Regression with Slowly Varying Regressors

By

Peter C.B. Phillips

July 2001

COWLES FOUNDATION DISCUSSION PAPER No. 1310

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

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

Regression with Slowly Varying Regressors*

Peter C. B. Phillips
Cowles Foundation, Yale University University of Auckland & University of York

First draft: 22 June 2000
This version: 20 May 2001

Abstract

Slowly varying regressors are asymptotically collinear in linear regression. Usual regression formulae for asymptotic standard errors 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 slowly varying functions and central limit theorems with slowly varying weights are given that assist in the development of a regression theory. Multivariate regression and polynomial regression with slowly varying 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 trend regression are discussed.

JEL Classification: C22
AMS 1991 subject classifications: 26A12, 62J05, 62J02

Key words and phrases: Asymptotic expansion, Collinearity, Karamata representation, Slow variation, Smooth variation, Trend regression.

1 Introduction

Empirical models of 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 author thanks Sidney Resnick for references on second order regular variation and the NSF for research support under Grant No. SBR 97-30295. The paper was typed by the author in SW2.5.

the regressors, such as those which are often characterized (e.g., by 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

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

where u_s is an unobserved error process. 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, and Hurvich, Deo and Brodsky, 1998, and the references therein). In that case (discussed in Example 3.2(a) below), y_s is the periodogram of the data measured at the Fourier frequencies $\lambda_s = \frac{2\pi s}{n}$, $s = 1,.., m \leq n$, 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 $D_n = diag\left(\sqrt{n}, \sqrt{n}\log n\right)$, and $F_n^{-1} = diag\left(\frac{\sqrt{n}}{\log n}, \sqrt{n}\right)$, we have (c.f. equations (23) and (24) below)

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

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 function $L(s)$. In effect, the intercept and any slowly varying 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

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

or similar regressions involving polynomials in a slowly varying function. In such cases, one finds that the sample moment matrix of the regressors, while 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 slowly varying 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

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

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 (c.f., equation (52) below). It follows that statistical models like (2) manifest asymptotic collinearity analogous to that of the linear regression (1). Wu (1980, 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 treatment of regressions of this type. The discussion is conducted in terms of slowly varying 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 organised 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 slowly varying 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). Section 7 and 8 contain supplementary technical material and proofs. Notation is listed in Section 9.

2 Assumptions and Preliminary Results

It will be convenient to use some standard theory of slowly varying 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, theorem 1.3.1, p.12), any slowly varying (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 which is arbitrarily smooth (e.g., BGT, theorem 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 below 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 assume the following.

2.1 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 $\eta(x) \to 0$ as $x \to \infty$.

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

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

and, more generally,

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

(BGT, p.44). The class for $L(x)$ covered in SSV includes all of the common slowly varying 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.

Since 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 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 slowly varying function K, $K(n)/n^\eta \to 0$ for arbitrary $\eta > 0$, we have, given some $\alpha > 0$ and any positive integer k

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

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 the regression errors u_s satisfy the following linear process condition.

2.2 Assumption (LP) *For all $t > 0$, u_t has Wold representation*

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

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

It is well known (e.g., Phillips and Solo, 1992, theorem 3.4) that LP is sufficient for the partial sums $S_t = \sum_{s=1}^{t} u_s$ to satisfy the functional law $\frac{1}{\sqrt{n}} S_{[n\cdot]} \to_d B(\cdot)$, 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

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

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 Csörgõ, M. and L. Horváth (1993). A strong approximation justifying (8) in the case where u_t is a linear processes is given in Phillips (1999) for time series data under LP. Akonom (1993) gave (8) with an $o_p(n^{-\frac{1}{2}+\frac{1}{p}})$ error under LP using the weaker moment requirement that $\mu_p = E\,|u_t|^p < \infty$ for some $p > 2$.

Under 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 slowly varying. In particular, we have the following.

2.3 Lemma *If $L(t)$ satisfies SSV, $\overline{L} = n^{-1}\sum_{t=1}^{n} L(t)$ and u_t satisfies LP, then:*

(a) $\frac{1}{\sqrt{n}L(n)} \sum_{t=1}^{n} L(t)\,u_t \to_d B(1) =_d N\left(0, \sigma^2\right)$ *as* $n \to \infty$.

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

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

2.4 Heuristics As shown in (56) and Lemma 7.3 below, one of the implications of SSV is that we have the following asymptotic representation of $L(t)$ for $t = nr$ with $r > 0$

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

Such a function may be called second order slowly varying (c.f., de Haan and Resnick, 1995, who discuss second order regular variation). For the sample mean $\overline{L}$, we have

$$\overline{L} = L(n) - L(n)\,\varepsilon(n) + o\left(L(n)\,\varepsilon(n)\right).$$

In consequence, the standardized sums that appear in (a), (b) and (c) of Lemma 2.3 have the approximate asymptotic forms

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

to which we may apply a standard central limit argument like that of (9) above. These cases indicate that, as far as first order asymptotic theory is concerned, weighted means of u_t with arbitrary slowly varying weights behave in a common way, at least up to a normalization factor that depends on the asymptotic form of the slowly varying function and its corresponding ε-function. The 'common' form that appears in these expressions is that of a logarithmic trend function $\log t$, while the influence of the particular slowly varying 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

$$y_s = \alpha + \beta L(s) + u_s \quad s = 1, ..., n$$

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

$$\overline{L} = \frac{1}{n}\sum_{s=1}^{n}L(s), \quad \frac{1}{n}\sum_{s=1}^{n}\left(L(s)-\overline{L}\right)^2 = \frac{1}{n}\sum_{s=1}^{n}L(s)^2 - \left(\frac{1}{n}\sum_{s=1}^{n}L(s)\right)^2. \tag{11}$$

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

$$\sum_{t=1}^{n}L(t)^k = \int_1^n L(t)^k\,dt + O(n^\eta), \tag{12}$$

where $\eta > 0$ is arbitrarily small, and Lemma 7.3 gives the explicit asymptotic expansion

$$\begin{aligned}
\frac{1}{n}\sum_{t=1}^{n}L(t)^k &= L(n)^k - kL(n)^k\varepsilon(n) + k^2L(n)^k\varepsilon(n)^2 - kL(n)^k\varepsilon(n)\eta(n) \\
&\quad + o\left(L(n)^k\varepsilon(n)\left[\eta(n)+\varepsilon(n)\right]\right).
\end{aligned} \tag{13}$$

Using (12) and (13) in (11) leads to the following asymptotic expansions for these sample moments

$$\begin{aligned}\overline{L} &= L(n) - L(n)\varepsilon(n) + o(L(n)\varepsilon(n)), \\ \frac{1}{n}\sum_{t=1}^{n}\left(L(t)-\overline{L}\right)^2 &= L(n)^2\varepsilon(n)^2 + o\left(L(n)^2\varepsilon(n)\left[\eta(n)+\varepsilon(n)\right]\right). \end{aligned} \tag{14}$$

Then,

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

and

$$\begin{aligned}\sqrt{n}\varepsilon(n)(\widehat{\alpha}-\alpha) &= \varepsilon(n)\frac{1}{\sqrt{n}}\sum_{t=1}^{n}u_t - \sqrt{n}L(n)\varepsilon(n)\left(\widehat{\beta}-\beta\right)\left[1+(\varepsilon(n))\right] \\ &= -\sqrt{n}L(n)\varepsilon(n)\left(\widehat{\beta}-\beta\right) + o_p(1). \end{aligned} \tag{15}$$

The limit theory for the regression coefficients then follows directly from (15) and Lemma 2.3.

3.1 Theorem *If $L(t)$ satisfies SSV and u_t satisfies LP, then*

$$\begin{bmatrix}\sqrt{n}\varepsilon(n)(\widehat{\alpha}-\alpha) \\ \sqrt{n}L(n)\varepsilon(n)\left(\widehat{\beta}-\beta\right)\end{bmatrix} \to_d N\left(0,\sigma^2\begin{bmatrix}1 & -1 \\ -1 & 1\end{bmatrix}\right). \tag{16}$$

3.2 Examples

(a) $L(s) = \log s$ This gives the semi-logarithmic model. Here, $\varepsilon(n) = \frac{1}{\log n}$, $L(n)\varepsilon(n) = 1$, and (16) is

$$\begin{bmatrix}\frac{\sqrt{n}}{\log n}(\widehat{\alpha}-\alpha) \\ \sqrt{n}\left(\widehat{\beta}-\beta\right)\end{bmatrix} \to_d N\left(0,\sigma^2\begin{bmatrix}1 & -1 \\ -1 & 1\end{bmatrix}\right). \tag{17}$$

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

$$\log(I_X(\lambda_s)) = \widehat{c} - 2\widehat{d}\log\lambda_s + \text{residual}, \quad s = 1, ..., m \tag{18}$$

where $I_X(\lambda_s)$ is the periodogram of a time series $(X_t)_{t=1}^n$ and $\lambda_s = \frac{2\pi s}{n}$ 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 (18) is taken over a band of frequencies that shrink to the origin, so that $\frac{1}{m} + \frac{m}{n} \to 0$. Then (18) has the alternate form

$$\log(I_X(\lambda_s)) = \left(\widehat{c} - 2\widehat{d}\log\frac{2\pi}{n}\right) - 2\widehat{d}\log s + \text{residual} = \widehat{c}_n - 2\widehat{d}\log s + \text{residual} \tag{19}$$

where $\widehat{c}_n = \widehat{c} - 2\widehat{d}\log\frac{2\pi}{n}$. Set $c = \log C$, $c_n = c - 2d\log\frac{2\pi.}{n}$. The moment matrix of the regressors in (19) is asymptotically singular, just as in (1). Although the details of the central limit theory differ from Lemma 2.3 because of the properties of the residual terms in (18) (c.f. Robinson, 1995, and Hurwich et al., 1998), we nevertheless end up with a result analogous to (17) but with sample size m, viz.

$$\left[\begin{array}{c} \frac{\sqrt{m}}{\log m}\left(\widehat{c}_n - c_n\right) \\ 2\sqrt{m}\left(\widehat{d} - d\right) \end{array}\right] \rightarrow_d N\left(0, \frac{\pi^2}{6}\left[\begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array}\right]\right).$$

Since $\widehat{c}_n - c_n = (\widehat{c} - c) - 2\left(\widehat{d} - d\right)\log\frac{2\pi.}{n}$, we have

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

from which we deduce that

$$\left[\begin{array}{c} \frac{\sqrt{m}}{\log n}\left(\widehat{c} - c\right) \\ 2\sqrt{m}\left(\widehat{d} - d\right) \end{array}\right] \rightarrow_d N\left(0, \frac{\pi^2}{6}\left[\begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array}\right]\right),$$

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

f

(b) $L(s) = \frac{1}{\log s}$ This example arises when the regressor decays slowly. Here $\varepsilon(n) = -\frac{1}{\log n}$, $L(n)\varepsilon(n) = -\frac{1}{\log^2 n}$ and (16) is

$$\left[\begin{array}{c} \frac{\sqrt{n}}{\log n}\left(\widehat{\alpha} - \alpha\right) \\ \frac{\sqrt{n}}{\log^2 n}\left(\widehat{\beta} - \beta\right) \end{array}\right] \rightarrow_d N\left(0, \sigma^2\left[\begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array}\right]\right). \tag{20}$$

(c) $L(s) = \log\log s$ Here, $\varepsilon(n) = \frac{1}{\log\log n}\frac{1}{\log n}$, $L(n)\varepsilon(n) = \frac{1}{\log n}$, and (16) is

$$\left[\begin{array}{c} \frac{\sqrt{n}}{\log\log n\log n}\left(\widehat{\alpha} - \alpha\right) \\ \frac{\sqrt{n}}{\log n}\left(\widehat{\beta} - \beta\right) \end{array}\right] \rightarrow_d N\left(0, \sigma^2\left[\begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array}\right]\right). \tag{21}$$

(d) $L(s) = \frac{1}{\log\log s}$ Here, $\varepsilon(n) = -\frac{1}{\log\log n}\frac{1}{\log n}$, $L(n)\varepsilon(n) = -\frac{1}{\log^2\log n}\frac{1}{\log n}$, and (16) is

$$\left[\begin{array}{c} \frac{\sqrt{n}}{\log\log n\log n}\left(\widehat{\alpha} - \alpha\right) \\ \frac{\sqrt{n}}{\log^2\log n\log n}\left(\widehat{\beta} - \beta\right) \end{array}\right] \rightarrow_d N\left(0, \sigma^2\left[\begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array}\right]\right)$$

(e) $L(s) = \log^{\gamma} s$, $\gamma > 0$. In this case, $\varepsilon(n) = \frac{\gamma}{\log n}$, $L(n)\varepsilon(n) = \gamma \log^{\gamma-1} n$, and (16) is

$$\left[\begin{array}{c} \frac{\gamma\sqrt{n}}{\log n}(\widehat{\alpha}-\alpha) \\ \gamma\sqrt{n}\log^{\gamma-1} n\left(\widehat{\beta}-\beta\right) \end{array}\right] \rightarrow_d N\left(0, \sigma^2 \left[\begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array}\right]\right)$$

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. When $L(n) \to \infty$ as $n \to \infty$, the convergence rate of $\widehat{\beta}$ exceeds that of $\widehat{\alpha}$, because the signal from the regressor is stronger than that of a constant regressor; when $L(n) \to 0$ as $n \to \infty$, the convergence rate of $\widehat{\beta}$ is less than that of $\widehat{\alpha}$, because the signal from the regressor is weaker than that of a constant regressor.

3.3 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 (12) and (13), we have

$$\sum_{s=1}^{n} \left[\begin{array}{cc} 1 & L(s) \\ L(s) & L(s)^2 \end{array}\right] = n \left[\begin{array}{cc} 1 & L_{12}(n) \\ L_{12}(n) & L_{22}(n) \end{array}\right] + O(n^{\eta}), \tag{22}$$

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

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

Upon standardization with the diagonal matrix $D_n = diag\left(\sqrt{n}, \sqrt{n}L(n)\right)$, (22) becomes

$$\begin{aligned} D_n^{-1}\sum_{s=1}^{n} \left[\begin{array}{cc} 1 & L(s) \\ L(s) & L(s)^2 \end{array}\right] D_n^{-1} &= \left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array}\right] + \left[\begin{array}{cc} 0 & -\varepsilon(n) \\ -\varepsilon(n) & -2\varepsilon(n) \end{array}\right] + o(\varepsilon(n)) \\ &\rightarrow \left[\begin{array}{cc} 1 & 1 \\ 1 & 1 \end{array}\right]. \end{aligned} \tag{23}$$

Similarly, upon inversion, we have

$$\begin{aligned} &\left(\sum_{s=1}^{n} \left[\begin{array}{cc} 1 & L(s) \\ L(s) & L(s)^2 \end{array}\right]\right)^{-1} \\ = &\frac{1}{n\sum_{s=1}^{n}\left(L(s)-\overline{L}\right)^2} \sum_{s=1}^{n} \left[\begin{array}{cc} L(s)^2 & -L(s) \\ -L(s) & 1 \end{array}\right] \end{aligned}$$

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

$$= \begin{bmatrix} \frac{1}{n\varepsilon(n)^2} & -\frac{1}{nL(n)\varepsilon(n)^2} \\ -\frac{1}{nL(n)\varepsilon(n)^2} & \frac{1}{nL(n)^2\varepsilon(n)^2} \end{bmatrix} [1 + o(nL(n)\varepsilon(n)[\eta(n) + \varepsilon(n)])], \tag{24}$$

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 formulae 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 smoothly slowly varying function $L(s)$ and the data are generated by

$$y_s = \sum_{j=0}^{p} \beta_j L(s)^j + u_s = \beta' L_s + u_s, \tag{25}$$

where the regression error u_s satisfies LP. This model may be analyzed using the approach of the previous section. But, as the degree p increases in (25), the analysis becomes complicated because higher order expansions than (13) of the sample moments of $L(s)$ are needed in order to develop a complete asymptotic theory. An alternate approach is to rewrite the model (25) 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

$$\begin{aligned} y_s &= \alpha + \beta \log n + \beta \log \frac{s}{n} + u_s \\ &= \alpha_n + \beta \log \frac{s}{n} + u_s, \quad \text{say.} \end{aligned} \tag{26}$$

The regressors $\{1, \log \frac{s}{n}\}$ in (26) are not collinear. Writing $k(r) = [1, \log r]'$ and using standard manipulations, we obtain

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

Since, $\widehat{\alpha_n} - \alpha_n = \widehat{\alpha} - \alpha + \left(\widehat{\beta} - \beta\right) \log n$, we deduce that

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

which leads directly to the earlier result (17).

Extending this process to the model (25) gives the representation

$$\begin{aligned} y_s &= \sum_{j=0}^{p} \beta_j \left\{ L(n) \left[\frac{L(s)}{L(n)} - 1 \right] + L(n) \right\}^j + u_s \\ &= \sum_{j=0}^{p} \beta_j L(n)^j \sum_{i=0}^{j} \binom{j}{i} \left[\frac{L(s)}{L(n)} - 1 \right]^i + u_s \\ &= \sum_{j=0}^{p} \alpha_{nj} \left[\frac{L(s)}{L(n)} - 1 \right]^j + u_s \end{aligned}$$

where

$$\alpha_{n0} = \sum_{j=0}^{p} \beta_j L(n)^j \tag{27}$$

$$\alpha_{nk} = \sum_{j=k}^{p} \beta_j L(n)^j \binom{j}{k} \quad k = 1, ..., p-1 \tag{28}$$

$$\alpha_{np} = \beta_p L(n)^p . \tag{29}$$

Define

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

and the model (25) becomes

$$y_s = \sum_{j=0}^{p} \alpha_{nj} K_{nj}\left(\frac{s}{n}\right) + u_s := \alpha_n' K_n\left(\frac{s}{n}\right) + u_s. \tag{30}$$

Least squares estimation gives

$$\widehat{\alpha}_n - \alpha_n = \left[\sum_{t=1}^{n} K_n\left(\frac{t}{n}\right) K_n\left(\frac{t}{n}\right)' \right]^{-1} \left[\sum_{t=1}^{n} K_n\left(\frac{t}{n}\right) u_t \right]. \tag{31}$$

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

4.1 Theorem

(a) *If $L(t)$ satisfies SSV, then*

$$\frac{1}{n} D_{n\varepsilon}^{-1} \sum_{t=1}^{n} K_n\left(\frac{t}{n}\right) \to \int_0^1 \ell_p(r)\, dr = \left[\begin{array}{cccccc} 1, & -1, & 2!, & -3!, & \ldots, & (-1)^p p! \end{array} \right]',$$

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

$$\begin{aligned} & \frac{1}{n} D_{n\varepsilon}^{-1} \sum_{t=1}^{n} K_n\left(\frac{t}{n}\right) K_n\left(\frac{t}{n}\right)' D_{n\varepsilon}^{-1} \\ \to & \int_0^1 \ell_p(r)\, \ell_p(r)'\, dr \\ = & \begin{bmatrix} 1 & -1 & 2! & -3! & \cdots & (-1)^p p! \\ -1 & 2! & -3! & 4! & \cdots & (-1)^{p+1}(p+1)! \\ 2! & -3! & 4! & -5! & \cdots & (-1)^{p+2}(p+2)! \\ -3! & 4! & -5! & 6! & \cdots & (-1)^{p+3}(p+3)! \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ (-1)^p p! & (-1)^{p+1}(p+1)! & (-1)^{p+2}(p+2)! & (-1)^{p+3}(p+3)! & \cdots & (2p)! \end{bmatrix}, \end{aligned} \tag{32}$$

which is positive definite.

(b) *If* $L(t)$ *satisfies SSV and* u_t *satisfies LP then*

$$\sqrt{n} D_{n\varepsilon}\left[\widehat{\alpha}_n - \alpha_n\right] \to_d N\left(0, \sigma^2\left[\int_0^1 \ell_p(r)\, \ell_p(r)'\, dr\right]^{-1}\right). \tag{33}$$

Next, we rewrite this limit distribution in terms of the original coefficients using relations (27) - (29). It transpires that only the final component, $\widehat{\alpha}_{np}$, in $\widehat{\alpha}_n$ (which translates to the component $\widehat{\beta}_p$ in the original coordinates) determines the nondegenerate part of the limit theory for the full set of coefficients.

4.2 Theorem *If* $L(t)$ *satisfies SSV and* u_t *satisfies LP then*

$$\sqrt{n}\varepsilon(n)^p D_{nL}\left(\widehat{\beta} - \beta\right) = \mu_{p+1}\sqrt{n} L(n)^p \varepsilon(n)^p \left(\widehat{\beta}_p - \beta_p\right) + o_p(1) \to_d N\left(0, v^{p+1,p+1}\mu_{p+1}\mu_{p+1}'\right),$$

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

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

$$\frac{v^{p+1,p+1}}{n\varepsilon(n)^{2p}} D_{nL}^{-1}\mu_{p+1}\mu_{p+1}' D_{nL}^{-1}, \tag{34}$$

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, (34) is the asymptotic form of $(L'L)^{-1}$.

4.3 Theorem *If $L(t)$ satisfies SSV, then:*

(a)

$$L'L = nD_{nL}i_{p+1}i'_{p+1}D_{nL}\left[1+o(1)\right], \tag{35}$$

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

(b)

$$\left(L'L\right)^{-1} = \frac{e'_{p+1}\left[\int_0^1 \ell_p(r)\,\ell_p(r)'\,dr\right]^{-1} e_{p+1}}{n\varepsilon(n)^{2p}} D_{nL}^{-1}\mu_{p+1}\mu'_{p+1}D_{nL}^{-1}\left[1+o(1)\right], \tag{36}$$

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

It follows from (36) 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 $s^2\left(L'L\right)^{-1}$ whenever s^2 is a consistent estimate of σ^2.

5 Regression with Multiple SSV Regressors

Multiple regression with different slowly varying functions as regressors is also of some interest in applications. One such formulation is given in example 5.2 below and involves a slowly varying 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 slowly varying 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

$$y_s = \beta_0 + \beta_1 L_1(s) + \beta_2 L_2(s) + u_s = \beta' L_s + u_s, \text{ say} \tag{37}$$

where the regression error u_s satisfies LP. An asymptotic theory of regression in this model is obtained by showing that (37) has an alternate, asymptotically equivalent, form involving a quadratic function of the simpler regressor $\log\left(\frac{s}{n}\right)$. Analysis similar to the previous section then applies.

Rewrite (37) as follows

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

To transform the regressors in this version of the model, we note from Lemma 7.10 that L_j has a higher order representation in terms of its ε- and η- functions that has the asymptotic form

$$\frac{L_j(rn)}{L_j(n)} - 1 = \varepsilon_j(n)\log r + \frac{1}{2}\varepsilon_j(n)\left[\varepsilon_j(n) + \eta_j(n)\right]\log^2 r\left[1 + o(1)\right], \quad r > 0. \tag{38}$$

Equation (38) shows L_j to be third order slowly varying in the sense that

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

thereby extending the concept of second order slow variation that appears in the earlier expression (10). Using the expansion (38) we write

$$\begin{aligned} y_s &= \beta_0 + \beta_1 L_1(n) + \beta_2 L_2(n) \\ &\quad + \beta_1 L_1(n)\varepsilon_1(n)\log\frac{s}{n} + \frac{1}{2}\beta_1\varepsilon_1(n)\left[\varepsilon_1(n) + \eta_1(n)\right]\log^2\frac{s}{n}\left[1 + o(1)\right] \\ &\quad + \beta_2 L_2(n)\varepsilon_2(n)\log\frac{s}{n} + \frac{1}{2}\beta_2\varepsilon_2(n)\left[\varepsilon_2(n) + \eta_2(n)\right]\log^2\frac{s}{n}\left[1 + o(1)\right] + u_s \\ &= \alpha_{n0} + \alpha_{n1}\log\left(\frac{s}{n}\right) + \alpha_{n2}\log^2\left(\frac{s}{n}\right)\left[1 + o(1)\right] + u_s, \text{ say} \end{aligned} \tag{39}$$

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

$$\begin{aligned} \begin{bmatrix} \alpha_{n0} \\ \alpha_{n1} \\ \alpha_{n2} \end{bmatrix} &= \begin{bmatrix} 1 & L_1(n) & L_2(n) \\ 0 & L_1(n)\varepsilon_1(n) & L_2(n)\varepsilon_2(n) \\ 0 & \frac{1}{2}L_1(n)\varepsilon_1(n)\left[\varepsilon_1(n) + \eta_1(n)\right] & \frac{1}{2}L_2(n)\varepsilon_2(n)\left[\varepsilon_2(n) + \eta_2(n)\right] \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 & 1 \\ 0 & \varepsilon_1(n) & \varepsilon_2(n) \\ 0 & \frac{1}{2}\varepsilon_1(n)\left[\varepsilon_1(n) + \eta_1(n)\right] & \frac{1}{2}\varepsilon_2(n)\left[\varepsilon_2(n) + \eta_2(n)\right] \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & L_1(n) & 0 \\ 0 & 0 & L_2(n) \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}. \end{aligned} \tag{40}$$

For further asymptotic analysis, we impose the condition

$$\delta(n) = \left[\varepsilon_2(n) + \eta_2(n)\right] - \left[\varepsilon_1(n) + \eta_1(n)\right] \neq 0 \tag{41}$$

which is necessary if we are to solve (40) for the original coefficients in (37). If (41) does not hold, then the regressors L_1 and L_2 are collinear to the second order in (38). 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 (41), uses only the second order form (38) and gives the limit theory for the original coefficients in (37).

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

$$\sqrt{n}\delta(n)\begin{bmatrix} \varepsilon_{\min}(n)\left(\widehat{\beta}_0 - \beta_0\right) \\ \varepsilon_1(n) L_1(n)\left[\widehat{\beta}_1 - \beta_1\right] \\ \varepsilon_2(n) L_2(n)\left[\widehat{\beta}_2 - \beta_2\right] \end{bmatrix} \sim \begin{bmatrix} 1_\varepsilon \\ -1 \\ 1 \end{bmatrix} \sqrt{n}\left[\widehat{\alpha}_{n2} - \alpha_{n2}\right] \to_d N\left(0, \frac{\sigma^2}{2!}\begin{bmatrix} 1 & -1_\varepsilon & 1_\varepsilon \\ -1_\varepsilon & 1 & -1 \\ 1_\varepsilon & -1 & 1 \end{bmatrix}\right), \tag{42}$$

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

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

5.2 Discussion

(a) Equation (39) indicates that multiple regression with different slowly varying 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 slowly varying 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.

(b) The scaling factor $\varepsilon_{\min}(n)$ in (42) 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.

(c) If $L_i(x) = \log x$ for some i then there is no second order term in (38) and $\varepsilon_i(n) + \eta_i(n) = 0$ in that case. The first matrix in (40) is simpler in this case and can be made upper triangular by permuting coefficients if necessary.

(d) Just as in the polynomial regression case, the limit distribution (42) is singular and has rank unity.

5.3 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 $\varepsilon_1(n) = -\frac{1}{\log n}$, $\eta_1(n) = -\frac{1}{\log n}$, $\varepsilon_2(n) = \frac{1}{\log\log n}\frac{1}{\log n}$ and $\eta_2(n) = -\frac{1}{\log\log n} - \frac{1}{\log n}$. Then

$$\begin{aligned}
\varepsilon_1(n)+\eta_1(n) &= -\frac{2}{\log n}, \\
\varepsilon_2(n)+\eta_2(n) &= -\frac{1}{\log\log n}+o\left(\frac{1}{\log\log n}\right), \\
\delta(n) &= -\frac{1}{\log\log n}+o\left(\frac{1}{\log\log n}\right), \\
\varepsilon_{\min} &= \varepsilon_2(n)=\frac{1}{\log\log n}\frac{1}{\log n}.
\end{aligned}$$

We deduce that

$$\frac{\sqrt{n}}{\log\log n}\left[\begin{array}{c} \frac{1}{\log\log n}\frac{1}{\log n}\left(\widehat{\beta}_0-\beta_0\right) \\ \frac{(-1)}{\log^2 n}\left[\widehat{\beta}_1-\beta_1\right] \\ \frac{1}{\log n}\left[\widehat{\beta}_2-\beta_2\right] \end{array}\right] \sim \left[\begin{array}{c} -1 \\ -1 \\ 1 \end{array}\right]\frac{\sqrt{n}}{\log n\log\log n}\left[\widehat{\beta}_2-\beta_2\right]\rightarrow_d N\left(0,\frac{\sigma^2}{2!}\left[\begin{array}{ccc} 1 & 1 & -1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{array}\right]\right).$$

The coefficient of the growth term converges fastest, but at less than a $\sqrt{n}$ 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.

5.4 The General Case Consider the model

$$y_s=\sum_{j=0}^{p}\beta_j L_j(s)+u_s=\beta' L_s+u_s,\ \text{say}, \tag{43}$$

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

$$y_s=\sum_{j=0}^{p}\beta_j L_j(n)+\sum_{j=1}^{p}\beta_j L_j(n)\left[\frac{L_j\left(n\frac{s}{n}\right)}{L_j(n)}-1\right]+u_s. \tag{44}$$

Assume that each L_j has a higher order representation extending (38) in terms of the following asymptotic expansion

$$\frac{L_j(rn)}{L_j(n)}-1=\sum_{i=1}^{p-1}\varepsilon_{ji}(n)\log^i r+\varepsilon_{jp}(n)\log^p r\left[1+o(1)\right], \quad r>0, \tag{45}$$

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

$$\varepsilon_{ji}(n)=o\left(\varepsilon_{ji-1}(n)\right), \tag{46}$$

for each j and each $\iota>1$, so the coefficients, $\varepsilon_{ji}(n)$, in (45) decrease in order of magnitude as i increases. Such a higher order expansion can be developed under conditions analogous to SSV in which each function in the sequence L, ε, η, ... itself has a Karamata representation with an ε- function that is SSV. Applying (45) in (44), we obtain the transformed model

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

The coefficients in this system satisfy

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

where

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

Define

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

and note that, in view of (46), we have

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

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,

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

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

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

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

so that

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

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

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

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

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

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

5.6 Theorem *If $L(t)$ satisfies SSV, u_t satisfies LP and $\det\Xi_n \neq 0$, then*

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

$$\sim i_{p+1}\left[\sqrt{n}\left(\widehat{\alpha}_{np}-\alpha_{np}\right)\right] \to_d N\left(0, \frac{\sigma^2}{(p!)^2} i_{p+1} i_{p+1}'\right) \tag{47}$$

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 (47) the scale coefficients $\frac{L_j(n)\varepsilon_j(n)}{\eta^{jp}(n)}$ as well as $\varepsilon_{\min}$ are implicitly signed. That is, the elements $\frac{\varepsilon_j(n)}{\eta^{jj}(n)}$ may have positive or negative signs. In consequence, since 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 Trends

In the nonlinear trend model (2), let u_s satisfy LP, let $\theta_0 = (\beta_0, \gamma_0)$ be the true values of the parameters and asssume 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 (1980, Example 4, p.507 & p.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 $\widehat{\theta} = (\widehat{\beta}, \widehat{\gamma})$ 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 (while 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 $(\widehat{\beta}, \widehat{\gamma})$ solve the extremum problem

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

and satisfy the first order conditions

$$S_n\left(\widehat{\beta}, \widehat{\gamma}\right) = 0, \tag{48}$$

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

$$0 = S_n(\theta_0) + H_n(\theta_0)\left(\widehat{\theta} - \theta_0\right) + [H_n^* - H_n(\theta_0)]\left(\widehat{\theta} - \theta_0\right), \tag{49}$$

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

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

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

6.1 Lemma *Let L be a slowly varying function satisfying SSV, and suppose u_s satisfies 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 neighbourhood 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*

$$\sup_{\theta \in N_n^0} |f_\theta(r;\theta)| \le g_\theta(r;\theta_0), \quad \sup_{\theta \in N_n^0} |f_r(r;\theta)| \le g_r(r;\theta_0), \quad \sup_{\theta \in N_n^0} |f_{r\theta}(r;\theta)| \le g_{r\theta}(r;\theta_0)$$

by functions $g_\theta(r;\theta_0)$, $g_r(r;\theta_0)$ and $g_{r\theta}(r;\theta_0)$ all of which are absolutely integrable over $[0,1]$. Then

$$\frac{1}{\sqrt{n}L(n)} \sum_{s=1}^{n} f\left(\frac{s}{n};\theta\right) L(s)\, u_s \to_d \int_0^1 f(r;\theta_0)\, dB(r) = N\left(0, \sigma^2 \int_0^1 f(r;\theta_0)^2\, dr\right), \tag{50}$$

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

6.2 Lemma *Suppose u_s satisfies 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*

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

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

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

Then

(a)

$$D_n^{-1} S_n(\theta_0) \to_d -\int_0^1 \begin{bmatrix} r^{\gamma_0} \\ \beta_0 r^{\gamma_0} \end{bmatrix} dB(r) = N\left(0, \frac{\sigma^2}{2\gamma_0+1} \begin{bmatrix} 1 & \beta_0 \\ \beta_0 & \beta_0^2 \end{bmatrix}\right), \tag{51}$$

(b)

$$D_n^{-1} H_n(\theta_0) D_n^{-1} \to_p \frac{1}{2\gamma_0+1} \begin{bmatrix} 1 & \beta_0 \\ \beta_0 & \beta_0^2 \end{bmatrix}, \tag{52}$$

(c)

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

(d)

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

(e)

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

6.3 Remarks

(a) Part (a) reveals that the order of convergence of the first member of (49), the score $S_n(\theta_0)$, is determined by the scaling factor D_n^{-1}. However, from part (b), 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 (49) to find a limit theory for a standardized form of $(\widehat{\theta} - \theta_0)$. Part (d) shows that upon standardization by F_n, rather than D_n, the inverse hessian matrix converges but also has a singular limit.

(b) Part (e) is useful in showing that, after rescaling, the third term of (49) can be neglected in the asymptotic behavior of $\widehat{\theta} - \theta_0$.

(c) As the following result shows, the appropriate scaling factor for (49) is the matrix F_n^{-1}, not D_n^{-1}, even though $D_n^{-1} S_n(\theta_0)$ is $O_p\left(1\right)$.

6.4 Theorem *In the model (2), let u_s satisfy 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 $\widehat{\theta} = (\widehat{\beta}, \widehat{\gamma})$ is consistent and has the following limit distribution as $n \to \infty$*

$$\begin{aligned} F_n\left(\widehat{\theta} - \theta_0\right) &\to_d \left(2\gamma^0+1\right)^3 \left[\begin{array}{c} 1 \\ -1/\beta_0 \end{array}\right] \int_0^1 r^{\gamma_0}\left[\log r + \frac{1}{2\gamma_0+1}\right] dB\left(r\right) \\ &= \left[\begin{array}{c} 1 \\ -1/\beta_0 \end{array}\right] N\left(0, \sigma^2\left(2\gamma^0+1\right)^3\right). \end{aligned}$$

6.5 Remarks

(a) The estimator $\widehat{\theta}$ 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 $\widehat{\theta}$. With the new scaling we have

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

and then

$$F_n\left(\widehat{\theta}-\theta_0\right) = -\left[I + \left(F_nH_n(\theta_0)^{-1}F_n\right)F_n^{-1}\left[H_n^* - H_n(\theta_0)\right]F_n^{-1}\right]^{-1}\left(F_nH_n(\theta_0)^{-1}F_n\right)F_n^{-1}S_n(\theta_0).$$

From Lemma 7.2 (d), the matrix $F_nH_n(\theta_0)^{-1}F_n = O_p(1)$ and is singular, and the matrix $F_n^{-1}\left[H_n^* - H_n(\theta_0)\right]F_n^{-1}$ is $o_p(1)$. Then

$$F_n\left(\widehat{\theta}-\theta_0\right) = -\left(F_nH_n(\theta_0)^{-1}F_n\right)F_n^{-1}S_n(\theta_0) + o_p(1),$$

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 $\left(F_nH_n(\theta_0)^{-1}F_n\right)F_n^{-1}S_n(\theta_0)$ is $O_p(1)$.

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

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

7 Technical Supplement

7.1 Lemma (Averages of SV Functions) If $L(t)$ satisfies SSV, then for $B \geq 1$

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

where $\eta > 0$ is arbitrarily small.

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

$$\sum_{t=B}^{n} L(t) = \int_B^n L(t)\,dt + \frac{1}{2}\left[L(B) + L(n)\right] + \int_B^n \left\{t - [t] - \frac{1}{2}\right\} L'(t)\,dt. \tag{53}$$

Since

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

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

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

Then, the final term in (53) may be bounded as follows

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

It follows that

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

for any $\eta > 0$ as $n \to \infty$.

7.3 Lemma

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

7.4 Proof From 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 integration by parts, we find

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

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

giving the stated result.

7.5 Example In the logarithmic case $L(t) = \log t$, and $\varepsilon(t) = \frac{1}{\log t}$ and $\eta(t) = -\frac{1}{\log t}$. Lemma 7.3 then gives the expansion

$$
\begin{aligned}
\int_1^n \log^k t dt &= n\log^k n - kn\log^{k-1} n + k^2 n\log^{k-2} n - kn\log^{k-2} n + o\left(n\log^{k-2} n\right) \\
&= n\log^k n - kn\log^{k-1} n + k(k-1)n\log^{k-2} n + o\left(n\log^{k-2} n\right), \qquad (54)
\end{aligned}
$$

whereas successive integration by parts (Lemma 7.9) 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 (54) is accurate to the third order.

7.6 Lemma

$$
\int_1^n \left[L(t) - \overline{L}\right]^2 dt = nL(n)^2 \varepsilon(n)^2 + o\left(nL(n)^2 \varepsilon(n)\left[\eta(n)+\varepsilon(n)\right]\right)
$$

7.7 Proof Applying the expansion from Lemma 7.3, we get

$$
\begin{aligned}
&\int_1^n \left[L(t) - \overline{L}\right]^2 dt \\
&= \int_1^n L(t)^2 dt - \frac{1}{n}\left(\int_1^n L(t) dt\right)^2 \\
&= \left[nL(n)^2 - 2nL(n)^2 \varepsilon(n) + 4nL(n)^2 \varepsilon(n)^2 - 2nL(n)^2 \varepsilon(n)\eta(n) + o\left(nL(n)^2 \varepsilon(n)\eta(n)\right)\right] \\
&\quad -\frac{1}{n}\left[nL(n) - nL(n)\varepsilon(n) + nL(n)\varepsilon(n)^2 - nL(n)\varepsilon(n)\eta(n) + o\left(nL(n)\varepsilon(n)\eta(n)\right)\right]^2 \\
&= \left[nL(n)^2 - 2nL(n)^2 \varepsilon(n) + 4nL(n)^2 \varepsilon(n)^2 - 2nL(n)^2 \varepsilon(n)\eta(n) + o\left(nL(n)^2 \varepsilon(n)^2\right)\right] \\
&\quad -\left[nL(n)^2 - 2nL(n)^2 \varepsilon(n) + 3nL(n)^2 \varepsilon(n)^2 - 2nL(n)^2 \varepsilon(n)\eta(n) + o\left(nL(n)^2 \varepsilon(n)^2\right)\right] \\
&= nL(n)^2 \varepsilon(n)^2 + o\left(nL(n)^2 \varepsilon(n)\left[\eta(n)+\varepsilon(n)\right]\right).
\end{aligned}
$$

7.8 Lemma *If $L(t)$ is a smoothly varying and satisfies 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 $n \to \infty$.

7.9 Proof In view of SSV, we have

$$\log \frac{L(rn)}{L(n)} = -\int_{rn}^n \frac{\varepsilon(t)}{t} dt, \tag{55}$$

and, since $|\varepsilon(t)|$ is slowly varying, it follows by Karamata's theorem (e.g., Proposition 1.5.9a, p.26 of BGT) 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

$$\frac{L(rn)}{L(n)} - 1 = \exp\{\varepsilon(n) \log r [1 + o(1)]\} - 1 = \varepsilon(n) \log r [1 + o(1)]. \tag{56}$$

The function L is second order slowly varying (see de Haan and Resnick, 1995, for second order regular variation) in the sense that

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

Integration by parts gives

$$\int_0^1 \log^k r dr = (-1)^k k!, \tag{57}$$

and so

$$\begin{aligned} \int_0^1 \left[\frac{L(rn)}{L(n)} - 1\right]^k dr &= \varepsilon(n)^k \int_0^1 \log^k r dr [1 + o(1)] \\ &= (-1)^k \varepsilon(n)^k k! [1 + o(1)], \end{aligned} \tag{58}$$

giving the stated result.

7.10 Lemma *If $L(t)$ satisfies SSV, then for all $r > 0$*

$$\frac{L(rn)}{L(n)} - 1 = \varepsilon(n) \log r + \frac{1}{2}\varepsilon(n) [\varepsilon(n) + \eta(n)] \log^2 r [1 + o(1)] \tag{59}$$

as $n \to \infty$.

7.11 Proof Since both L and ε are SSV functions we have both (55) and

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

and, as in (56) above, we get for ε

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

Then

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

and we deduce that

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

as stated.

7.12 Example $L(n) = \frac{1}{\log n}$, $\varepsilon(n) = -\frac{1}{\log n}$, $\eta(n) = -\frac{1}{\log n}$. 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 (59) above.

7.13 Lemma

$$\int_{1}^{n} \log^k t dt = n \sum_{j=0}^{k} (-k)_j \log^{k-j} n$$

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

7.14 Proof This follows by successive integration by parts.

7.15 Lemma

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

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

7.16 Proof Both parts follow by direct calculation. First,

$$\begin{aligned}\sum_{j=0}^{p-k-1}\binom{p}{p-j}\binom{p-j}{k}(-1)^{p-j} &= \sum_{j=0}^{p-k-1}\frac{p!}{j!k!\,(p-j-k)!}(-1)^{p-j} = \binom{p}{k}\sum_{j=0}^{p-k-1}\frac{(p-k)!}{j!\,(p-k-j)!}(-1)^{p-j} \\ &= \binom{p}{k}\left[\sum_{j=0}^{p-k}\frac{(p-k)!}{j!\,(p-k-j)!}(-1)^{p-j} - (-1)^{p-k}\right] = -(-1)^{p-k}\binom{p}{k},\end{aligned}$$

giving (a). Second,

$$\sum_{j=k}^{p}(-1)^{p-j}\binom{p}{j}\binom{j}{k} = \binom{p}{k}\sum_{j=k}^{p}(-1)^{p-j}\frac{(p-k)!}{(p-j)!\,(j-k)!} = (-1)^{k}\binom{p}{k}\sum_{\ell=0}^{p-k}(-1)^{p-\ell}\binom{p-k}{\ell} = 0$$

giving (b).

7.17 Lemma

(a)

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

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

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

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

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

7.18 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 $H_pF_p^2H_p'$ and let $j=k\le i$. By direct calculation and using the representation

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

we find that this element is

$$\begin{aligned}
&\sum_{m=0}^{(i-1)\wedge(j-1)} (-1)^{i-1+m}\binom{i-1}{m}(i-1)!\,(j-1)!\binom{j-1}{m}(-1)^{j-1+m} \\
=\;& (-1)^{i+k}(i-1)!\,(k-1)!\sum_{m=0}^{k-1}(-1)^{2m}\frac{(1-i)_m\,(1-k)_m}{m!\,(1)_m} \\
=\;& (-1)^{i+k}(i-1)!\,(k-1)!\;{}_2F_1(1-i,1-k,1;1), \qquad (60)
\end{aligned}$$

where ${}_2F_1(a,b,c;z)=\sum_{j=0}^{\infty}\frac{(a)_j(b)_j}{j!(c)_j}z^j$ 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, (60) 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 (a). Parts (b) and (c) follow directly.

8 Proofs

8.1 Proof of Lemma 2.3

Part (a). By partial summation we have

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}L(t)\,u_t = L(n)\frac{S_n}{\sqrt{n}}-\frac{1}{\sqrt{n}}\sum_{t=1}^{n}[L(t)-L(t-1)]\,S_{t-1}, \qquad (61)$$

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

$$\begin{aligned}
\frac{1}{\sqrt{n}L(n)}\sum_{t=1}^{n}L(t)\,u_t &= \frac{S_n}{\sqrt{n}}-\frac{1}{\sqrt{n}L(n)}\sum_{t=1}^{n}[L(t)-L(t-1)]\,S_{t-1} \\
&= \frac{S_n}{\sqrt{n}}-\frac{1}{L(n)}\sum_{t=1}^{n}\left[L\left(n\frac{t}{n}\right)-L\left(n\frac{t-1}{n}\right)\right]\frac{S_{t-1}}{\sqrt{n}}. \qquad (62)
\end{aligned}$$

We now use the embedding of the standardized partial sum $\frac{S_{t-1}}{\sqrt{n}}$ in Brownian motion given in LP(b), viz.

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

Then

$$\frac{1}{L(n)}\sum_{t=1}^{n}\left[L\left(n\frac{t}{n}\right)-L\left(n\frac{t-1}{n}\right)\right]\frac{S_{t-1}}{\sqrt{n}}=\frac{1}{L(n)}\int_0^1 B(r)\,dL(nr)+o_{a.s.}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right). \tag{63}$$

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

$$\frac{2}{L(n)^2}\int_0^1\int_0^r s\,dL(ns)\,dL(nr). \tag{64}$$

Observe that

$$\begin{aligned}\int_0^r s\,dL(ns) &= \int_0^r nsL'(ns)\,ds=\int_0^r L(ns)\,\varepsilon(ns)\,ds\\ &= \frac{1}{n}\int_0^{nr}L(t)\,\varepsilon(t)\,dt=\frac{1}{n}\left[nrL(nr)\,\varepsilon(nr)+o\left(nrL(nr)\,\varepsilon(nr)\right)\right]\end{aligned} \tag{65}$$

For the last equality, note that $L(t)\,\varepsilon(t)$ is (up to sign) a smoothly slowly varying function. We can then use Karamata's theorem,viz. that for $\alpha>-1$ and a slowly varying function ℓ, we have the asymptotic equivalence

$$\int_a^x t^{\alpha}\ell(t)\,dt\sim\frac{x^{\alpha+1}}{\alpha+1}\ell(x)\ \text{ as } x\to\infty \tag{66}$$

(e.g., BGT, Proposition 1.5.8, p. 26), setting $\alpha=0$ to obtain (65). Using (65) in (64), the dominant term is

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

by applying (66) again. It follows that

$$\frac{1}{L(n)}\int_0^1 B(r)\,dL(nr)=o_p(1) \tag{67}$$

as $n\to\infty$. We deduce from (62), (63) and (67) that

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

Part (b) We have

$$\sum_{t=1}^{n}\left(L(t)-\overline{L}\right)u_t=\sum_{t=1}^{n}L(t)\,u_t-\overline{L}S_n$$

and Lemma 7.3 gives

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

and using (62) and (63)

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

so that

$$\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(L(t)-\overline{L}\right)u_t\\
=\;& L(n)\frac{S_n}{\sqrt{n}}-\int_0^1 B(r)\,dL(nr)+o_{a.s.}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)\\
&-\left[L(n)-L(n)\,\varepsilon(n)+L(n)\,\varepsilon(n)^2-L(n)\,\varepsilon(n)\,\eta(n)+o\left(L(n)\,\varepsilon(n)\,\eta(n)\right)\right]\frac{S_n}{\sqrt{n}}\\
=\;& -\int_0^1 B(r)\,dL(nr)+o_{a.s.}\left(\frac{1}{n^{\frac{1}{2}-\frac{1}{p}}}\right)+L(n)\,\varepsilon(n)\frac{S_n}{\sqrt{n}}+O_p\left(L(n)\,\varepsilon(n)^2\right)\\
=\;& -\int_0^1\frac{B(r)}{r}\frac{L'(nr)\,nr}{L(nr)}L(nr)\,dr+L(n)\,\varepsilon(n)\int_0^1 dB(r)+O_p\left(L(n)\,\varepsilon(n)^2\right)\\
=\;& -\int_0^1\frac{B(r)}{r}\varepsilon(nr)\,L(nr)\,dr+L(n)\,\varepsilon(n)\int_0^1 dB(r)+O_p\left(L(n)\,\varepsilon(n)^2\right). \qquad (68)
\end{aligned}$$

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 (66), we have

$$\begin{aligned}
\int_0^1\frac{B(r)}{r}\varepsilon(nr)\,L(nr)\,dr &= \varepsilon(n)\,L(n)\int_0^1\frac{B(r)}{r}dr\,[1+o_p(1)]\\
&= -\varepsilon(n)\,L(n)\int_0^1(\log r)\,dB(r)\,[1+o_p(1)]. \qquad (69)
\end{aligned}$$

It follows from (69) that (68) is

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

as stated.

Part (c) Start by considering $n^{-\frac{1}{2}}\sum_{t=1}^{n}K_{nj}\left(\frac{t}{n}\right)u_t$. By partial summation and the strong approximation LP(b), we obtain, as in (63),

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

From (56) 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,

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

as required.

8.2 Proof of Theorem 4.1 Using Euler summation with $f(t)=\left[\frac{L(t)}{L(n)}-1\right]^j$ we obtain, as in Lemma 7.1,

$$\begin{aligned}&\frac{1}{n}\sum_{t=1}^{n}K_{nj}\left(\frac{t}{n}\right)\\ =\;& \frac{1}{n}\sum_{t=1}^{n}\left[\frac{L(t)}{L(n)}-1\right]^j\\ =\;& \frac{1}{n}\int_1^n\left[\frac{L(t)}{L(n)}-1\right]^j dt+\frac{1}{2n}\{f(1)+f(n)\}+\frac{j}{n}\int_1^n\left\{t-[t]-\frac{1}{2}\right\}\left[\frac{L(t)}{L(n)}-1\right]^{j-1}\frac{L'(t)}{L(t)}dt\\ =\;& \frac{1}{n}\int_1^n\left[\frac{L(t)}{L(n)}-1\right]^j dt+O\left(\frac{1}{n^{1-\delta}}\right)\\ =\;& \int_{\frac{1}{n}}^1\left[\frac{L(rn)}{L(n)}-1\right]^j dr+O\left(\frac{1}{n^{1-\delta}}\right)=\int_0^1\left[\frac{L(rn)}{L(n)}-1\right]^j dr+O\left(\frac{1}{n^{1-\delta}}\right),\end{aligned}\tag{70}$$

for arbitrarily small $\delta>0$, in view of (6). Hence, from (58) in the proof of Lemma 7.8, 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

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

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

To prove part (b), we note by Lemma 2.3 (c) 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 above elements for $j=0,1,...,p$, we get

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

By the Cramér - Wold device, we deduce that

$$\frac{1}{\sqrt{n}}D_{n\varepsilon}^{-1}\sum_{t=1}^{n}K_n\left(\frac{t}{n}\right)u_t\to_d\int_0^1\ell_p(r)dB(r),\quad \ell_p(r)=\left(1,\log r,...,\log^p r\right). \tag{71}$$

Then, from (31), (32) and (71) we obtain

$$\sqrt{n}D_{n\varepsilon}\left[\widehat{\alpha}_n-\alpha_n\right]=\left[\frac{1}{n}D_{n\varepsilon}^{-1}\sum_{t=1}^{n}K_n\left(\frac{t}{n}\right)K_n\left(\frac{t}{n}\right)'D_{n\varepsilon}^{-1}\right]^{-1}\left[\frac{1}{\sqrt{n}}D_{n\varepsilon}^{-1}\sum_{t=1}^{n}K_n\left(\frac{t}{n}\right)u_t\right]\to_d N\left(0,\sigma^2V^{-1}\right).$$

8.3 Proof of Theorem 4.2 From (29) and (33), we get for the final coefficient

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

where $V^{-1}=\left(v^{i,j}\right)$ and $V=\int_0^1\ell_p(r)\ell_p(r)' dr$. A calculation (see Lemma 7.17(c)) 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

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

and so

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

leading to

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

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

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

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

$$\alpha_{nk} = \sum_{j=k}^{p}\beta_j L(n)^j\binom{j}{k} = \binom{k}{k}\beta_k L(n)^k + \binom{k+1}{k}\beta_{k+1}L(n)^{k+1} + \ldots + \binom{p}{k}\beta_p L(n)^p \quad (72)$$

so that

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

We establish by induction (for decreasing k) that

$$\sqrt{n}L(n)^k \varepsilon(n)^p \left(\widehat{\beta}_k - \beta_k\right) = (-1)^{p-k} \binom{p}{k} \sqrt{n}L(n)^p \varepsilon(n)^p \left(\widehat{\beta}_p - \beta_p\right) + o_p(1). \quad (73)$$

We have already shown (73) to be valid for $k = p-1$ and $p-2$. Assume it is valid for $k+1$. Then, using (72) and Lemma 7.13 we have

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

showing that the result holds for k as well.

Equation (73) gives an asymptotic correspondence between the elements of the least squares estimate $\widehat{\beta}$ and its final component $\widehat{\beta}_p$ that has the form

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

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

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

giving the stated result. The explicit formula $v^{p+1,p+1} = 1/(p!)^2$ follows from Lemma 7.17(c).

8.4 Proof of Theorem 4.3 First, transform the regressor space in (25) as follows

$$y_s = \beta' L_s + u_s = \beta' J_n J_n^{-1} L_s + u_s = a'_n X_s + u_s, \quad (74)$$

where

$$\begin{aligned}
J'_n &= \begin{bmatrix}
1 & L(n) & L(n)^2 & \cdots & L(n)^{p-1} & L(n)^p \\
0 & L(n)\varepsilon(n) & \binom{2}{1}L(n)^2\varepsilon(n) & \cdots & \binom{p-1}{1}L(n)^{p-1}\varepsilon(n) & \binom{p}{1}L(n)^p\varepsilon(n) \\
0 & 0 & L(n)^2\varepsilon(n)^2 & \cdots & \binom{p-1}{2}L(n)^{p-1}\varepsilon(n) & \binom{p}{2}L(n)^p\varepsilon(n)^2 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & L(n)^{p-1}\varepsilon(n)^{p-1} & \binom{p}{p-1}L(n)^p\varepsilon(n)^{p-1} \\
0 & 0 & 0 & \cdots & 0 & L(n)^p\varepsilon(n)^p
\end{bmatrix} \\
&= E_n H' D_{nL}, \text{ say}
\end{aligned}$$

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

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

In (74) $a_n = J_n'\beta = E_n\alpha_n$ where α_n is the parameter vector in (30) whose elements α_{nj} are given in (27)-(29). Since $X_s = J_n^{-1}L_s = E_n^{-1}H^{-1}D_{nL}^{-1}L_s$, we may rewrite (74) as

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

In view of (56), the vector $E_n^{-1}K_n\left(\frac{s}{n}\right)$ 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

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

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

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

Since $\varepsilon(n) \to 0$ as $n \to \infty$, the matrix $E_n = e_1e_1' + o(1)$, where $e_1 = (1, 0, ..., 0)'$, and final expression above is

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

where i_{p+1} is the $p+1$ sum vector (i.e., it has unity in each component). This gives the first result (35).

Next, consider the inverse sample moment matrix

$$\begin{aligned}
\left(L'L\right)^{-1} &= J_n^{-1\prime}\left[\sum_{s=1}^{n}\ell_p\left(\frac{s}{n}\right)\ell_p\left(\frac{s}{n}\right)'\right]^{-1}J_n^{-1}\left[1+o\left(1\right)\right]\\
&= \frac{1}{n}J_n^{-1\prime}\left[\int_0^1\ell_p\left(r\right)\ell_p\left(r\right)'dr\right]^{-1}J_n^{-1}\left[1+o\left(1\right)\right]\\
&= \frac{1}{n}D_{nL}^{-1}H^{-1\prime}E_n^{-1}\left[\int_0^1\ell_p\left(r\right)\ell_p\left(r\right)'dr\right]^{-1}E_n^{-1}H^{-1}D_{nL}^{-1}\left[1+o\left(1\right)\right]
\end{aligned}$$

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

$$\begin{aligned}
&\frac{1}{n\varepsilon\left(n\right)^{2p}}D_{nL}^{-1}H^{-1\prime}e_{p+1}e_{p+1}'\left[\int_0^1\ell_p\left(r\right)\ell_p\left(r\right)'dr\right]^{-1}e_{p+1}e_{p+1}'H^{-1}D_{nL}^{-1}\\
=\;&\frac{e_{p+1}'\left[\int_0^1\ell_p\left(r\right)\ell_p\left(r\right)'dr\right]^{-1}e_{p+1}}{n\varepsilon\left(n\right)^{2p}}D_{nL}^{-1}H^{-1\prime}e_{p+1}e_{p+1}'H^{-1}D_{nL}^{-1}\\
=\;&\frac{e_{p+1}'\left[\int_0^1\ell_p\left(r\right)\ell_p\left(r\right)'dr\right]^{-1}e_{p+1}}{n\varepsilon\left(n\right)^{2p}}D_{nL}^{-1}\mu_{p+1}\mu_{p+1}'D_{nL}^{-1}, \qquad (75)
\end{aligned}$$

giving the stated result. Line (75) holds because $H^{-1\prime}e_{p+1} = \mu_{p+1}$, the final column of $H^{-1\prime}$, as is apparent from the fact that $H'\mu_{p+1} = e_{p+1}$ which can be verified by direct multiplication using Lemma 7.12(b).

8.5 Proof of Theorem 5.1 Solving (40) for β_1 and β_2, we get

$$\begin{aligned}
\begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix} &= \begin{bmatrix}L_1\left(n\right) & 0\\ 0 & L_2\left(n\right)\end{bmatrix}^{-1}\begin{bmatrix}\varepsilon_1\left(n\right) & \varepsilon_2\left(n\right)\\ \frac{1}{2}\varepsilon_1\left(n\right)\left[\varepsilon_1\left(n\right)+\eta_1\left(n\right)\right] & \frac{1}{2}\varepsilon_2\left(n\right)\left[\varepsilon_2\left(n\right)+\eta_2\left(n\right)\right]\end{bmatrix}^{-1}\begin{bmatrix}\alpha_{n1}\\ \alpha_{n2}\end{bmatrix}\\
&= \begin{bmatrix}L_1\left(n\right)\varepsilon_1\left(n\right) & 0\\ 0 & L_2\left(n\right)\varepsilon_2\left(n\right)\end{bmatrix}^{-1}\begin{bmatrix}1 & 1\\ \frac{1}{2}\left[\varepsilon_1\left(n\right)+\eta_1\left(n\right)\right] & \frac{1}{2}\left[\varepsilon_2\left(n\right)+\eta_2\left(n\right)\right]\end{bmatrix}^{-1}\begin{bmatrix}\alpha_{n1}\\ \alpha_{n2}\end{bmatrix}\\
&= \begin{bmatrix}L_1\left(n\right)\varepsilon_1\left(n\right) & 0\\ 0 & L_2\left(n\right)\varepsilon_2\left(n\right)\end{bmatrix}^{-1}\frac{2}{\delta\left(n\right)}\begin{bmatrix}\frac{1}{2}\left[\varepsilon_2\left(n\right)+\eta_2\left(n\right)\right] & -1\\ -\frac{1}{2}\left[\varepsilon_1\left(n\right)+\eta_1\left(n\right)\right] & 1\end{bmatrix}\begin{bmatrix}\alpha_{n1}\\ \alpha_{n2}\end{bmatrix}.
\end{aligned}$$

and so

$$\frac{\sqrt{n}}{2}\begin{bmatrix}\delta\left(n\right)\varepsilon_1\left(n\right)L_1\left(n\right)\left[\widehat{\beta}_1-\beta_1\right]\\ \delta\left(n\right)\varepsilon_2\left(n\right)L_2\left(n\right)\left[\widehat{\beta}_2-\beta_2\right]\end{bmatrix} = \begin{bmatrix}\frac{1}{2}\left[\varepsilon_2\left(n\right)+\eta_2\left(n\right)\right] & -1\\ -\frac{1}{2}\left[\varepsilon_1\left(n\right)+\eta_1\left(n\right)\right] & 1\end{bmatrix}\begin{bmatrix}\sqrt{n}\left[\widehat{\alpha}_{n1}-\alpha_{n1}\right]\\ \sqrt{n}\left[\widehat{\alpha}_{n2}-\alpha_{n2}\right]\end{bmatrix}.$$

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

$$\frac{\sqrt{n}}{2}\begin{bmatrix}\delta\left(n\right)\varepsilon_1\left(n\right)L_1\left(n\right)\left[\widehat{\beta}_1-\beta_1\right]\\ \delta\left(n\right)\varepsilon_2\left(n\right)L_2\left(n\right)\left[\widehat{\beta}_2-\beta_2\right]\end{bmatrix} = \begin{bmatrix}-1\\ 1\end{bmatrix}\left[\sqrt{n}\left[\widehat{\alpha}_{n2}-\alpha_{n2}\right]\right]+o_p\left(1\right)\to_d N\left(0,\frac{\sigma^2}{\left(2!\right)^2}\begin{bmatrix}1 & -1\\ -1 & 1\end{bmatrix}\right), \qquad (76)$$

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

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

which, in combination with (76), leads to

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

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

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

which gives the stated result upon scaling.

8.6 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 2.3 (a) have

$$\frac{1}{\sqrt{n}L(n)}\sum_{t=1}^n f\left(\frac{t}{n};\theta\right)L(t)\,u_t = f(1;\theta)\,\frac{S_n}{\sqrt{n}} - \frac{1}{L(n)}\sum_{t=1}^n\left[f\left(\frac{t}{n};\theta\right)L\left(n\frac{t}{n}\right) - f\left(\frac{t-1}{n};\theta\right)L\left(n\frac{t-1}{n}\right)\right]\frac{S_{t-1}}{\sqrt{n}}. \tag{77}$$

Assume the probability space is constructed so that we can embed the standardized partial sum $\frac{S_{t-1}}{\sqrt{n}}$ in Brownian motion as in LP(b), viz.

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

Then, the first term of (77) clearly satisfies

$$f(1;\theta)\,\frac{S_n}{\sqrt{n}} \to_d f(1;\theta_0)\,B(1)\,, \tag{78}$$

as $n \to \theta$ uniformly over $\theta \in N_n^0$. The second term of (77) is

$$
\begin{aligned}
& \frac{1}{L(n)} \sum_{t=1}^{n} \left[f\left(\frac{t}{n};\theta\right) L\left(n\frac{t}{n}\right) - f\left(\frac{t-1}{n};\theta\right) L\left(n\frac{t-1}{n}\right)\right] \frac{S_{t-1}}{\sqrt{n}} \\
= & \frac{1}{L(n)} \sum_{t=1}^{n} f\left(\frac{t}{n};\theta\right) \left[L\left(n\frac{t}{n}\right) - L\left(n\frac{t-1}{n}\right)\right] \frac{S_{t-1}}{\sqrt{n}} \\
& + \frac{1}{L(n)} \sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) \left[f\left(\frac{t}{n};\theta\right) - f\left(\frac{t-1}{n};\theta\right)\right] \frac{S_{t-1}}{\sqrt{n}} \\
= & T_1 + T_2. \qquad (79)
\end{aligned}
$$

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

$$
\begin{aligned}
& \sup_{\theta\in N_n^0} \left| \frac{1}{L(n)} \sum_{t=1}^{n} \left[f\left(\frac{t}{n};\theta\right) - f\left(\frac{t}{n};\theta_0\right)\right] \left[L\left(n\frac{t}{n}\right) - L\left(n\frac{t-1}{n}\right)\right] \frac{S_{t-1}}{\sqrt{n}} \right| \\
= & \sup_{\theta\in N_n^0} \left| \sum_{t=1}^{n} \left[f_\theta\left(\frac{t}{n};\theta^*\right)\right] \frac{\left[L\left(n\frac{t}{n}\right) - L\left(n\frac{t-1}{n}\right)\right]}{L(n)} \frac{S_{t-1}}{\sqrt{n}} \right| |\theta-\theta_0| \\
\leq & \sup_{\theta\in N_n^0} |\theta-\theta_0| \frac{1}{n} \sum_{t=1}^{n} \sup_{\theta\in N_n^0} \left| f_\theta\left(\frac{t}{n};\theta^*\right)\right| \left|\frac{L'(t^*)}{L(n)}\right| \left|\frac{S_{t-1}}{\sqrt{n}}\right| \\
= & \sup_{\theta\in N_n^0} |\theta-\theta_0| \frac{1}{n} \sum_{t=1}^{n} g_\theta\left(\frac{t}{n};\theta_0\right) \left|\frac{L'(t^*)}{L(n)}\right| \left|\frac{S_{t-1}}{\sqrt{n}}\right| \\
= & \sup_{\theta\in N_n^0} |\theta-\theta_0| \left[\int_0^1 g_\theta(r;\theta_0) \left|\frac{B(r)}{r}\right| \left|\frac{L'(nr)\,nr}{L(nr)}\right| \left|\frac{L(nr)}{L(n)}\right| dr + o_p(1)\right] \\
= & o_p(1), \qquad (80)
\end{aligned}
$$

as

$$
\left|\frac{L(nr)}{L(n)}\right| \to 1, \quad \left|\frac{L'(nr)\,nr}{L(nr)}\right| \to 0, \text{ for all } r>0 \text{ and } \sup_{\theta_1,\theta_2\in N_n^0} |\theta_1-\theta_2| \to 0
$$

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 2.3 (a),

$$
\frac{1}{L(n)} \sum_{t=1}^{n} f\left(\frac{t}{n};\theta_0\right) \left[L\left(n\frac{t}{n}\right) - L\left(n\frac{t-1}{n}\right)\right] \frac{S_{t-1}}{\sqrt{n}} = \frac{1}{L(n)} \int_0^1 f(r;\theta_0)\, B(r)\, dL(nr) = O_p(\varepsilon(n)) = o_p(1),
$$

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

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

$$
\begin{aligned}
T_2 &= \frac{1}{L(n)} \sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) \left[f\left(\frac{t}{n};\theta\right) - f\left(\frac{t-1}{n};\theta\right)\right] \frac{S_{t-1}}{\sqrt{n}} \\
&= \frac{1}{L(n)\,n} \sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) f_r\left(\frac{t^*}{n};\theta\right) \frac{S_{t-1}}{\sqrt{n}}. \qquad (81)
\end{aligned}
$$

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

$$\begin{aligned}
&\frac{1}{L(n)n}\sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) f_r\left(\frac{t^*}{n};\theta\right)\frac{S_{t-1}}{\sqrt{n}} - \frac{1}{L(n)n}\sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) f_r\left(\frac{t^*}{n};\theta_0\right)\frac{S_{t-1}}{\sqrt{n}} \\
= &\frac{1}{L(n)n}\sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right)\left[f_r\left(\frac{t^*}{n};\theta\right) - f_r\left(\frac{t^*}{n};\theta_0\right)\right]\frac{S_{t-1}}{\sqrt{n}} \\
= &\frac{1}{L(n)n}\sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) f_{r\theta}\left(\frac{t^*}{n};\theta^*\right)\frac{S_{t-1}}{\sqrt{n}}(\theta-\theta_0),
\end{aligned}$$

and, just as in (80) above, we find that

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

Moreover,

$$\frac{1}{L(n)n}\sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) f_r\left(\frac{t^*}{n};\theta_0\right)\frac{S_{t-1}}{\sqrt{n}} = \int_0^1 \frac{L(nr)}{L(n)} f(r;\theta_0)B(r)\,dr + o_p(1) \to_p \int_0^1 f(r;\theta_0)B(r)\,dr, \tag{82}$$

as $n\to\infty$. It follows from (81) and (82) that

$$T_2 = \frac{1}{L(n)n}\sum_{t=1}^{n} L\left(n\frac{t-1}{n}\right) f_r\left(\frac{t^*}{n};\theta_0\right)\frac{S_{t-1}}{\sqrt{n}} + o_p(1) \to_d \int_0^1 f_r(r;\theta_0)B(r)\,dr, \tag{83}$$

uniformly over $\theta\in N_n^0$. We deduce from (77), (79), (78) and (83) that

$$\begin{aligned}
\frac{1}{\sqrt{n}L(n)}\sum_{t=1}^{n} f\left(\frac{t}{n};\theta\right)L(t)u_t &= f(1;\theta_0)\frac{S_n}{\sqrt{n}} - (T_1+T_2) \to_d f(1;\theta_0)B(1) - \int_0^1 f_r(r;\theta_0)B(r)\,dr \\
&= \int_0^1 f(r;\theta_0)\,dB(r),
\end{aligned} \tag{84}$$

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

8.7 Proof of Lemma 6.2

Part (a) For any slowly varying function L satisfying SSV and any function $f\in C^1$, we can show in the same way as Lemma 6.1 that

$$\frac{1}{\sqrt{n}L(n)}\sum_{s=1}^{n} f\left(\frac{s}{n};\theta_0\right)L(s)u_s \to_d \int_0^1 f(r;\theta_0)\,dB(r) = N\left(0,\sigma^2\int_0^1 f(r)^2\,dr\right), \tag{85}$$

extending (9). The limit (51) follows directly.

Part (b) Using Lemma 6.1 and (85), the asymptotic form of $D_n^{-1}H_n(\theta_0)D_n^{-1}$ is

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

$$\begin{aligned}
&= \begin{bmatrix} \frac{1}{n}\sum_{s=1}^n \left(\frac{s}{n}\right)^{2\gamma_0} & \frac{1}{n\log n}\sum_{s=1}^n \beta_0\left(\frac{s}{n}\right)^{2\gamma_0}\log s \\ \frac{1}{n\log n}\sum_{s=1}^n \beta_0\left(\frac{s}{n}\right)^{2\gamma_0}\log s & \frac{1}{n\log^2 n}\sum_{s=1}^n (\beta_0)^2\left(\frac{s}{n}\right)^{2\gamma_0}\log^2 s \end{bmatrix} + O_p\left(\frac{1}{\sqrt{n}}\right) \\
&= \begin{bmatrix} \int_0^1 r^{2\gamma_0}dr & \beta_0\int_0^1 r^{2\gamma_0}dr + \frac{\beta_0}{\log n}\int_0^1 r^{2\gamma_0}\log r dr \\ \beta_0\int_0^1 r^{2\gamma_0}dr + \frac{\beta_0}{\log n}\int_0^1 r^{2\gamma_0}\log r dr & \frac{\beta_0^2}{\log^2 n}\int_0^1 r^{2\gamma_0}(\log n+\log r)^2 dr \end{bmatrix} \qquad (86) \\
&\quad + O_p\left(\frac{1}{\sqrt{n}}\right) \\
&\to_p \begin{bmatrix} \int_0^1 r^{2\gamma_0}dr & \beta_0\int_0^1 r^{2\gamma_0}dr \\ \beta_0\int_0^1 r^{2\gamma_0}dr & \beta_0^2\int_0^1 r^{2\gamma_0}dr \end{bmatrix} = \frac{1}{2\gamma_0+1}\begin{bmatrix} 1 & \beta_0 \\ \beta_0 & \beta_0^2 \end{bmatrix},
\end{aligned}$$

as stated.

Part (c) Upon calculation of (86) and rescaling, we deduce that

$$\begin{aligned}
F_n^{-1}H_n(\theta_0)F_n^{-1} &= \log^2 n D_n^{-1}H_n(\theta_0)D_n^{-1} \\
&= \frac{\log^2 n}{2\gamma_0+1}\begin{bmatrix} 1 & \beta_0 - \frac{\beta_0}{(2\gamma_0+1)\log n} \\ \beta_0 - \frac{\beta_0}{(2\gamma_0+1)\log n} & \beta_0^2 + \frac{2\beta_0^2}{(2\gamma_0+1)\log n} + \frac{2\beta_0^2}{(2\gamma_0+1)^2\log^2 n} \end{bmatrix} + o_p(1) \\
&= \frac{\log^2 n}{2\gamma_0+1}\left\{\begin{bmatrix} 1 & \beta_0 \\ \beta_0 & \beta_0^2 \end{bmatrix} + \begin{bmatrix} 0 & -\frac{\beta_0}{(2\gamma_0+1)\log n} \\ -\frac{\beta_0}{(2\gamma_0+1)\log n} & \frac{2\beta_0^2}{(2\gamma_0+1)\log n} \end{bmatrix}\right\} + O_p(1),
\end{aligned}$$

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

Part (d) First calculate

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

Then

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

We deduce that

$$
\begin{aligned}
& F_n H_n(\theta_0)^{-1} F_n \\
= & \frac{\left(2\gamma^0+1\right)^3}{\beta_0^2}\left[\begin{array}{ll} \beta_0^2 - \frac{2\beta_0^2}{\log n} \frac{1}{(2\gamma^0+1)} + \frac{\beta_0^2}{\log^2 n} \frac{2}{(2\gamma^0+1)^2} & -\beta_0 + \frac{\beta_0}{\log n} \frac{1}{(2\gamma^0+1)} \\ -\beta_0 + \frac{\beta_0}{\log n} \frac{1}{(2\gamma^0+1)} & 1 \end{array}\right] + O_p\left(\frac{1}{\sqrt{n}}\right) \qquad (87) \\
\rightarrow_p & \frac{\left(2\gamma^0+1\right)^3}{\beta_0^2}\left[\begin{array}{ll} \beta_0^2 & -\beta_0 \\ -\beta_0 & 1 \end{array}\right],
\end{aligned}
$$

as given.

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

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

and define the matrix

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

We show that

$$\sup_{\theta \in N_n^0}\left|\left|C_n^{-1}\left[H_n(\theta) - H_n(\theta_0)\right] C_n^{-1}\right|\right| = o_p(1). \qquad (88)$$

Note that in N_n^0 we have

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

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

$$\sup_{\theta \in N_n^0} |r^{\gamma}| \leq r^{-\frac{1}{2} + \varepsilon}. \qquad (89)$$

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

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

Next

$$\frac{1}{n\log n}\sum_{s=1}^{n}\left(\frac{s}{n}\right)^{2\gamma^*}\log s=\frac{1}{n}\sum_{s=1}^{n}\left(\frac{s}{n}\right)^{2\gamma^*}+\frac{1}{n\log n}\sum_{s=1}^{n}\left(\frac{s}{n}\right)^{2\gamma^*}\log\frac{s}{n}\rightarrow\int_0^1 r^{2\gamma_0}dr$$

uniformly for $\theta\in N_n^0$ in view of the majorization (89). It follows that for large enough n

$$\begin{aligned} \sup_{\theta\in N_n^0}\left|\frac{1}{n^{2\gamma_0+1+2\delta}}\sum_{s=1}^{n}\left(s^{2\gamma}-s^{2\gamma_0}\right)\right| &= O\left(\frac{2\log n}{n^{2(\gamma_0-\gamma^*)+2\delta}}\sup_{\theta\in N_n^0}|\gamma-\gamma_0|\right) \\ &= O\left(\frac{1}{n^{\delta}}\frac{1}{n^{\gamma_0+\frac{1}{2}-\delta}\log n}\right)=o(1)\,. \end{aligned} \tag{90}$$

Next,

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

We find

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

in the same way as (90) above. For term T_2, in view of (85) we have for each $\gamma\in N_n^0$

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

By Lemma 7.1 we have

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

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

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

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

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

as $n\rightarrow\infty$. The argument for the term $a_{22,n}$ is entirely analogous and (88) therefore follows.

8.8 Proof of Theorem 6.4 Standard asymptotic arguments of nonlinear regression for nonstationary dependent time series (e.g., Wooldridge, 1994, theorem 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 (48) in an open, shrinking neighbourhood of θ_0 follows as in the proof of Wooldridge's theorem 8.1 using (49) and Lemma 7.2 (e). 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 6.3(a); (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 7.2(c), $F_n^{-1}H_n(\theta_0)F_n^{-1}$ is positive definite for all large n and has eigenvalues of order $O\left(\log^2 n\right)$ and $O(\log n)$ and the smallest eigenvalue $\lambda_{\min}\left(F_n^{-1}H_n(\theta_0)F_n^{-1}\right) = O(\log n) \to \infty$. With these changes, the remainder of the consistency argument in Wooldridge's theorem 8.1 holds and we obtain $F_n\left(\widehat{\theta}-\theta_0\right) = O_p(1)$.

Next, scaling the first order conditions (49), we have

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

and then

$$F_n\left(\widehat{\theta}-\theta_0\right) = -\left[I + \left(F_nH_n(\theta_0)^{-1}F_n\right)F_n^{-1}\left[H_n^* - H_n(\theta_0)\right]F_n^{-1}\right]^{-1}\left(F_nH_n(\theta_0)^{-1}F_n\right)F_n^{-1}S_n(\theta_0). \tag{91}$$

Note that $F_n = \frac{1}{\log n}D_n = diag\left[\frac{n^{\gamma_0+\frac{1}{2}}}{\log n}, n^{\gamma_0+\frac{1}{2}}\right]$ and since $D_n^{-1}S_n(\theta_0) = O_p(1)$ from Lemma 7.2(a), we have $F_n^{-1}S_n(\theta_0) = O_p(\log n)$. However, from (87) in the proof of Lemma 7.2(d) we have

$$\begin{aligned} & F_nH_n(\theta_0)^{-1}F_n \\ = \; & \frac{\left(2\gamma^0+1\right)^3}{\beta_0^2}\begin{bmatrix} \beta_0^2 - \frac{2\beta_0^2}{\log n}\frac{1}{(2\gamma^0+1)} + \frac{\beta_0^2}{\log^2 n}\frac{2}{(2\gamma^0+1)^2} & -\beta_0 + \frac{\beta_0}{\log n}\frac{1}{(2\gamma^0+1)} \\ -\beta_0 + \frac{\beta_0}{\log n}\frac{1}{(2\gamma^0+1)} & 1 \end{bmatrix} + O_p\left(\frac{1}{\sqrt{n}}\right) \\ = \; & O_p(1), \end{aligned} \tag{92}$$

and from Lemma 7.2(e) we have

$$\sup_{\theta\in N_n^0} F_n^{-1}\left[H_n(\theta) - H_n(\theta_0)\right]F_n^{-1} = o_p(1),$$

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 $\delta > 0$. Since $F_n\left(\widehat{\theta}-\theta_0\right) = O_p(1)$, it follows that $\widehat{\theta}, \theta^* \in N_n^0$ with probability approaching unity as $n \to \infty$, where θ^* is a generic mean value between $\widehat{\theta}$ and θ_0. Hence

$$F_n^{-1}\left[H_n^* - H_n(\theta_0)\right]F_n^{-1} = o_p(1), \tag{93}$$

and so, combining (93) and (92) we have

$$\left(F_nH_n(\theta_0)^{-1}F_n\right)F_n^{-1}\left[H_n^* - H_n(\theta_0)\right]F_n^{-1} = o_p(1). \tag{94}$$

Then, from (91), (94) and (87) we deduce that

$$
\begin{aligned}
& F_n\left(\widehat{\theta}-\theta_0\right) \\
= & -\left(F_n H_n(\theta_0)^{-1} F_n\right) F_n^{-1} S_n(\theta_0) + o_p(1) \\
= & -\frac{\left(2\gamma^0+1\right)^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\} \\
& \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) \\
= & -\frac{\left(2\gamma^0+1\right)^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 + o_p(1) \\
= & -\frac{\left(2\gamma^0+1\right)^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) \\
= & -\frac{\left(2\gamma^0+1\right)^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) \\
= & \left(2\gamma^0+1\right)^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) \\
\to_d & \left(2\gamma^0+1\right)^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\left(0, \sigma^2\left(2\gamma^0+1\right)^3\right),
\end{aligned}
$$

giving the stated result.

9 Notation

$\to_{a.s.}$	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
$(a)_k$	$a(a+1)\ldots(a+k-1)$	$r \wedge s$	$min(r,s)$
$B(r)$	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

10 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.

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

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

Geluk, J. L. De Haan, S. Resnick and C. Stărică (1997). "Second Order regular variation, convolution and the central limit theorem." *Stochastic Processes and their Applications*, 69, 139-159.

Grenander, U. and M. Rosenblatt (1957). *Statistical Analysis of Stationary Time Series.* New York: Wiley.

Haan, L. de and Resnick, S. (1995) "Second order regular variation and rates of converence in extreme value theory". *Annals of Probability*,

Hannan, E. J. (1970). *Multiple Time Series.* New York: John Wiley and Sons.

Hurvich, C. M., R. Deo and 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. (1956). *Infinite Sequences and Series.* New York: Dover.

Phillips, P. C. B. (1999): "Discrete Fourier Transforms of Fractional Processes", *Yale University, mimeographed.*

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

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

Shorack, G. R. And J. A. Wellner (1986). *Empirical Processes with Applications to Statistics.* New York: Wiley.

Wooldridge, J. M. (1994). "Estimation and Inference for Dependent Processes." In R. F. Engle and D. L. McFadden (Eds.) *Handbook of Econometrics. Vol. IV.* North Holland: Amsterdam.

Wu, C_F. (1981). "Asymptotic theory of nonlinear least squares estimation". *Annals of Statistics*, 9: 501-513.