges but also has a singular limit.

(b) Part (v) is useful in showing that, after rescaling, the third term of (52) can be neglected in the asymptotic behavior of $\hat{\theta} - \theta_0$.
(c) As the following result shows, the appropriate scaling factor for (52) is the matrix F_n^{-1}, not D_n^{-1}, even though $D_n^{-1} S_n(\theta_0)$ is $O_p(1)$.

THEOREM 6.3. *In the model (2), let u_s satisfy Assumption LP and let the true parameter vector $\theta_0 = (\beta_0, \gamma_0)$ lie in the interior of $\Theta = [0, b] \times [-\frac{1}{2}, c]$ where $0 < b, c < \infty$. Then, the least squares estimator $\hat{\theta} = (\hat{\beta}, \hat{\gamma})$ is consistent and has the following limit distribution as $n \to \infty$:*

$$F_n(\hat{\theta} - \theta_0) \to_d (2\gamma^0 + 1)^3 \begin{bmatrix} 1 \\ -1/\beta_0 \end{bmatrix} \int_0^1 r^{\gamma_0} \left[\log r + \frac{1}{2\gamma_0 + 1} \right] dB(r)$$

$$= \begin{bmatrix} 1 \\ -1/\beta_0 \end{bmatrix} N(0, \sigma^2 (2\gamma^0 + 1)^3).$$

Remarks.

(a) The estimator $\hat{\theta}$ has a convergence rate that is slower by a factor of $\log n$ than that of the score $S_n(\theta_0)$. The reason is that the (conventionally standardized) Hessian $D_n^{-1} H_n(\theta_0) D_n^{-1}$ has an inverse that diverges at the rate $\log^2 n$ and this divergence slows down the convergence rate of the estimator. Both the score and the Hessian need to be rescaled to achieve the appropriate convergence rate for $\hat{\theta}$. With the new scaling we have

$$0 = F_n^{-1} S_n(\theta_0) + F_n^{-1} H_n(\theta_0) F_n^{-1} F_n(\hat{\theta} - \theta_0)$$

$$+ F_n^{-1} [H_n^* - H_n(\theta_0)] F_n^{-1} F_n(\hat{\theta} - \theta_0),$$

and then

$$F_n(\hat{\theta} - \theta_0) = -[I + (F_n H_n(\theta_0)^{-1} F_n) F_n^{-1} [H_n^* - H_n(\theta_0)] F_n^{-1}]^{-1}$$

$$\times (F_n H_n(\theta_0)^{-1} F_n) F_n^{-1} S_n(\theta_0).$$

From Lemma 6.2(iv), the matrix $F_n H_n(\theta_0)^{-1} F_n = O_p(1)$ and has a singular limit. Also, as shown in (98) in Section 8, the matrix $F_n^{-1}[H_n^* - H_n(\theta_0)] F_n^{-1}$ is $o_p(1)$. Then

$$F_n(\hat{\theta} - \theta_0) = -(F_n H_n(\theta_0)^{-1} F_n) F_n^{-1} S_n(\theta_0) + o_p(1),$$

from which the limit distribution follows. Interestingly, even though the individual elements of $F_n^{-1} S_n(\theta_0)$ diverge, the relevant linear combination $(F_n H_n(\theta_0)^{-1} F_n) F_n^{-1} S_n(\theta_0)$ is $O_p(1)$.

(b) The variance matrix for $\hat{\theta}$ is singular but is consistently estimated by $s^2 H_n(\hat{\theta})^{-1}$, where s^2 is a consistent estimator of σ^2, because

$$F_n H_n(\hat{\theta})^{-1} F_n = \frac{(2\gamma^0+1)^3}{\beta_0^2} \begin{bmatrix} \beta_0^2 - \dfrac{2\beta_0^2}{\log n}\dfrac{1}{(2\gamma^0+1)} + \dfrac{\beta_0^2}{\log^2 n}\dfrac{2}{(2\gamma^0+1)^2} & -\beta_0 + \dfrac{\beta_0}{\log n}\dfrac{1}{(2\gamma^0+1)} \\ -\beta_0 + \dfrac{\beta_0}{\log n}\dfrac{1}{(2\gamma^0+1)} & 1 \end{bmatrix} + o_p\left(\frac{1}{\sqrt{n}}\right).$$

7. TECHNICAL SUPPLEMENT

LEMMA 7.1 (Averages of SV functions). *If $L(t)$ satisfies Assumption SSV, then for $B \geq 1$*

$$\sum_{t=B}^{n} L(t) = \int_B^n L(t)\,dt + O(n^\delta) \quad \text{as } n \to \infty,$$

where $\delta > 0$ is arbitrarily small.

Proof. Using Euler summation (e.g., Knopp, 1990, p. 521) we have

$$\sum_{t=B}^{n} L(t) = \int_B^n L(t)\,dt + \frac{1}{2}[L(B) + L(n)] + \int_B^n \left\{t - [t] - \frac{1}{2}\right\} L'(t)\,dt. \tag{56}$$

Because

$$\frac{tL'(t)}{L(t)} = \varepsilon(t) \to 0 \quad \text{and} \quad \frac{L(t)}{t^\delta} \to 0,$$

for all $\delta > 0$, we may choose a constant C such that for all $t \geq C$ and any $\delta > 0$

$$\left|\varepsilon(t)\frac{L(t)}{t^\delta}\right| < 1.$$

Then, the final term in (56) may be bounded as follows:

$$\begin{aligned} \left|\int_B^n \left\{t - [t] - \frac{1}{2}\right\} L'(t)\,dt\right| &\leq \frac{1}{2}\int_B^n \frac{1}{t}\,|\varepsilon(t)L(t)|\,dt \\ &\leq \frac{1}{2}\int_C^n \frac{1}{t^{1-\eta}}\,dt + \frac{1}{2}\left|\int_B^C |\varepsilon(t)L(t)|\frac{1}{t}\,dt\right| \\ &= \frac{1}{2\eta}[t^\delta]_C^n + O(1) \\ &= O(n^\delta). \end{aligned}$$

It follows that

$$\sum_{t=B}^{n} L(t) = \int_B^n L(t)\,dt + O(n^{\delta} + L(n)) = \int_B^n L(t)\,dt + O(n^{\delta}),$$

for any $\delta > 0$ as $n \to \infty$. ■

LEMMA 7.2.

$$\frac{1}{n}\sum_{t=1}^{n} L(t)^k = L(n)^k - kL(n)^k\varepsilon(n) + k^2 L(n)^k\varepsilon(n)^2 + kL(n)^k\varepsilon(n)\eta(n)$$
$$- k^3[L(n)^k\varepsilon(n)^3 + 3L(n)^k\varepsilon(n)^2\eta(n) + L(n)^k\varepsilon(n)\eta(n)^2]$$
$$+ o(L(n)^k[\varepsilon(n)^3 + \varepsilon(n)^2\eta(n) + \varepsilon(n)\eta(n)^2]).$$

Proof. From Assumption SSV(b), $|\varepsilon(x)|$ is SSV and

$$\varepsilon(x) = c_{\varepsilon}\exp\left(\int_1^x \frac{\eta(t)}{t}\,dt\right),$$

where $\eta(n) \to 0$ as $n \to \infty$. Like ε, $\eta \in C^{\infty}$, and if $|\eta|$ is SSV

$$\frac{x^m \eta^{(m)}(x)}{\eta(x)} \to 0.$$

Then, using repeated integration by parts and the formulas $L'(t) = L(t)\varepsilon(t)/t$ and $\varepsilon'(t) = \varepsilon(t)\eta(t)/t$, we find

$$\int_1^n L(t)^k\,dt$$
$$= [tL(t)^k]_1^n - k\int_1^n tL(t)^k \frac{\varepsilon(t)}{t}\,dt$$
$$= nL(n)^k - k\int_1^n L(t)^k\varepsilon(t)\,dt + O(1)$$
$$= nL(n)^k - k[tL(t)^k\varepsilon(t)]_1^n + k\int_1^n t\left[kL(t)^k\frac{\varepsilon(t)^2}{t} + L(t)^k\varepsilon(t)\frac{\eta(t)}{t}\right]dt$$
$$+ O(1)$$
$$= nL(n)^k - knL(n)^k\varepsilon(n) + k\int_1^n [kL(t)^k\varepsilon(t)^2 + L(t)^k\varepsilon(t)\eta(t)]\,dt + O(1)$$

$$
\begin{aligned}
&= nL(n)^k - knL(n)^k\varepsilon(n) + k^2 nL(n)^k\varepsilon(n)^2 + knL(n)^k\varepsilon(n)\eta(n) \\
&\quad - k^2\int_1^n t\left[kL(t)^k\,\frac{\varepsilon(t)^3}{t} + L(t)^k 2\varepsilon(t)^2\,\frac{\eta(t)}{t}\right]dt \\
&\quad - k\int_1^n t\left[kL(t)^k\,\frac{\varepsilon(t)^2}{t}\,\eta(t) + L(t)^k\varepsilon(t)\,\frac{\eta(t)^2}{t} + L(t)^k\varepsilon(t)\eta'(t)\right]dt \\
&\quad + O(1) \\
&= nL(n)^k - knL(n)^k\varepsilon(n) + k^2 nL(n)^k\varepsilon(n)^2 + knL(n)^k\varepsilon(n)\eta(n) \\
&\quad - k^3[nL(n)^k\varepsilon(n)^3 + nL(n)^k 2\varepsilon(n)^2\eta(n)] \\
&\quad - k^2[nL(n)^k\varepsilon(n)^2\eta(n) + nL(n)^k\varepsilon(n)\eta(n)^2 + nL(n)^k\varepsilon(n)\eta'(n)] \\
&\quad + o(nL(n)^k[\varepsilon(n)^3 + \varepsilon(n)^2\eta(n) + \varepsilon(n)\eta(n)^2 + \varepsilon(n)\eta'(n)]) \\
&= nL(n)^k - knL(n)^k\varepsilon(n) + k^2 nL(n)^k\varepsilon(n)^2 + knL(n)^k\varepsilon(n)\eta(n) \\
&\quad - k^3[nL(n)^k\varepsilon(n)^3 + 3nL(n)^k\varepsilon(n)^2\eta(n) + nL(n)^k\varepsilon(n)\eta(n)^2 \\
&\qquad\quad + nL(n)^k\varepsilon(n)\eta'(n)] \\
&\quad + o(nL(n)^k[\varepsilon(n)^3 + \varepsilon(n)^2\eta(n) + \varepsilon(n)\eta(n)^2]) \\
&= nL(n)^k - knL(n)^k\varepsilon(n) + k^2 nL(n)^k\varepsilon(n)^2 + knL(n)^k\varepsilon(n)\eta(n) \\
&\quad - k^3[nL(n)^k\varepsilon(n)^3 + 3nL(n)^k\varepsilon(n)^2\eta(n) + nL(n)^k\varepsilon(n)\eta(n)^2] \\
&\quad + o(nL(n)^k[\varepsilon(n)^3 + \varepsilon(n)^2\eta(n) + \varepsilon(n)\eta(n)^2]),
\end{aligned}
$$

because $n\eta'(n)/\eta(n) = o(1)$ and $\eta(n)^2 = o(\varepsilon(n))$ as $n \to \infty$, thereby giving the stated result. ■

Example

In the logarithmic case $L(t) = \log t$, and $\varepsilon(t) = [1/(\log t)]$ and $\eta(t) = -[1/(\log t)]$. Lemma 7.2 then gives the expansion

$$
\begin{aligned}
\int_1^n \log^k t dt &= n\log^k n - kn\log^{k-1} n + k^2 n\log^{k-2} n - kn\log^{k-2} n \\
&\quad + o(n\log^{k-2} n) \\
&= n\log^k n - kn\log^{k-1} n + k(k-1)n\log^{k-2} n + o(n\log^{k-2} n),
\end{aligned}
\tag{57}
$$

whereas successive integration by parts gives the exact result

$$\int_1^n \log^k t dt = n \sum_{j=0}^{k} (-k)_j \log^{k-j} n,$$

so that the expansion in (57) is accurate to the third order.

LEMMA 7.3.

$$\frac{1}{n}\int_1^n [L(t) - \bar{L}]^2 \, dt = L(n)^2 \varepsilon(n)^2 \{1 + o(1)\}.$$

Proof. Applying the expansion from Lemma 7.2 and using expressions (15) and (16), we get

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n} (L(t) - \bar{L})^2 \\
&= \frac{1}{n}\int_1^n L(t)^2 \, dt - \left(\frac{1}{n}\int_1^n L(t)\, dt\right)^2 \\
&= -2L(n)^2\varepsilon(n) + 4L(n)^2\varepsilon(n)^2 + 2L(n)^2\varepsilon(n)\eta(n) \\
&\quad - 8[L(n)^2\varepsilon(n)^3 + 3L(n)^2\varepsilon(n)^2\eta(n) + L(n)^2\varepsilon(n)\eta(n)^2] \\
&\quad - \{L(n)^2\varepsilon(n)^2 + L(n)^2\varepsilon(n)^2\eta(n)^2 - 2L(n)^2\varepsilon(n)\} \\
&\quad - \{2L(n)^2\varepsilon(n)^2 + 2L(n)^2\varepsilon(n)\eta(n)\} \\
&\quad - \{-2L(n)^2\varepsilon(n)^3 + 6L(n)^2\varepsilon(n)^2\eta(n) + 2L(n)^2\varepsilon(n)\eta(n)^2\} \\
&= L(n)^2\varepsilon(n)^2 - 8[L(n)^2\varepsilon(n)^3 + 3L(n)^2\varepsilon(n)^2\eta(n) + L(n)^2\varepsilon(n)\eta(n)^2] \\
&\quad - \{L(n)^2\varepsilon(n)^2\eta(n)^2 - 2L(n)^2\varepsilon(n)^3 + 6L(n)^2\varepsilon(n)^2\eta(n) \\
&\qquad + 2L(n)^2\varepsilon(n)\eta(n)^2\} \\
&\quad + o(L(n)^2[\varepsilon^3(n) + \varepsilon(n)^2\eta(n) + \varepsilon(n)\eta(n)^2]) \\
&= L(n)^2\varepsilon(n)^2\{1 + o(1)\}, \qquad \textbf{(58)}
\end{aligned}$$

because $\eta(n)^2 = o(\varepsilon(n))$. ■

LEMMA 7.4. *If $L(t)$ is SV and satisfies Assumption SSV and (6), then*

$$\int_0^1 \left[\frac{L(rn)}{L(n)} - 1\right]^k dr = (-1)^k k!\, \varepsilon(n)^k [1 + o(1)],$$

as $n \to \infty$.

Proof. In view of Assumption SSV, we have

$$\log \frac{L(rn)}{L(n)} = -\int_{rn}^{n} \frac{\varepsilon(t)}{t}\, dt, \tag{59}$$

and, because $|\varepsilon(t)|$ is SV, it follows by Karamata's theorem (e.g., BGT, Prop. 1.5.9a, p. 26) that for all $r > 0$

$$\begin{aligned}\int_{rn}^{n} \frac{\varepsilon(t)}{t}\, dt &= \varepsilon(n) \int_{rn}^{n} \frac{dt}{t}[1 + o(1)] \\ &= -\varepsilon(n) \log r[1 + o(1)] \quad \text{as } n \to \infty.\end{aligned}$$

Then

$$\frac{L(rn)}{L(n)} - 1 = \exp\{\varepsilon(n)\log r[1 + o(1)]\} - 1 = \varepsilon(n)\log r[1 + o(1)]. \tag{60}$$

The function L is second-order SV (for second-order regular variation, see de Haan and Resnick, 1996), in the sense that

$$\lim_{n\to\infty} \frac{\dfrac{L(rn)}{L(n)} - 1}{\varepsilon(n)} = \log r, \qquad r > 0.$$

Integration by parts gives

$$\int_0^1 \log^k r dr = (-1)^k k!, \tag{61}$$

and so

$$\begin{aligned}\int_0^1 \left[\frac{L(rn)}{L(n)} - 1\right]^k dr &= \varepsilon(n)^k \int_0^1 \log^k r dr[1 + o(1)] \\ &= (-1)^k \varepsilon(n)^k k![1 + o(1)],\end{aligned} \tag{62}$$

giving the stated result. ■

LEMMA 7.5. *If $L(t)$ satisfies Assumption SSV, then for all $r > 0$*

$$\frac{L(rn)}{L(n)} - 1 = \varepsilon(n)\log r + \frac{1}{2}\varepsilon(n)[\varepsilon(n) + \eta(n)]\log^2 r + o(\varepsilon(n)\eta(n) + \varepsilon(n)^2), \tag{63}$$

as $n \to \infty$.

Proof. Because both L and ε are SSV functions we have both (59) and

$$\log \frac{\varepsilon(rn)}{\varepsilon(n)} = -\int_{rn}^{n} \frac{\eta(t)}{t}\, dt,$$

and, as in (60), we get for ε

$$\frac{\varepsilon(rn)}{\varepsilon(n)} = 1 + \eta(n)\log r + o(\eta(n)).$$

Then

$$\begin{aligned}
\log \frac{L(rn)}{L(n)} &= -\int_{rn}^{n} \frac{\varepsilon(t)}{t}\, dt = -\varepsilon(n)\int_{rn}^{n} \frac{\varepsilon\left(n\frac{t}{n}\right)}{\varepsilon(n)} \frac{dt}{t} \\
&= -\varepsilon(n)\int_{rn}^{n}\left\{1 + \eta(n)\log\frac{t}{n} + o(\eta(n))\right\}\frac{dt}{t} \\
&= \varepsilon(n)\log r - \varepsilon(n)\eta(n)\int_{r}^{1} \log s\, \frac{ds}{s} + o(\varepsilon(n)\eta(n)) \\
&= \varepsilon(n)\log r + \frac{1}{2}\,\varepsilon(n)\eta(n)\log^2 r + o(\varepsilon(n)\eta(n)),
\end{aligned}$$

and we deduce that

$$\begin{aligned}
\frac{L(rn)}{L(n)} - 1 &= \exp\left\{\varepsilon(n)\log r + \frac{1}{2}\,\varepsilon(n)\eta(n)\log^2 r + o(\varepsilon(n)\eta(n))\right\} - 1 \\
&= \varepsilon(n)\log r + \frac{1}{2}\,\varepsilon(n)[\varepsilon(n) + \eta(n)]\log^2 r + o(\varepsilon(n)\eta(n) + \varepsilon(n)^2),
\end{aligned}$$

as stated. ∎

Example

$L(n) = [1/(\log n)]$, $\varepsilon(n) = -[1/(\log n)]$, and $\eta(n) = -[1/(\log n)]$. Then, by direct expansion we have for large n

$$\begin{aligned}
\frac{L(rn)}{L(n)} - 1 &= \frac{-\log r}{\log r + \log n} = \frac{-\log r}{\log n}\left[1 + \frac{\log r}{\log n}\right]^{-1} \\
&= -\frac{\log r}{\log n}\sum_{j=0}^{\infty}(-1)^j\left(\frac{\log r}{\log n}\right)^j,
\end{aligned}$$

which agrees with the third-order expansion given in (63).

LEMMA 7.6.

$$\int_1^n \log^k t dt = n \sum_{j=0}^{k} (-k)_j \log^{k-j} n,$$

where $(-k)_j = (-k)(-k+1)\dots(-k+j-1)$.

Proof. This follows by successive integration by parts. ∎

LEMMA 7.7.

(i) $\sum_{j=0}^{p-k-1} \binom{p}{p-j}\binom{p-j}{k}(-1)^{p-j} = (-1)^{p-k+1}\binom{p}{k}$.

(ii) $\sum_{j=k}^{p}(-1)^{p-j}\binom{p}{j}\binom{j}{k} = 0$.

Proof. Both parts follow by direct calculation. ∎

LEMMA 7.8.

(i)

$$\int_0^1 \ell_p(r)\ell_p(r)'\,dr = H_p F_p^2 H_p',$$

where

$$H_p = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & -2 & 1 & 0 & \dots & 0 & 0 \\ -1 & 3 & -3 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ (-1)^{p-1} & (-1)^{p}\binom{p-1}{1} & (-1)^{p+1}\binom{p-1}{2} & (-1)^{p+2}\binom{p-1}{3} & & 1 & 0 \\ (-1)^{p} & (-1)^{p+1}\binom{p}{1} & (-1)^{p+2}\binom{p}{2} & (-1)^{p+3}\binom{p}{3} & \dots & (-1)^{2p-1}\binom{p}{p-1} & 1 \end{bmatrix}$$

and

$$F_p = diag[1, 1, 2!, 3!, \dots, (p-1)!, p!].$$

(ii) $\det[\int_0^1 \ell_p(r)\ell_p(r)'\,dr] = \Pi_{j=1}^{p}(j!)^2$.

(iii) $([\int_0^1 \ell_p(r)\ell_p(r)'\,dr]^{-1})_{p+1,p+1} = 1/(p!)^2$.

Proof. Note that the (i,j)th element of the matrix $\int_0^1 \ell_p(r)\ell_p(r)'\,dr$ is $(-1)^{i+j-2}(i+j-2)!$. Consider the (i,j)th element of the matrix product $H_p F_p^2 H_p'$ and let $j = k \le i$. By direct calculation and using the representation

$$\binom{a}{\ell} = (-1)^{\ell}\frac{(-b)_{\ell}}{\ell!}, \qquad (-b)_{\ell} = (-b)(-b+1)\dots(-b+\ell-1),$$

we find that this element is

$$\sum_{m=0}^{(i-1)\wedge(j-1)} (-1)^{i-1+m} \binom{i-1}{m} (i-1)!(j-1)! \binom{j-1}{m} (-1)^{j-1+m}$$

$$= (-1)^{i+k}(i-1)!(k-1)! \sum_{m=0}^{k-1} (-1)^{2m} \frac{(1-i)_m (1-k)_m}{m!(1)_m}$$

$$= (-1)^{i+k}(i-1)!(k-1)!\, {}_2F_1(1-i, 1-k, 1; 1), \tag{64}$$

where ${}_2F_1(a,b,c;z) = \sum_{j=0}^{\infty} ((a)_j (b)_j / j!(c)_j) z^j$ is the hypergeometric function. Noting that the series terminates (because $1-k$ is zero or a negative integer) and applying the summation formula (e.g., Erdélyi, 1953, p. 61),

$${}_2F_1(a,b,c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)},$$

where Γ is the gamma function, (64) reduces to

$$(-1)^{i+k}(i-1)!(k-1)! \frac{\Gamma(i+k-1)}{\Gamma(i)\Gamma(k)} = (-1)^{i+k-2}(i+k-2)!,$$

giving the required result and part (i). Parts (ii) and (iii) follow directly. ∎

8. PROOFS

Proof of Lemma 2.1. *Part (i).* By partial summation we have

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} L(t) u_t = L(n) \frac{S_n}{\sqrt{n}} - \frac{1}{\sqrt{n}} \sum_{t=1}^{n} [L(t) - L(t-1)] S_{t-1}, \tag{65}$$

where $S_t = \sum_{s=1}^{t} u_s$. So

$$\frac{1}{\sqrt{n}L(n)} \sum_{t=1}^{n} L(t) u_t = \frac{S_n}{\sqrt{n}} - \frac{1}{\sqrt{n}L(n)} \sum_{t=1}^{n} [L(t) - L(t-1)] S_{t-1}$$

$$= \frac{S_n}{\sqrt{n}} - \frac{1}{L(n)} \sum_{t=1}^{n} \left[L\left(n\frac{t}{n}\right) - L\left(n\frac{t-1}{n}\right) \right] \frac{S_{t-1}}{\sqrt{n}}. \tag{66}$$

We now use the embedding of the standardized partial sum $S_{t-1}/\sqrt{n}$ in Brownian motion given in equation (8) following Assumption LP, namely,

$$\sup_{1 \le t \le n} \left| \frac{S_{t-1}}{\sqrt{n}} - B\left(\frac{t-1}{n}\right) \right| = o_{a.s.}\left(\frac{1}{n^{(1/2)-(1/p)}} \right).$$

Then

$$\frac{1}{L(n)}\sum_{t=1}^{n}\left[L\left(n\frac{t}{n}\right)-L\left(n\frac{t-1}{n}\right)\right]\frac{S_{t-1}}{\sqrt{n}}$$
$$=\frac{1}{L(n)}\int_0^1 B(r)\,dL(nr)+o_{a.s.}\left(\frac{1}{n^{(1/2)-(1/p)}}\right). \quad \textbf{(67)}$$

Next $(1/L(n))\int_0^1 B(r)\,dL(nr)$ has mean zero and variance

$$\frac{2}{L(n)^2}\int_0^1\int_0^r sdL(ns)\,dL(nr). \quad \textbf{(68)}$$

Observe that

$$\int_0^r sdL(ns)=\int_0^r nsL'(ns)\,ds=\int_0^r L(ns)\varepsilon(ns)\,ds$$
$$=\frac{1}{n}\int_0^{nr}L(t)\varepsilon(t)\,dt=\frac{1}{n}[nrL(nr)\varepsilon(nr)+o(nrL(nr)\varepsilon(nr))]. \quad \textbf{(69)}$$

For the last equality, note that $L(t)\varepsilon(t)$ is (up to sign) on SSV function. We can then use Karamata's theorem, namely, that for $\alpha > -1$ and an SV function ℓ, we have the asymptotic equivalence

$$\int_a^x t^{\alpha}\ell(t)\,dt \sim \frac{x^{\alpha+1}}{\alpha+1}\,\ell(x) \quad \text{as } x\to\infty \quad \textbf{(70)}$$

(e.g., BGT, Prop. 1.5.8, p. 26), setting $\alpha = 0$ to obtain (69). Using (69) in (68), the dominant term is

$$\frac{2}{L(n)^2}\int_0^1 rL(nr)\varepsilon(nr)\,dL(nr)=\frac{2}{L(n)^2}\int_0^1 nrL'(nr)L(nr)\varepsilon(nr)\,dr$$
$$=\frac{2}{L(n)^2}\int_0^1 L(nr)^2\varepsilon(nr)^2\,dr$$
$$=\frac{2}{nL(n)^2}\int_0^n L(t)^2\varepsilon(t)^2\,dt$$
$$=2\varepsilon(n)^2+o(\varepsilon(n)^2)=o(1),$$

by applying (70) again. It follows that

$$\frac{1}{L(n)}\int_0^1 B(r)\,dL(nr)=o_p(1), \quad \textbf{(71)}$$

as $n \to \infty$. We deduce from (66), (67), and (71) that

$$\frac{1}{\sqrt{n}L(n)} \sum_{t=1}^{n} L(t)u_t = \frac{S_n}{\sqrt{n}} + o_p(1) \to_d N(0,\sigma^2).$$

Part (ii). We have

$$\sum_{t=1}^{n} (L(t) - \bar{L})u_t = \sum_{t=1}^{n} L(t)u_t - \bar{L}S_n.$$

Lemma 7.2 gives

$$\bar{L} = L(n) - L(n)\varepsilon(n) + L(n)\varepsilon(n)^2 - L(n)\varepsilon(n)\eta(n) + o(L(n)\varepsilon(n)\eta(n)),$$

and using (66) and (67)

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} L(t)u_t = L(n)\frac{S_n}{\sqrt{n}} - \int_0^1 B(r)\,dL(nr) + o_{a.s.}\left(\frac{1}{n^{(1/2)-(1/p)}}\right),$$

so that

$$\begin{aligned}
&\frac{1}{\sqrt{n}} \sum_{t=1}^{n} (L(t) - \bar{L})u_t \\
&\quad = L(n)\frac{S_n}{\sqrt{n}} - \int_0^1 B(r)\,dL(nr) + o_{a.s.}\left(\frac{1}{n^{(1/2)-(1/p)}}\right) \\
&\qquad - [L(n) - L(n)\varepsilon(n) + L(n)\varepsilon(n)^2 - L(n)\varepsilon(n)\eta(n) \\
&\qquad\quad + o(L(n)\varepsilon(n)\eta(n))]\frac{S_n}{\sqrt{n}} \\
&\quad = -\int_0^1 B(r)\,dL(nr) + o_{a.s.}\left(\frac{1}{n^{(1/2)-(1/p)}}\right) + L(n)\varepsilon(n)\frac{S_n}{\sqrt{n}} \\
&\qquad + O_p(L(n)\varepsilon(n)^2) \\
&\quad = -\int_0^1 \frac{B(r)}{r}\frac{L'(nr)nr}{L(nr)} L(nr)\,dr + L(n)\varepsilon(n)\int_0^1 dB(r) + O_p(L(n)\varepsilon(n)^2) \\
&\quad = -\int_0^1 \frac{B(r)}{r}\varepsilon(nr)L(nr)\,dr + L(n)\varepsilon(n)\int_0^1 dB(r) + O_p(L(n)\varepsilon(n)^2).
\end{aligned} \tag{72}$$

Now, in view of the local law of the iterated logarithm for Brownian motion, we have

$$\limsup_{r\to 0} \frac{B(r)}{\sqrt{2r \log\log 1/r}} = 1.$$

So, as in (70), we have

$$\int_0^1 \frac{B(r)}{r} \varepsilon(nr)L(nr)\,dr = \varepsilon(n)L(n)\int_0^1 \frac{B(r)}{r}\,dr[1+o_p(1)]$$
$$= -\varepsilon(n)L(n)\int_0^1 (\log r)\,dB(r)[1+o_p(1)]. \quad \textbf{(73)}$$

It follows from (73) that (72) is

$$\int_0^1 (1+\log r)\,dB(r) + o_p(1) \to_d N\left(0,\sigma^2\int_0^1 (1+\log r)^2\,dr\right) =_d N(0,\sigma^2),$$

as stated.

Part (iii). Start by considering $n^{-1/2}\sum_{t=1}^n K_{nj}(t/n)u_t$. By partial summation and the strong approximation in equation (8) following Assumption LP, we obtain, as in (67),

$$\frac{1}{\sqrt{n}}\sum_{t=1}^n\left[\frac{L(t)}{L(n)}-1\right]^j u_t = -\sum_{t=1}^n \frac{S_{t-1}}{\sqrt{n}}\Delta\left[\frac{L\left(n\frac{t}{n}\right)}{L(n)}-1\right]^j$$
$$= -\int_0^1 B(r)\,d\left[\frac{L(nr)}{L(n)}-1\right]^j + o_{a.s.}\left(\frac{1}{n^{(1/2)-(1/p)}}\right)$$
$$= \int_0^1\left[\frac{L(nr)}{L(n)}-1\right]^j dB(r) + o_{a.s.}\left(\frac{1}{n^{(1/2)-(1/p)}}\right).$$

From (60) we have

$$\frac{L(rn)}{L(n)}-1 = \exp\{\varepsilon(n)\log r[1+o(1)]\}-1 = \varepsilon(n)\log r[1+o(1)],$$

so that

$$\int_0^1\left[\frac{L(nr)}{L(n)}-1\right]^j dB(r) = \varepsilon(n)^j\int_0^1 \log^j r dB(r)[1+o(1)].$$

Thus,

$$\frac{1}{\sqrt{n}\varepsilon(n)^j}\sum_{t=1}^n\left[\frac{L(t)}{L(n)}-1\right]^j u_t \to_d \int_0^1 \log^j r dB(r) =_d N\left(0,\sigma^2\int_0^1 \log^{2j} r dr\right)$$
$$= N(0,\sigma^2(2j)!),$$

as required. ■

Proof of Theorem 4.1. Using Euler summation with $f(t) = [(L(t)/L(n)) - 1]^j$ we obtain, as in Lemma 7.1,

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n} K_{nj}\left(\frac{t}{n}\right) \\
&= \frac{1}{n}\sum_{t=1}^{n}\left[\frac{L(t)}{L(n)} - 1\right]^j \\
&= \frac{1}{n}\int_1^n \left[\frac{L(t)}{L(n)} - 1\right]^j dt + \frac{1}{2n}\{f(1) + f(n)\} \\
&\quad + \frac{j}{n}\int_1^n \left\{t - [t] - \frac{1}{2}\right\}\left[\frac{L(t)}{L(n)} - 1\right]^{j-1} \frac{L'(t)}{L(t)}\, dt \\
&= \frac{1}{n}\int_1^n \left[\frac{L(t)}{L(n)} - 1\right]^j dt + O\left(\frac{1}{n^{1-\delta}}\right) \\
&= \int_{1/n}^1 \left[\frac{L(rn)}{L(n)} - 1\right]^j dr + O\left(\frac{1}{n^{1-\delta}}\right) \\
&= \int_0^1 \left[\frac{L(rn)}{L(n)} - 1\right]^j dr + O\left(\frac{1}{n^{1-\delta}}\right),
\end{aligned} \tag{74}$$

for arbitrarily small $\delta > 0$, in view of (6). Hence, from (62) in the proof of Lemma 7.4, we have

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n} K_{nj}\left(\frac{t}{n}\right) &= \varepsilon(n)^j \int_0^1 \log^j r dr[1 + o(1)] + O\left(\frac{1}{n^{1-\delta}}\right) \\
&= (-1)^j j!\, \varepsilon(n)^j [1 + o(1)],
\end{aligned}$$

so that

$$\frac{1}{n\varepsilon(n)^j}\sum_{t=1}^{n} K_{nj}\left(\frac{t}{n}\right) \to \int_0^1 \log^j r dr = (-1)^j j!,$$

from which the stated limit results follow. The matrix $\int_0^1 \ell_p(r)\ell_p(r)'\, dr$ is positive definite because

$$\int_0^1 [a'\ell_p(r)]^2\, dr = 0$$

implies $a'\ell_p(r) = 0$ for all r, which implies $a = 0$, and part (i) is established.

To prove part (ii), we note by Lemma 2.1(iii) that

$$\frac{1}{\sqrt{n}\varepsilon(n)^j}\sum_{t=1}^{n}K_{nj}\left(\frac{t}{n}\right)u_t \to_d \int_0^1 \log^j r dB(r),$$

and proceeding in the same way as in the proof of that lemma but with an arbitrary linear combination of the preceding elements for $j = 0,1,\dots,p$, we get

$$\begin{aligned}\sum_{j=0}^{p}\frac{b_j}{\sqrt{n}\varepsilon(n)^j}\sum_{t=1}^{n}K_{nj}\left(\frac{t}{n}\right)u_t &= \sum_{j=0}^{p}b_j\int_0^1 \log^j r dB(r)[1+o_{a.s.}(1)]\\ &\to_d \sum_{j=0}^{p}b_j\int_0^1 \log^j r dB(r).\end{aligned}$$

By the Cramér–Wold device, we deduce that

$$\frac{1}{\sqrt{n}}D_{n\varepsilon}^{-1}\sum_{t=1}^{n}K_n\left(\frac{t}{n}\right)u_t \to_d \int_0^1 \ell_p(r)\,dB(r), \qquad \ell_p(r) = (1,\log r,\dots,\log^p r). \tag{75}$$

Then, from (34), (35), and (75) we obtain

$$\begin{aligned}\sqrt{n}D_{n\varepsilon}[\hat{\alpha}_n - \alpha_n] &= \left[\frac{1}{n}D_{n\varepsilon}^{-1}\sum_{t=1}^{n}K_n\left(\frac{t}{n}\right)K_n\left(\frac{t}{n}\right)'D_{n\varepsilon}^{-1}\right]^{-1}\\ &\quad\times\left[\frac{1}{\sqrt{n}}D_{n\varepsilon}^{-1}\sum_{t=1}^{n}K_n\left(\frac{t}{n}\right)u_t\right]\to_d N(0,\sigma^2V^{-1}).\end{aligned}$$ ■

Proof of Theorem 4.2. From (32) and (36), we get for the final coefficient

$$\sqrt{n}L(n)^p\varepsilon(n)^p(\hat{\beta}_p - \beta_p) = \sqrt{n}\varepsilon(n)^p[\hat{\alpha}_{np} - \alpha_{np}] \to_d N(0,v^{p+1,p+1}),$$

where $V^{-1} = (v^{i,j})$ and $V = \int_0^1 \ell_p(r)\ell_p(r)'\,dr$. A calculation (see Lemma 7.8(iii)) gives the final diagonal element of the inverse matrix V^{-1},

$$v^{p+1,p+1} = \frac{1}{(p!)^2}.$$

For the next coefficient, we have

$$\begin{aligned}\alpha_{np-1} &= \binom{p-1}{p-1}\beta_{p-1}L(n)^{p-1} + \binom{p}{p-1}\beta_p L(n)^p\\ &= \beta_{p-1}L(n)^{p-1} + p\beta_p L(n)^p,\end{aligned}$$

and so

$$\beta_{p-1}L(n)^{p-1} = \alpha_{np-1} - p\beta_p L(n)^p,$$

leading to

$$\begin{aligned}
&\sqrt{n}L(n)^{p-1}\varepsilon(n)^p(\hat{\beta}_{p-1} - \beta_{p-1}) \\
&\quad = \sqrt{n}\varepsilon(n)^p(\hat{\alpha}_{np-1} - \alpha_{np-1}) - \binom{p}{p-1}\sqrt{n}L(n)^p\varepsilon(n)^p(\hat{\beta}_p - \beta_p) \\
&\quad = O_p(\varepsilon(n)) - \binom{p}{p-1}\sqrt{n}L(n)^p\varepsilon(n)^p(\hat{\beta}_p - \beta_p) \\
&\quad = -\binom{p}{p-1}\sqrt{n}L(n)^p\varepsilon(n)^p(\hat{\beta}_p - \beta_p) + o_p(1) \\
&\quad = -\binom{p}{p-1}\sqrt{n}\varepsilon(n)^p[\hat{\alpha}_{np} - \alpha_{np}] + o_p(1) \\
&\quad \to_d -\binom{p}{p-1}N(0, v^{p+1,p+1}).
\end{aligned}$$

Next, for $k = p - 2$ we have

$$\beta_{p-2}L(n)^{p-2} = \alpha_{np-2} - \left[\binom{p-1}{p-2}\beta_{p-1}L(n)^{p-1} + \binom{p}{p-2}\beta_p L(n)^p\right],$$

so that

$$\begin{aligned}
&\sqrt{n}L(n)^{p-2}\varepsilon(n)^p(\hat{\beta}_{p-2} - \beta_{p-2}) \\
&\quad = \sqrt{n}\varepsilon(n)^p(\hat{\alpha}_{np-2} - \alpha_{np-2}) \\
&\qquad - \binom{p-1}{p-2}L(n)^{p-1}\sqrt{n}\varepsilon(n)^p(\hat{\beta}_{p-1} - \beta_{p-1}) \\
&\qquad - \binom{p}{p-2}L(n)^p\sqrt{n}\varepsilon(n)^p(\hat{\beta}_p - \beta_p) \\
&\quad = O(\varepsilon(n)^2) + \left[\binom{p-1}{p-2}\binom{p}{p-1} - \binom{p}{p-2}\right]L(n)^p\sqrt{n}\varepsilon(n)^p(\hat{\beta}_p - \beta_p) \\
&\quad = O(\varepsilon(n)^2) + \left[(p-1)p - \frac{1}{2!}p(p-1)\right]L(n)^p\sqrt{n}\varepsilon(n)^p(\hat{\beta}_p - \beta_p) \\
&\quad = \binom{p}{p-2}L(n)^p\sqrt{n}\varepsilon(n)^p(\hat{\beta}_p - \beta_p) + o_p(1).
\end{aligned}$$

More generally, proceeding in this way for $p-1>k\geq 0$ (under the convention that $\binom{j}{0}=1$), we have

$$\alpha_{nk}=\sum_{j=k}^{p}\beta_j L(n)^j\binom{j}{k}=\binom{k}{k}\beta_k L(n)^k+\binom{k+1}{k}\beta_{k+1}L(n)^{k+1}$$
$$+\cdots+\binom{p}{k}\beta_p L(n)^p, \tag{76}$$

so that

$$\beta_k L(n)^k=\alpha_{nk}-\left[\binom{k+1}{k}\beta_{k+1}L(n)^{k+1}+\cdots+\binom{p}{k}\beta_p L(n)^p\right].$$

We establish by induction (for decreasing k) that

$$\sqrt{n}L(n)^k\varepsilon(n)^p(\hat{\beta}_k-\beta_k)=(-1)^{p-k}\binom{p}{k}\sqrt{n}L(n)^p\varepsilon(n)^p(\hat{\beta}_p-\beta_p)+o_p(1). \tag{77}$$

We have already shown (77) to be valid for $k=p-1$ and $p-2$. Assume that it is valid for $k+1$. Then, using (76) and Lemma 7.6 we have

$$\sqrt{n}L(n)^k\varepsilon(n)^p(\hat{\beta}_k-\beta_k)$$
$$=\sqrt{n}\varepsilon(n)^p(\hat{\alpha}_{nk}-\alpha_{nk})$$
$$-\left[\binom{k+1}{k}L(n)^{k+1}\varepsilon(n)^p(\hat{\beta}_{k+1}-\beta_{k+1})\right.$$
$$\left.+\cdots+\binom{p}{k}L(n)^p\varepsilon(n)^p(\hat{\beta}_p-\beta_p)\right]$$
$$=O(\varepsilon(n)^{p-k})-\left[(-1)^{p-k-1}\binom{p}{k+1}\binom{k+1}{k}\right.$$
$$\left.+\cdots+(-1)\binom{p}{p-1}\binom{p-1}{k}+\binom{p}{k}\right]$$
$$\times\sqrt{n}L(n)^p\varepsilon(n)^p(\hat{\beta}_p-\beta_p)$$
$$=O(\varepsilon(n)^{p-k})-\sum_{j=0}^{p-k-1}\binom{p}{p-j}\binom{p-j}{k}(-1)^{p-j}\sqrt{n}L(n)^p\varepsilon(n)^p(\hat{\beta}_p-\beta_p)$$
$$=(-1)^{p-k}\binom{p}{k}\sqrt{n}L(n)^p\varepsilon(n)^p(\hat{\beta}_p-\beta_p)+o_p(1),$$

showing that the result holds for k also.

Equation (77) gives an asymptotic correspondence between the elements of the least squares estimate $\hat{\beta}$ and its final component $\hat{\beta}_p$ that has the form

$$\sqrt{n}\varepsilon(n)^p D_{nL}(\hat{\beta}-\beta) = \mu_{p+1}\sqrt{n}L(n)^p\varepsilon(n)^p(\hat{\beta}_p-\beta_p) + o_p(1),$$

where $D_{nL} = diag(1, L(n), \ldots, L(n)^p)$ and $\mu'_{p+1} = \left[(-1)^p, (-1)^{p-1}\binom{p}{1}, \ldots, (-1)\binom{p}{p-1}, 1\right]$. We deduce that

$$\sqrt{n}\varepsilon(n)^p D_{nL}(\hat{\beta}-\beta) \to_d N(0, v^{p+1,p+1}\mu_{p+1}\mu'_{p+1}),$$

giving the stated result. The explicit formula $v^{p+1,p+1} = 1/(p!)^2$ follows from Lemma 7.8 (iii). ■

Proof of Theorem 4.3. First, transform the regressor space in (28) as follows:

$$y_s = \beta' L_s + u_s = \beta' J_n J_n^{-1} L_s + u_s = a_n' X_s + u_s, \tag{78}$$

where

$$J_n' = \begin{bmatrix} 1 & L(n) & L(n)^2 & \ldots & L(n)^{p-1} & L(n)^p \\ 0 & L(n)\varepsilon(n) & \binom{2}{1}L(n)^2\varepsilon(n) & \ldots & \binom{p-1}{1}L(n)^{p-1}\varepsilon(n) & \binom{p}{1}L(n)^p\varepsilon(n) \\ 0 & 0 & L(n)^2\varepsilon(n)^2 & \ldots & \binom{p-1}{2}L(n)^{p-1}\varepsilon(n) & \binom{p}{2}L(n)^p\varepsilon(n)^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & L(n)^{p-1}\varepsilon(n)^{p-1} & \binom{p}{p-1}L(n)^p\varepsilon(n)^{p-1} \\ 0 & 0 & 0 & \ldots & 0 & L(n)^p\varepsilon(n)^p \end{bmatrix}$$

$$= E_n H' D_{nL}, \quad \text{say},$$

and $E_n = diag[1, \varepsilon(n), \varepsilon(n)^2, \ldots, \varepsilon(n)^p]$, $D_{nL} = diag[1, L(n), L(n)^2, \ldots, L(n)^p]$, and

$$H' = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 & 1 \\ 0 & 1 & \binom{2}{1} & \ldots & \binom{p-1}{1} & \binom{p}{1} \\ 0 & 0 & 1 & \ldots & \binom{p-1}{2} & \binom{p}{2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 1 & \binom{p}{p-1} \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{bmatrix}.$$

In (78) $a_n = J_n'\beta = E_n\alpha_n$ where α_n is the parameter vector in (33) whose elements α_{nj} are given in (30)–(32). Because $X_s = J_n^{-1}L_s = E_n^{-1}H^{-1}D_{nL}^{-1}L_s$, we may rewrite (78) as

$$y_s = a_n'E_n^{-1}K_n\left(\frac{s}{n}\right) + u_s.$$

In view of (60), the vector $E_n^{-1}K_n(s/n)$ has elements

$$\frac{1}{\varepsilon(n)^j}K_{nj}\left(\frac{s}{n}\right) = \frac{1}{\varepsilon(n)^j}\left[\frac{L\left(\frac{s}{n}n\right)}{L(n)} - 1\right]^j = \log^j\left(\frac{s}{n}\right)[1+o(1)],$$

and so

$$\begin{aligned} a_n'E_n^{-1}K_n\left(\frac{s}{n}\right) &= \beta'J_nE_n^{-1}K_n\left(\frac{s}{n}\right) = \beta'J_n\ell_p\left(\frac{s}{n}\right)[1+o(1)] \\ &= \beta'D_{nL}HE_n\ell_p\left(\frac{s}{n}\right)[1+o(1)]. \end{aligned}$$

The sample second moment matrix, $L'L$, of the regressors can now be written as

$$\begin{aligned} L'L &= J_n\left[\sum_{s=1}^{n}\ell_p\left(\frac{s}{n}\right)\ell_p\left(\frac{s}{n}\right)'\right]J_n'[1+o(1)] \\ &= nJ_n\left[\int_0^1 \ell_p(r)\ell_p(r)'\,dr\right]J_n'[1+o(1)] \\ &= D_{nL}HE_n\left[\sum_{s=1}^{n}\ell_p\left(\frac{s}{n}\right)\ell_p\left(\frac{s}{n}\right)'\right]E_nH'D_{nL}[1+o(1)] \\ &= nD_{nL}H\left\{E_n\int_0^1 \ell_p(r)\ell_p(r)'\,drE_n\right\}H'D_{nL}[1+o(1)]. \end{aligned}$$

Because $\varepsilon(n) \to 0$ as $n \to \infty$, the matrix $E_n = e_1e_1' + o(1)$, where $e_1 = (1,0,\ldots,0)'$, and the final expression given previously is

$$\begin{aligned} &nD_{nL}H\left\{e_1e_1'\int_0^1 \ell_p(r)\ell_p(r)'\,dre_1e_1' + o(1)\right\}H'D_{nL}[1+o(1)] \\ &\quad = nD_{nL}He_1e_1'H'D_{nL}[1+o(1)] \\ &\quad = nD_{nL}i_{p+1}i_{p+1}'D_{nL}[1+o(1)], \end{aligned}$$

where i_{p+1} is the $p+1$ sum vector (i.e., it has unity in each component). This gives the first result (38).

Next, consider the inverse sample moment matrix

$$(L'L)^{-1} = J_n^{-1\prime}\left[\sum_{s=1}^{n} \ell_p\left(\frac{s}{n}\right)\ell_p\left(\frac{s}{n}\right)'\right]^{-1} J_n^{-1}[1+o(1)]$$

$$= \frac{1}{n} J_n^{-1\prime}\left[\int_0^1 \ell_p(r)\ell_p(r)'\,dr\right]^{-1} J_n^{-1}[1+o(1)]$$

$$= \frac{1}{n} D_{nL}^{-1} H^{-1\prime} E_n^{-1}\left[\int_0^1 \ell_p(r)\ell_p(r)'\,dr\right]^{-1} E_n^{-1} H^{-1} D_{nL}^{-1}[1+o(1)].$$

Now observe that E_n^{-1} is dominated by its final diagonal element, and so we can write $E_n^{-1} = (1/\varepsilon(n)^p)e_{p+1}e_{p+1}'[1+o(1)]$ where $e_{p+1} = (0,0,\ldots,1)'$. The final expression given previously is asymptotically equivalent to

$$\frac{1}{n\varepsilon(n)^{2p}} D_{nL}^{-1} H^{-1\prime} e_{p+1} e_{p+1}'\left[\int_0^1 \ell_p(r)\ell_p(r)'\,dr\right]^{-1} e_{p+1} e_{p+1}' H^{-1} D_{nL}^{-1}$$

$$= \frac{e_{p+1}'\left[\int_0^1 \ell_p(r)\ell_p(r)'\,dr\right]^{-1} e_{p+1}}{n\varepsilon(n)^{2p}} D_{nL}^{-1} H^{-1\prime} e_{p+1} e_{p+1}' H^{-1} D_{nL}^{-1}$$

$$= \frac{e_{p+1}'\left[\int_0^1 \ell_p(r)\ell_p(r)'\,dr\right]^{-1} e_{p+1}}{n\varepsilon(n)^{2p}} D_{nL}^{-1} \mu_{p+1} \mu_{p+1}' D_{nL}^{-1}, \tag{79}$$

giving the stated result. The second equality (79) holds because $H^{-1\prime}e_{p+1} = \mu_{p+1}$, the final column of $H^{-1'}$, as is apparent from the fact that $H'\mu_{p+1} = e_{p+1}$, which can be verified by direct multiplication using Lemma 7.7(ii). ■

Proof of Theorem 5.1. Solving (43) for β_1 and β_2, we get

$$\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} = \begin{bmatrix} L_1(n) & 0 \\ 0 & L_2(n)\end{bmatrix}^{-1}$$

$$\times \begin{bmatrix} \varepsilon_1(n) & \varepsilon_2(n) \\ \frac{1}{2}\varepsilon_1(n)[\varepsilon_1(n)+\eta_1(n)] & \frac{1}{2}\varepsilon_2(n)[\varepsilon_2(n)+\eta_2(n)]\end{bmatrix}^{-1} \begin{bmatrix}\alpha_{n1}\\ \alpha_{n2}\end{bmatrix}$$

$$= \begin{bmatrix} L_1(n)\varepsilon_1(n) & 0 \\ 0 & L_2(n)\varepsilon_2(n) \end{bmatrix}^{-1}$$

$$\times \begin{bmatrix} 1 & 1 \\ \frac{1}{2}[\varepsilon_1(n)+\eta_1(n)] & \frac{1}{2}[\varepsilon_2(n)+\eta_2(n)] \end{bmatrix}^{-1} \begin{bmatrix} \alpha_{n1} \\ \alpha_{n2} \end{bmatrix}$$

$$= \begin{bmatrix} L_1(n)\varepsilon_1(n) & 0 \\ 0 & L_2(n)\varepsilon_2(n) \end{bmatrix}^{-1} \frac{2}{\delta(n)} \begin{bmatrix} \frac{1}{2}[\varepsilon_2(n)+\eta_2(n)] & -1 \\ -\frac{1}{2}[\varepsilon_1(n)+\eta_1(n)] & 1 \end{bmatrix}$$

$$\times \begin{bmatrix} \alpha_{n1} \\ \alpha_{n2} \end{bmatrix},$$

and so

$$\frac{\sqrt{n}}{2} \begin{bmatrix} \delta(n)\varepsilon_1(n)L_1(n)[\hat{\beta}_1-\beta_1] \\ \delta(n)\varepsilon_2(n)L_2(n)[\hat{\beta}_2-\beta_2] \end{bmatrix} = \begin{bmatrix} \frac{1}{2}[\varepsilon_2(n)+\eta_2(n)] & -1 \\ -\frac{1}{2}[\varepsilon_1(n)+\eta_1(n)] & 1 \end{bmatrix}$$

$$\times \begin{bmatrix} \sqrt{n}[\hat{\alpha}_{n1}-\alpha_{n1}] \\ \sqrt{n}[\hat{\alpha}_{n2}-\alpha_{n2}] \end{bmatrix}.$$

Because $\varepsilon_j(n)+\eta_j(n)=o(1)$ for $j=1,2$, we have

$$\frac{\sqrt{n}}{2} \begin{bmatrix} \delta(n)\varepsilon_1(n)L_1(n)[\hat{\beta}_1-\beta_1] \\ \delta(n)\varepsilon_2(n)L_2(n)[\hat{\beta}_2-\beta_2] \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} [\sqrt{n}[\hat{\alpha}_{n2}-\alpha_{n2}]]$$

$$+ o_p(1) \to_d N\left(0, \frac{\sigma^2}{(2!)^2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right), \tag{80}$$

where the coefficient $1/(2!)^2$ comes from the third diagonal element of the inverse matrix $[\int_0^1 \ell_2(r)\ell_2(r)'\,dr]^{-1}$. Finally, the constant term satisfies

$$\beta_0 = \alpha_0 - L_1(n)\beta_1 - L_2(n)\beta_2,$$

which, in combination with (80), leads to

$$
\begin{aligned}
&\frac{\sqrt{n}}{2}\,\delta(n)\varepsilon_{\min}(n)(\hat{\beta}_0-\beta_0)\\
&\quad=\begin{cases} -L_2(n)\,\dfrac{\sqrt{n}}{2}\,\delta(n)\varepsilon_2(n)(\hat{\beta}_2-\beta_2)+o_p(1) & \text{if } \varepsilon_2(n)=o(\varepsilon_1(n))\\[2ex] -L_1(n)\,\dfrac{\sqrt{n}}{2}\,\delta(n)\varepsilon_1(n)(\hat{\beta}_1-\beta_1)+o_p(1) & \text{if } \varepsilon_1(n)=o(\varepsilon_2(n)) \end{cases}\\
&\quad=\begin{cases} -\sqrt{n}[\hat{\alpha}_{n2}-\alpha_{n2}]+o_p(1) & \text{if } \varepsilon_2(n)=o(\varepsilon_1(n))\\ \sqrt{n}[\hat{\alpha}_{n2}-\alpha_{n2}]+o_p(1) & \text{if } \varepsilon_1(n)=o(\varepsilon_2(n)) \end{cases}\\
&\quad=1_\varepsilon\sqrt{n}[\hat{\alpha}_{n2}-\alpha_{n2}]+o_p(1)\to_d N\left(0,\frac{\sigma^2}{(2!)^2}\right),
\end{aligned}
$$

where

$$
\varepsilon_{\min}(n)=\begin{cases} \varepsilon_2(n) & \text{if } \varepsilon_2(n)=o(\varepsilon_1(n))\\ \varepsilon_1(n) & \text{if } \varepsilon_1(n)=o(\varepsilon_2(n)) \end{cases}
$$

and

$$
1_\varepsilon=\begin{cases} -1 & \text{if } \varepsilon_2(n)=o(\varepsilon_1(n))\\ 1 & \text{if } \varepsilon_1(n)=o(\varepsilon_2(n)) \end{cases}.
$$

We deduce that

$$
\begin{aligned}
\frac{\sqrt{n}\delta(n)}{2}\begin{bmatrix} \varepsilon_{\min}(n)(\hat{\beta}_0-\beta_0)\\ \varepsilon_1(n)L_1(n)(\hat{\beta}_1-\beta_1)\\ \varepsilon_2(n)L_2(n)(\hat{\beta}_2-\beta_2) \end{bmatrix}&=\begin{bmatrix} \mp 1\\ -1\\ 1 \end{bmatrix}\sqrt{n}[\hat{\alpha}_{n2}-\alpha_{n2}]\\
&\to_d N\left(0,\frac{\sigma^2}{(2!)^2}\begin{bmatrix} 1 & \pm 1 & \mp 1\\ \pm 1 & 1 & -1\\ \mp 1 & -1 & 1 \end{bmatrix}\right),
\end{aligned}
$$

which gives the stated result upon scaling. ∎

Proof of Lemma 6.1. Setting $S_t=\sum_{s=1}^t u_s$, using partial summation, and proceeding as in the proof of Lemma 2.1(i) we have

$$
\begin{aligned}
&\frac{1}{\sqrt{n}L(n)}\sum_{t=1}^n f\left(\frac{t}{n};\theta\right)L(t)u_t\\
&\quad=f(1;\theta)\,\frac{S_n}{\sqrt{n}}-\frac{1}{L(n)}\sum_{t=1}^n\left[f\left(\frac{t}{n};\theta\right)L\left(n\,\frac{t}{n}\right)-f\left(\frac{t-1}{n};\theta\right)L\left(n\,\frac{t-1}{n}\right)\right]\\
&\qquad\times\frac{S_{t-1}}{\sqrt{n}}. \qquad \textbf{(81)}
\end{aligned}
$$

Assume that the probability space is constructed so that we can embed the standardized partial sum $S_{t-1}/\sqrt{n}$ in Brownian motion as in equation (8) following Assumption LP, namely,

$$\sup_{1\le t\le n}\left|\frac{S_{t-1}}{\sqrt{n}} - B\left(\frac{t-1}{n}\right)\right| = o_{a.s.}\left(\frac{1}{n^{(1/2)-(1/p)}}\right).$$

Then, the first term of (81) clearly satisfies

$$f(1;\theta)\frac{S_n}{\sqrt{n}} \to_d f(1;\theta_0)B(1), \tag{82}$$

as $n\to\infty$ uniformly over $\theta\in N_n^0$. The second term of (81) is

$$\begin{aligned}
&\frac{1}{L(n)}\sum_{t=1}^{n}\left[f\left(\frac{t}{n};\theta\right)L\left(n\frac{t}{n}\right) - f\left(\frac{t-1}{n};\theta\right)L\left(n\frac{t-1}{n}\right)\right]\frac{S_{t-1}}{\sqrt{n}}\\
&\quad = \frac{1}{L(n)}\sum_{t=1}^{n} f\left(\frac{t}{n};\theta\right)\left[L\left(n\frac{t}{n}\right) - L\left(n\frac{t-1}{n}\right)\right]\frac{S_{t-1}}{\sqrt{n}}\\
&\qquad + \frac{1}{L(n)}\sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right)\left[f\left(\frac{t}{n};\theta\right) - f\left(\frac{t-1}{n};\theta\right)\right]\frac{S_{t-1}}{\sqrt{n}}\\
&\quad = T_1 + T_2.
\end{aligned} \tag{83}$$

Start with T_1. We have $f(t/n;\theta) - f(t/n;\theta_0) = f_\theta(t/n;\theta^*)(\theta-\theta_0)$ for some $\theta^*\in N_n^0$, and so

$$\begin{aligned}
&\sup_{\theta\in N_n^0}\left|\frac{1}{L(n)}\sum_{t=1}^{n}\left[f\left(\frac{t}{n};\theta\right) - f\left(\frac{t}{n};\theta_0\right)\right]\left[L\left(n\frac{t}{n}\right) - L\left(n\frac{t-1}{n}\right)\right]\frac{S_{t-1}}{\sqrt{n}}\right|\\
&\quad = \sup_{\theta\in N_n^0}\left|\sum_{t=1}^{n}\left[f_\theta\left(\frac{t}{n};\theta^*\right)\right]\frac{\left[L\left(n\frac{t}{n}\right) - L\left(n\frac{t-1}{n}\right)\right]}{L(n)}\frac{S_{t-1}}{\sqrt{n}}\right||\theta-\theta_0|\\
&\quad \le \sup_{\theta\in N_n^0}|\theta-\theta_0|\frac{1}{n}\sum_{t=1}^{n}\sup_{\theta\in N_n^0}\left|f_\theta\left(\frac{t}{n};\theta^*\right)\right|\left|\frac{L'(t^*)}{L(n)}\right|\left|\frac{S_{t-1}}{\sqrt{n}}\right|\\
&\quad = \sup_{\theta\in N_n^0}|\theta-\theta_0|\frac{1}{n}\sum_{t=1}^{n} g_1\left(\frac{t}{n}\right)\left|\frac{L'(t^*)}{L(n)}\right|\left|\frac{S_{t-1}}{\sqrt{n}}\right|\\
&\quad = \sup_{\theta\in N_n^0}|\theta-\theta_0|\left[\int_0^1 g_1(r)\left|\frac{B(r)}{r}\right|\left|\frac{L'(nr)nr}{L(nr)}\right|\left|\frac{L(nr)}{L(n)}\right|dr + o_p(1)\right]\\
&\quad = o_p(1),
\end{aligned} \tag{84}$$

as

$$\left|\frac{L(nr)}{L(n)}\right| \to 1, \qquad \left|\frac{L'(nr)nr}{L(nr)}\right| \to 0 \quad \text{for all } r > 0 \quad \text{and}$$

$$\sup_{\theta_1,\theta_2 \in N_n^0} |\theta_1 - \theta_2| \to 0$$

as $n \to \infty$. It follows that

$$T_1 = \frac{1}{L(n)} \sum_{t=1}^{n} f\left(\frac{t}{n};\theta_0\right)\left[L\left(n\frac{t}{n}\right) - L\left(n\frac{t-1}{n}\right)\right]\frac{S_{t-1}}{\sqrt{n}} + o_p(1)$$

uniformly over $\theta \in N_n^0$. But, just as in the proof of Lemma 2.1(i),

$$\frac{1}{L(n)} \sum_{t=1}^{n} f\left(\frac{t}{n};\theta_0\right)\left[L\left(n\frac{t}{n}\right) - L\left(n\frac{t-1}{n}\right)\right]\frac{S_{t-1}}{\sqrt{n}}$$

$$= \frac{1}{L(n)} \int_0^1 f(r;\theta_0)B(r)\,dL(nr) = O_p(\varepsilon(n)) = o_p(1),$$

and so $T_1 = o_p(1)$ uniformly over $\theta \in N_n^0$.

Next, consider T_2. We have $f(t/n;\theta) - f((t-1)/n;\theta) = f_r(t^*/n;\theta)$ for some $t^* \in (t-1,t)$, and then

$$T_2 = \frac{1}{L(n)} \sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right)\left[f\left(\frac{t}{n};\theta\right) - f\left(\frac{t-1}{n};\theta\right)\right]\frac{S_{t-1}}{\sqrt{n}}$$

$$= \frac{1}{L(n)n} \sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) f_r\left(\frac{t^*}{n};\theta\right)\frac{S_{t-1}}{\sqrt{n}}. \tag{85}$$

Now $f_r(t^*/n;\theta) - f_r(t^*/n;\theta_0) = f_{r\theta}(t^*/n;\theta^*)(\theta - \theta_0)$ for some $\theta^* \in N_n^0$ and

$$\frac{1}{L(n)n} \sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) f_r\left(\frac{t^*}{n};\theta\right)\frac{S_{t-1}}{\sqrt{n}} - \frac{1}{L(n)n} \sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) f_r\left(\frac{t^*}{n};\theta_0\right)\frac{S_{t-1}}{\sqrt{n}}$$

$$= \frac{1}{L(n)n} \sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right)\left[f_r\left(\frac{t^*}{n};\theta\right) - f_r\left(\frac{t^*}{n};\theta_0\right)\right]\frac{S_{t-1}}{\sqrt{n}}$$

$$= \frac{1}{L(n)n} \sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) f_{r\theta}\left(\frac{t^*}{n};\theta^*\right)\frac{S_{t-1}}{\sqrt{n}}(\theta - \theta_0),$$

and, just as in (84), we find that

$$\sup_{\theta\in N_n^0}\left|\frac{1}{L(n)n}\sum_{t=1}^{n}L\left(n\frac{t-1}{n}\right)f_{r\theta}\left(\frac{t^*}{n};\theta^*\right)\frac{S_{t-1}}{\sqrt{n}}(\theta-\theta_0)\right|$$
$$\leq \sup_{\theta\in N_n^0}\theta-\theta_0\left|\left[\int_0^1 g_3(r)\left|\frac{B(r)}{r}\right|\left|\frac{L'(nr)nr}{L(nr)}\right|\left|\frac{L(nr)}{L(n)}\right|dr+o_p(1)\right]\right.$$
$$=o_p(1).$$

Moreover,

$$\frac{1}{L(n)n}\sum_{t=1}^{n}L\left(n\frac{t-1}{n}\right)f_r\left(\frac{t^*}{n};\theta_0\right)\frac{S_{t-1}}{\sqrt{n}}$$
$$=\int_0^1\frac{L(nr)}{L(n)}f(r;\theta_0)B(r)\,dr+o_p(1)\to_p\int_0^1 f(r;\theta_0)B(r)\,dr, \tag{86}$$

as $n\to\infty$. It follows from (85) and (86) that

$$T_2=\frac{1}{L(n)n}\sum_{t=1}^{n}L\left(n\frac{t-1}{n}\right)f_r\left(\frac{t^*}{n};\theta_0\right)\frac{S_{t-1}}{\sqrt{n}}+o_p(1)\to_d\int_0^1 f_r(r;\theta_0)B(r)\,dr, \tag{87}$$

uniformly over $\theta\in N_n^0$. We deduce from (81)–(83) and (87) that

$$\frac{1}{\sqrt{n}L(n)}\sum_{t=1}^{n}f\left(\frac{t}{n};\theta\right)L(t)u_t=f(1;\theta_0)\frac{S_n}{\sqrt{n}}-(T_1+T_2)\to_d f(1;\theta_0)B(1)$$
$$-\int_0^1 f_r(r;\theta_0)B(r)\,dr$$
$$=\int_0^1 f(r;\theta_0)\,dB(r), \tag{88}$$

uniformly over $\theta\in N_n^0$, giving the stated result. ∎

Proof of Lemma 6.2. *Part (i).* For any slowly varying function L satisfying Assumption SSV and any function $f\in C^1$, we can show in the same way as Lemma 6.1 that

$$\frac{1}{\sqrt{n}L(n)}\sum_{s=1}^{n}f\left(\frac{s}{n};\theta_0\right)L(s)u_s\to_d\int_0^1 f(r;\theta_0)\,dB(r)=N\left(0,\sigma^2\int_0^1 f(r)^2\,dr\right), \tag{89}$$

extending (9). The limit (54) follows directly.

Part (ii). Using Lemma 6.1 and (89), the asymptotic form of $D_n^{-1}H_n(\theta_0)D_n^{-1}$ is

$$\begin{bmatrix} \frac{1}{n}\sum_{s=1}^{n}\left(\frac{s}{n}\right)^{2\gamma_0} & \frac{1}{n\log n}\sum_{s=1}^{n}\left[\beta_0\left(\frac{s}{n}\right)^{2\gamma_0}\log s - u_s\left(\frac{s}{n}\right)^{\gamma_0}\log s\right] \\ \frac{1}{n\log n}\sum_{s=1}^{n}\left[\beta_0\left(\frac{s}{n}\right)^{2\gamma_0}\log s - u_s\left(\frac{s}{n}\right)^{\gamma_0}\log s\right] & \frac{1}{n\log^2 n}\sum_{s=1}^{n}\left[(\beta_0)^2\left(\frac{s}{n}\right)^{2\gamma_0}\log^2 s - u_s\beta_0\left(\frac{s}{n}\right)^{\gamma_0}\log^2 s\right] \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{n}\sum_{s=1}^{n}\left(\frac{s}{n}\right)^{2\gamma_0} & \frac{1}{n\log n}\sum_{s=1}^{n}\beta_0\left(\frac{s}{n}\right)^{2\gamma_0}\log s \\ \frac{1}{n\log n}\sum_{s=1}^{n}\beta_0\left(\frac{s}{n}\right)^{2\gamma_0}\log s & \frac{1}{n\log^2 n}\sum_{s=1}^{n}(\beta_0)^2\left(\frac{s}{n}\right)^{2\gamma_0}\log^2 s \end{bmatrix} + O_p\left(\frac{1}{\sqrt{n}}\right)$$

$$= \begin{bmatrix} \int_0^1 r^{2\gamma_0}\,dr & \beta_0\int_0^1 r^{2\gamma_0}\,dr + \frac{\beta_0}{\log n}\int_0^1 r^{2\gamma_0}\log r\,dr \\ \beta_0\int_0^1 r^{2\gamma_0}\,dr + \frac{\beta_0}{\log n}\int_0^1 r^{2\gamma_0}\log r\,dr & \frac{\beta_0^2}{\log^2 n}\int_0^1 r^{2\gamma_0}(\log n + \log r)^2\,dr \end{bmatrix}$$

$$+ O_p\left(\frac{1}{\sqrt{n}}\right)$$

$$\to_p \begin{bmatrix} \int_0^1 r^{2\gamma_0}\,dr & \beta_0\int_0^1 r^{2\gamma_0}\,dr \\ \beta_0\int_0^1 r^{2\gamma_0}\,dr & \beta_0^2\int_0^1 r^{2\gamma_0}\,dr \end{bmatrix} = \frac{1}{2\gamma_0+1}\begin{bmatrix} 1 & \beta_0 \\ \beta_0 & \beta_0^2 \end{bmatrix}, \tag{90}$$

as stated.

Part (iii). Upon calculation of (90) and rescaling, we deduce that

$$F_n^{-1}H_n(\theta_0)F_n^{-1} = \log^2 n D_n^{-1}H_n(\theta_0)D_n^{-1}$$

$$= \frac{\log^2 n}{2\gamma_0+1}\begin{bmatrix} 1 & \beta_0 - \frac{\beta_0}{(2\gamma_0+1)\log n} \\ \beta_0 - \frac{\beta_0}{(2\gamma_0+1)\log n} & \beta_0^2 + \frac{2\beta_0^2}{(2\gamma_0+1)\log n} + \frac{2\beta_0^2}{(2\gamma_0+1)^2\log^2 n} \end{bmatrix}$$

$$+ o_p(1)$$

$$= \frac{\log^2 n}{2\gamma_0+1}\left\{\begin{bmatrix} 1 & \beta_0 \\ \beta_0 & \beta_0^2 \end{bmatrix} + \begin{bmatrix} 0 & -\frac{\beta_0}{(2\gamma_0+1)\log n} \\ -\frac{\beta_0}{(2\gamma_0+1)\log n} & \frac{2\beta_0^2}{(2\gamma_0+1)\log n} \end{bmatrix}\right\} + O_p(1),$$

whose eigenvalues are evidently $O(\log^2 n)$ and $O(\log n)$, respectively, provided $\beta_0 \neq 0$.

Part (iv). First calculate

$$
\begin{aligned}
d_n(\theta_0) &= \det[D_n^{-1} H_n(\theta_0) D_n^{-1}] \\
&= \left(\frac{\beta_0^2}{\log^2 n}\int_0^1 r^{2\gamma_0}(\log n + \log r)^2\,dr\right)\int_0^1 r^{2\gamma_0}\,dr \\
&\quad - \left(\beta_0\int_0^1 r^{2\gamma_0}\,dr + \frac{\beta_0}{\log n}\int_0^1 r^{2\gamma_0}\log r dr\right)^2 + O_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \frac{\beta_0^2}{\log^2 n}\left[\int_0^1 r^{2\gamma_0}\log^2 r dr\left(\int_0^1 r^{2\gamma_0}\,dr\right) - \left(\int_0^1 r^{2\gamma_0}\log r dr\right)^2\right] \\
&\quad + o\left(\frac{1}{\log^2 n}\right) \\
&= \frac{\beta_0^2}{\log^2 n}\left[-\frac{2}{(2\gamma^0+1)}\int_0^1 r^{2\gamma_0}\log r dr\int_0^1 r^{2\gamma_0}\,dr - \left(\int_0^1 r^{2\gamma_0}\log r dr\right)^2\right] \\
&\quad + o\left(\frac{1}{\log^2 n}\right) \\
&= \frac{\beta_0^2}{\log^2 n}\left[\frac{2}{(2\gamma^0+1)^3}\int_0^1 r^{2\gamma_0}\,dr - \left(\frac{1}{2\gamma^0+1}\int_0^1 r^{2\gamma_0}\,dr\right)^2\right] \\
&\quad + o\left(\frac{1}{\log^2 n}\right) \\
&= \frac{\beta_0^2}{(2\gamma^0+1)^4\log^2 n}.
\end{aligned}
$$

Then

$$
\begin{aligned}
&D_n H_n(\theta_0)^{-1} D_n \\
&= \frac{1}{d_n(\theta_0)}\begin{bmatrix}
\beta_0^2\int_0^1 r^{2\gamma_0}\,dr + \frac{2\beta_0^2}{\log n}\int_0^1 r^{2\gamma_0}\log r dr + \frac{\beta_0^2}{\log^2 n}\int_0^1 r^{2\gamma_0}\log^2 r dr & -\beta_0\int_0^1 r^{2\gamma_0}\,dr - \frac{\beta_0}{\log n}\int_0^1 r^{2\gamma_0}\log r dr \\
-\beta_0\int_0^1 r^{2\gamma_0}\,dr - \frac{\beta_0}{\log n}\int_0^1 r^{2\gamma_0}\log r dr & \int_0^1 r^{2\gamma_0}\,dr
\end{bmatrix} \\
&\quad + O_p\left(\frac{\log^2 n}{\sqrt{n}}\right)
\end{aligned}
$$

$$= \frac{1}{d_n(\theta_0)} \begin{bmatrix} \frac{\beta_0^2}{\log^2 n} \frac{2}{(2\gamma^0+1)^3} - \frac{2\beta_0^2}{\log n} \frac{1}{(2\gamma^0+1)^2} + \frac{\beta_0^2}{2\gamma^0+1} & -\frac{\beta_0}{2\gamma^0+1} + \frac{\beta_0}{\log n} \frac{1}{(2\gamma^0+1)^2} \\ -\frac{\beta_0}{2\gamma^0+1} + \frac{\beta_0}{\log n} \frac{1}{(2\gamma^0+1)^2} & \frac{1}{2\gamma^0+1} \end{bmatrix}$$

$$+ O_p\left(\frac{\log^2 n}{\sqrt{n}}\right)$$

$$= \frac{(2\gamma^0+1)^3 \log^2 n}{\beta_0^2} \begin{bmatrix} \beta_0^2 - \frac{2\beta_0^2}{\log n} \frac{1}{(2\gamma^0+1)} + \frac{\beta_0^2}{\log^2 n} \frac{2}{(2\gamma^0+1)^2} & -\beta_0 + \frac{\beta_0}{\log n} \frac{1}{(2\gamma^0+1)} \\ -\beta_0 + \frac{\beta_0}{\log n} \frac{1}{(2\gamma^0+1)} & 1 \end{bmatrix}$$

$$+ O_p\left(\frac{\log^2 n}{\sqrt{n}}\right).$$

We deduce that

$$F_n H_n(\theta_0)^{-1} F_n$$

$$= \frac{(2\gamma^0+1)^3}{\beta_0^2} \begin{bmatrix} \beta_0^2 - \frac{2\beta_0^2}{\log n} \frac{1}{(2\gamma^0+1)} + \frac{\beta_0^2}{\log^2 n} \frac{2}{(2\gamma^0+1)^2} & -\beta_0 + \frac{\beta_0}{\log n} \frac{1}{(2\gamma^0+1)} \\ -\beta_0 + \frac{\beta_0}{\log n} \frac{1}{(2\gamma^0+1)} & 1 \end{bmatrix}$$

$$+ O_p\left(\frac{1}{\sqrt{n}}\right)$$

$$\to_p \frac{(2\gamma^0+1)^3}{\beta_0^2} \begin{bmatrix} \beta_0^2 & -\beta_0 \\ -\beta_0 & 1 \end{bmatrix}, \tag{91}$$

as given.

Part (v). Define $C_n = D_n/n^{\delta}$ for some small positive $\delta \in (0, \gamma_0 + \frac{1}{2})$, so that $C_n D_n^{-1} = O(n^{-\delta}) = o(1)$ and $\lambda_{\min}(C_n) \to \infty$ as $n \to \infty$, where $\lambda_{\min}$ denotes the smallest eigenvalue. Construct the following shrinking neighborhood of θ_0:

$$N_n^0 = \{\theta \in \Theta : \|C_n(\theta - \theta_0)\| \leq 1\}$$

and define the matrix

$$C_n^{-1}[H_n(\theta) - H_n(\theta_0)]C_n^{-1} = \begin{bmatrix} a_{11,n} & a_{12,n} \\ a_{21,n} & a_{22,n} \end{bmatrix}.$$

We show that

$$\sup_{\theta \in N_n^0} \|C_n^{-1}[H_n(\theta) - H_n(\theta_0)]C_n^{-1}\| = o_p(1). \tag{92}$$

Note that in N_n^0 we have

$$\sup_{\theta\in N_n^0}|\gamma-\gamma_0|\le\frac{1}{n^{\gamma_0+(1/2)-\delta}\log n},\quad \sup_{\theta\in N_n^0}|\beta-\beta_0|\le\frac{1}{n^{\gamma_0+(1/2)-\delta}}.$$

Also, because $\gamma_0>-\frac{1}{2}$ we can choose $\varepsilon>0$ such that $\gamma_0>-\frac{1}{2}+\varepsilon$, and then we have the dominating function

$$\sup_{\theta\in N_n^0}|r^{\gamma}|\le r^{-(1/2)+\varepsilon}. \tag{93}$$

Consider the individual elements of $C_n^{-1}[H_n(\theta)-H_n(\theta_0)]C_n^{-1}$ in turn. First, for γ^* between γ and γ_0, we have

$$\begin{aligned}a_{11,n}&=\frac{1}{n^{2\gamma_0+1+2\delta}}\sum_{s=1}^{n}(s^{2\gamma}-s^{2\gamma_0})=\frac{2}{n^{2\gamma_0+1+\delta}}\sum_{s=1}^{n}s^{2\gamma^*}\log s(\gamma-\gamma_0)\\&=\frac{2\log n}{n^{2(\gamma_0-\gamma^*)+2\delta}}\left[\frac{1}{n\log n}\sum_{s=1}^{n}\left(\frac{s}{n}\right)^{2\gamma^*}\log s\right](\gamma-\gamma_0).\end{aligned}$$

Next

$$\begin{aligned}\frac{1}{n\log n}\sum_{s=1}^{n}\left(\frac{s}{n}\right)^{2\gamma^*}\log s&=\frac{1}{n}\sum_{s=1}^{n}\left(\frac{s}{n}\right)^{2\gamma^*}+\frac{1}{n\log n}\sum_{s=1}^{n}\left(\frac{s}{n}\right)^{2\gamma^*}\log\frac{s}{n}\\&\to\int_0^1 r^{2\gamma_0}\,dr,\end{aligned}$$

uniformly for $\theta\in N_n^0$ in view of the majorization (93). It follows that for large enough n

$$\begin{aligned}\sup_{\theta\in N_n^0}\left|\frac{1}{n^{2\gamma_0+1+2\delta}}\sum_{s=1}^{n}(s^{2\gamma}-s^{2\gamma_0})\right|&=O\left(\frac{2\log n}{n^{2(\gamma_0-\gamma^*)+2\delta}}\sup_{\theta\in N_n^0}|\gamma-\gamma_0|\right)\\&=O\left(\frac{1}{n^{\delta}}\frac{1}{n^{\gamma_0+(1/2)-\delta}\log n}\right)=o(1).\end{aligned} \tag{94}$$

Next,

$$\begin{aligned}a_{12,n}&=\frac{1}{n^{2\gamma_0+1+2\delta}\log n}\sum_{s=1}^{n}[\beta s^{2\gamma}\log s-u_s s^{\gamma}\log s-(\beta_0 s^{\gamma_0}-\beta s^{\gamma})s^{\gamma}\log s]\\&=\frac{\beta}{n^{2\gamma_0+1+2\delta}\log n}\sum_{s=1}^{n}s^{2\gamma}\log s-\frac{1}{n^{2\gamma_0+1+2\delta}\log n}\sum_{s=1}^{n}u_s s^{\gamma}\log s\\&\quad-\frac{1}{n^{2\gamma_0+1+2\delta}\log n}\sum_{s=1}^{n}(\beta_0 s^{\gamma_0}-\beta s^{\gamma})s^{\gamma}\log s\\&=T_1+T_2+T_3.\end{aligned}$$

We find

$$\sup_{\theta \in N_n^0} (|T_1| + |T_3|) = o(1)$$

in the same way as (94). For term T_2, in view of (89) we have for each $\gamma \in N_n^0$

$$\frac{1}{n^{2\gamma_0+1+2\delta} \log n} \sum_{s=1}^{n} u_s s^{\gamma} \log s = \frac{n^{\gamma}}{n^{2\gamma_0+(1/2)+2\delta}} \frac{1}{\sqrt{n} n^{\gamma} \log n} \sum_{s=1}^{n} u_s \left(\frac{s}{n}\right)^{\gamma} \log s.$$

By Lemma 7.1 we have

$$\frac{1}{\sqrt{n} \log n} \sum_{s=1}^{n} u_s \left(\frac{s}{n}\right)^{\gamma} \log s \to_d \int_0^1 r^{\gamma_0} \, dB(r),$$

uniformly over $\theta \in N_n^0$, and

$$\frac{n^{\gamma}}{n^{2\gamma_0+(1/2)+2\delta}} = O_p\left(\frac{1}{n^{2\delta}}\right),$$

uniformly over $\theta \in N_n^0$ for large enough n. Hence,

$$\sup_{\theta \in N_n^0} |T_2| = o_p(1),$$

as $n \to \infty$. The argument for the term $a_{22,n}$ is entirely analogous, and (92) therefore follows. ■

Proof of Theorem 6.3. Standard asymptotic arguments of nonlinear regression for nonstationary dependent time series (e.g., Wooldridge, 1994, Thm. 8.1) may be applied. But, modifications to the arguments need to be made to attend to the singularity arising from the asymptotically collinear elements s^{γ_0} and $s^{\gamma_0} \log s$ that appear in the score $S_n(\theta_0)$. First, the demonstration that there is a consistent root of the first-order conditions (51) in an open, shrinking neighborhood of θ_0 follows as in the proof of Wooldridge's Theorem 8.1 using (52) and Lemma 6.2(v). There are two changes in the proof that are needed: (i) the standardizing matrix is F_n^{-1} in place of D_n^{-1}, as discussed in Remark (c) following Lemma 6.2; (ii) the scaled Hessian matrix $F_n^{-1} H_n(\theta_0) F_n^{-1}$ does not tend to a positive definite limit with finite eigenvalues bounded away from the origin. Instead, as shown in the proof of Lemma 6.2(iii), $F_n^{-1} H_n(\theta_0) F_n^{-1}$ is positive definite for all large n and has eigenvalues of order $O(\log^2 n)$ and $O(\log n)$, and the smallest eigenvalue $\lambda_{\min}(F_n^{-1} H_n(\theta_0) F_n^{-1}) = O(\log n) \to \infty$. With these changes, the remainder of the consistency argument in Wooldridge's Theorem 8.1 holds, and we obtain $F_n(\hat{\theta} - \theta_0) = O_p(1)$.

Next, scaling the first-order conditions (52), we have

$$0 = F_n^{-1} S_n(\theta_0) + F_n^{-1} H_n(\theta_0) F_n^{-1} F_n(\hat{\theta} - \theta_0) + F_n^{-1} [H_n^* - H_n(\theta_0)] F_n^{-1} F_n(\hat{\theta} - \theta_0),$$

and then

$$F_n(\hat{\theta} - \theta_0) = -[I + (F_n H_n(\theta_0)^{-1} F_n) F_n^{-1} [H_n^* - H_n(\theta_0)] F_n^{-1}]^{-1} \times (F_n H_n(\theta_0)^{-1} F_n) F_n^{-1} S_n(\theta_0). \quad \textbf{(95)}$$

Note that $F_n = (1/\log n) D_n = diag[n^{\gamma_0+1/2}/\log n, n^{\gamma_0+1/2}]$ and because $D_n^{-1} S_n(\theta_0) = O_p(1)$ from Lemma 6.2(i), we have $F_n^{-1} S_n(\theta_0) = O_p(\log n)$. However, from (91) in the proof of Lemma 6.2(iv) we have

$$F_n H_n(\theta_0)^{-1} F_n = \frac{(2\gamma^0+1)^3}{\beta_0^2} \begin{bmatrix} \beta_0^2 - \frac{2\beta_0^2}{\log n} \frac{1}{(2\gamma^0+1)} + \frac{\beta_0^2}{\log^2 n} \frac{2}{(2\gamma^0+1)^2} & -\beta_0 + \frac{\beta_0}{\log n} \frac{1}{(2\gamma^0+1)} \\ -\beta_0 + \frac{\beta_0}{\log n} \frac{1}{(2\gamma^0+1)} & 1 \end{bmatrix} + O_p\left(\frac{1}{\sqrt{n}}\right) = O_p(1), \quad \textbf{(96)}$$

and from Lemma 6.2(v) we have

$$\sup_{\theta \in N_n^0} F_n^{-1} [H_n(\theta) - H_n(\theta_0)] F_n^{-1} \leq \sup_{\theta \in N_n^0} C_n^{-1} [H_n(\theta) - H_n(\theta_0)] C_n^{-1} = o_p(1), \quad \textbf{(97)}$$

where $N_n^0 = \{\theta \in \Theta : \|C_n(\theta - \theta_0)\| \leq 1\}$, a shrinking neighborhood around θ_0 with $C_n = D_n/n^{\delta}$ for some small $\delta > 0$. Because $F_n(\hat{\theta} - \theta_0) = O_p(1)$, it follows that $\hat{\theta}, \theta^* \in N_n^0$ with probability approaching unity as $n \to \infty$, where θ^* is a generic mean value between $\hat{\theta}$ and θ_0. Hence

$$F_n^{-1} [H_n^* - H_n(\theta_0)] F_n^{-1} = o_p(1), \quad \textbf{(98)}$$

and so combining (96) and (98) we have

$$(F_n H_n(\theta_0)^{-1} F_n) F_n^{-1} [H_n^* - H_n(\theta_0)] F_n^{-1} = o_p(1). \quad \textbf{(99)}$$

Then, from (91), (95), and (99) we deduce that

$$
\begin{aligned}
&F_n(\hat\theta-\theta_0)\\
&\quad= -(F_nH_n(\theta_0)^{-1}F_n)F_n^{-1}S_n(\theta_0)+o_p(1)\\
&\quad= -\frac{(2\gamma^0+1)^3}{\beta_0^2}\left\{\begin{bmatrix} \beta_0^2-\frac{2\beta_0^2}{\log n}\frac{1}{(2\gamma^0+1)}+\frac{\beta_0^2}{\log^2 n}\frac{2}{(2\gamma^0+1)^2} & -\beta_0+\frac{\beta_0}{\log n}\frac{1}{(2\gamma^0+1)}\\ -\beta_0+\frac{\beta_0}{\log n}\frac{1}{(2\gamma^0+1)} & 1\end{bmatrix}+O_p\left(\frac{1}{\sqrt{n}}\right)\right\}\\
&\qquad\times\frac{1}{\sqrt{n}}\sum_{s=1}^{n}\begin{bmatrix}\left(\frac{s}{n}\right)^{\gamma_0}\log n\\ \beta_0\left(\frac{s}{n}\right)^{\gamma_0}\log s\end{bmatrix}u_s+o_p(1)\\
&\quad= -\frac{(2\gamma^0+1)^3}{\beta_0^2}\frac{1}{\sqrt{n}}\sum_{s=1}^{n}\begin{bmatrix}-\frac{2\beta_0^2}{\log n}\frac{\left(\frac{s}{n}\right)^{\gamma_0}\log n}{(2\gamma^0+1)}+\frac{\beta_0^2}{\log^2 n}\frac{2\left(\frac{s}{n}\right)^{\gamma_0}\log n}{(2\gamma^0+1)^2}-\beta_0^2\left(\frac{s}{n}\right)^{\gamma_0}\log\frac{s}{n}+\frac{\beta_0^2}{\log n}\frac{\left(\frac{s}{n}\right)^{\gamma_0}\log s}{(2\gamma^0+1)}\\ \beta_0\left(\frac{s}{n}\right)^{\gamma_0}\log\frac{s}{n}+\frac{\beta_0}{\log n}\frac{\left(\frac{s}{n}\right)^{\gamma_0}\log n}{(2\gamma^0+1)}\end{bmatrix}u_s\\
&\qquad+o_p(1)\\
&\quad= -\frac{(2\gamma^0+1)^3}{\beta_0^2}\frac{1}{\sqrt{n}}\sum_{s=1}^{n}\begin{bmatrix}-\beta_0^2\left(\frac{s}{n}\right)^{\gamma_0}\log\frac{s}{n}-\frac{\beta_0^2\left(\frac{s}{n}\right)^{\gamma_0}}{(2\gamma^0+1)}+\frac{\beta_0^2}{\log n}\frac{\left(\frac{s}{n}\right)^{\gamma_0}\log\frac{s}{n}}{(2\gamma^0+1)}\\ \beta_0\left(\frac{s}{n}\right)^{\gamma_0}\log\frac{s}{n}+\frac{\beta_0\left(\frac{s}{n}\right)^{\gamma_0}}{(2\gamma^0+1)}\end{bmatrix}u_s+o_p(1)\\
&\quad= -\frac{(2\gamma^0+1)^3}{\beta_0^2}\frac{1}{\sqrt{n}}\sum_{s=1}^{n}\begin{bmatrix}-\beta_0^2\left(\frac{s}{n}\right)^{\gamma_0}\log\frac{s}{n}-\frac{\beta_0^2}{(2\gamma^0+1)}\left(\frac{s}{n}\right)^{\gamma_0}\\ \beta_0\left(\frac{s}{n}\right)^{\gamma_0}\log\frac{s}{n}+\frac{\beta_0\left(\frac{s}{n}\right)^{\gamma_0}}{(2\gamma^0+1)}\end{bmatrix}u_s+O_p\left(\frac{1}{\log n}\right)+o_p(1)\\
&\quad= (2\gamma^0+1)^3\begin{bmatrix}1\\ -1/\beta_0\end{bmatrix}\frac{1}{\sqrt{n}}\sum_{s=1}^{n}\left[\left(\frac{s}{n}\right)^{\gamma_0}\log\frac{s}{n}+\frac{\left(\frac{s}{n}\right)^{\gamma_0}}{(2\gamma^0+1)}\right]u_s+O_p\left(\frac{1}{\log n}\right)+o_p(1)\\
&\quad\to_d (2\gamma^0+1)^3\begin{bmatrix}1\\ -1/\beta_0\end{bmatrix}\int_0^1 r^{\gamma_0}\left[\log r+\frac{1}{2\gamma_0+1}\right]dB(r)=\begin{bmatrix}1\\ -1/\beta_0\end{bmatrix}N(0,\sigma^2(2\gamma^0+1)^3),
\end{aligned}
$$

giving the stated result. ■

REFERENCES

Akonom, J. (1993) Comportement asymptotique du temps d'occupation du processus des sommes partielles. *Annals of the Institute of Henri Poincaré* 29, 57–81.

Barro, R.J. & X. Sala-i-Martin (2004) *Economic Growth*, 2nd ed. Prentice-Hall.

Bingham, N.H., C.M. Goldie, & J.L. Teugels (1987) *Regular Variation*. Cambridge University Press.

Csörgö, M. & L. Horváth (1993) *Weighted Approximations in Probability and Statistics.* Wiley.

de Haan, L. and S. Resnick (1996) Second order regular variation and rates of convergence in extreme value theory. *Annals of Probability* 24, 97–124.

Erdélyi, A. (1953) *Higher Transcendental Functions*, vol. 1. Krieger.

Grenander, U. & M. Rosenblatt (1957) *Statistical Analysis of Stationary Time Series.* Wiley.

Hannan, E.J. (1970) *Multiple Time Series*. Wiley.

Hurvich, C.M., R. Deo, & J. Brodsky (1998) The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis* 19, 19–46.

Knopp, K. (1990) *Theory and Application of Infinite Series*. Dover.

Phillips, P.C.B. (1999) Discrete Fourier Transforms of Fractional Processes. Cowles Foundation Discussion Paper no. 1243, Yale University.

Phillips, P.C.B. (2006) Unit root log periodogram regression. *Journal of Econometrics* (forthcoming).

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.

Phillips, P.C.B. & Y. Sun (2003) Regression with an evaporating logarithmic trend. *Econometric Theory* 19, 692–701.

Robinson, P.M. (1995) Log-periodogram regression of time series with long range dependence. *Annals of Statistics* 23, 1048–1072.

Shorack, G.R. & J.A. Wellner (1986) *Empirical Processes with Applications to Statistics.* Wiley.

Wang, Q., Y.-X. Lin, & C.M. Gulati (2003) Strong approximation for long memory processes with applications. *Journal of Theoretical Probability* 16, 377–389.

Wooldridge, J.M. (1994) Estimation and inference for dependent processes. In R.F. Engle & D.L. McFadden (eds.), *Handbook of Econometrics*, vol. 4. North-Holland.

Wu, C.-F. (1981) Asymptotic theory of nonlinear least squares estimation. *Annals of Statistics* 9, 501–513.