Efficient Inference on Fractionally Integrated Panel Data Models with Fixed Effects

Peter M. Robinson*†
London School of Economics

Carlos Velasco‡
Universidad Carlos III de Madrid

March 27, 2013

Abstract

A dynamic panel data model is considered that contains possibly stochastic individual components and a common fractional stochastic time trend. We propose four different ways of coping with the individual effects so as to estimate the fractional parameter. Like models with autoregressive dynamics, ours nests a unit root, but unlike the nonstandard asymptotics in the autoregressive case, estimates of the fractional parameter can be asymptotically normal. Establishing this property is made difficult due to bias caused by the individual effects, or by the consequences of eliminating them, and requires the number of time series observations T to increase, while the cross-sectional size, N, can either remain fixed or increase with T. The biases in the central limit theorem are asymptotically negligible only under stringent conditions on the growth of N relative to T, but these can be relaxed by bias correction. For three of the estimates the biases depend only on the fractional parameter. In hypothesis testing, bias correction of the estimates is readily carried out. We evaluate the biases numerically for a range of T and parameter values, develop and justify feasible bias-corrected estimates, and briefly discuss simplified but less effective corrections. A Monte Carlo study of finite-sample performance is included.

JEL Classifications: C12, C13, C23

Keywords: Panel data, Fractional time series, Estimation, Testing, Bias correction

....

* Corresponding author. Tel. 44 20 7955 7516. fax 44 20 7955 6592

E-mail address: p.m.robinson @lse.ac.uk

†Research supported by a Cátedra de Excelencia at Universidad Carlos III de Madrid, Spanish Plan Nacional de I+d+I Grant SEJ2007-62908/ECON, and ESRC Grant ES/J007242/1

‡Research supported by Spanish Ministerio de Economía y Competitividad Grant ECO2012-31748.

1. Introduction

Consider the following unobserved components panel data model for an observable array $\{y_{it}\}$:

$$y_{it} = \alpha_i + \Delta_{t+1}^{-d_0} \varepsilon_{it}, \tag{1}$$

for $i = 1, \ldots, N$, $t = 0, 1, \ldots, T$. Here, the unobserved individual effects $\{\alpha_i, i \geq 1\}$ are random variables that are subject to little, if any, more detailed specification in the sequel; the array $\{\varepsilon_{it}, i \geq 1, t \geq 0\}$ consists of random variables that are throughout assumed to be independent and identically distributed (iid) and to satisfy $E\varepsilon_{it} = 0$, $E\varepsilon_{it}^4 < \infty$; d_0 is an unknown positive number; for any positive integer s and any real d,

$$\Delta_s^d = \sum_{j=0}^{s-1} \pi_j (d) L^j, \; \pi_j (d) = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)},$$

with L the lag operator, $\Gamma(d) = (-1)^d \infty$ for $d = 0, -1, \ldots$, and the convention $\Gamma(0)/\Gamma(0) = 1$. The Δ_s^d notation is due to the usual definition of the difference operator $\Delta = 1 - L$, and Δ_s^d is obtained by truncating the fractional operator which is given (at least formally) by

$$\Delta^d = \sum_{j=0}^{\infty} \pi_j (d) L^j.$$

We thus have

$$\Delta_{t+1}^{-d} \varepsilon_{it} = \Delta^{-d} \{\varepsilon_{it} 1 (t \geq 0)\},$$

where $1\{.\}$is the indicator function.

We can write (1) as

$$y_{it} = \alpha_i + \sum_{j=0}^{t} \pi_j (-d_0) \varepsilon_{i,t-j}.$$

A special case of (1) is heavily featured in the dynamic panel data literature: $d_0 = 1$, whence

$$y_{it} = \alpha_i + \sum_{j=0}^{t} \varepsilon_{i,t-j}. \tag{2}$$

In that literature, however, the unit root model (2) is nested in the autoregressive scheme

$$y_{it} = \alpha_i + \sum_{j=0}^{t} \rho^j \varepsilon_{i,t-j}. \tag{3}$$

The typical alternatives to $\rho = 1$ covered by (3) are the stationary ones $\rho \in (-1, 1)$ or the explosive ones $\rho > 1$. Other versions of the autoregressive panel data model are

$$y_{it} = \alpha_i + \rho y_{i,t-1} + \varepsilon_{it}, \; t > 0, \tag{4}$$

and

$$y_{it} = \alpha_i + u_{it}, \quad u_{it} = \rho u_{i,t-1} + \varepsilon_{it}, \; t > 0, \tag{5}$$

with $\rho \in (-1, 1]$; note that (5) implies that

$$y_{it} = (1-\rho) \alpha_i + \rho y_{i_i t-1} + \varepsilon_{it}, \; t > 0,$$

so that α_i is eliminated when $\rho = 1$. The usual aim in (3), (4) or (5) is estimating ρ or unit root testing. As one recent reference, Han and Phillips (2010) develop inference based on generalized method-of-moment estimates.

In the fractional model (1), the moving average weights have decay or growth that is, unlike in (3), not exponential but algebraic, since, for any d,

$$\pi_j(d) = \frac{1}{\Gamma(-d)} j^{-d-1}(1 + O(j^{-1})) \text{ as } j \to \infty. \tag{6}$$

As is well known from the time series literature the fractional class has a smoothness at the unit root (and elsewhere) that the autoregressive class lacks. A consequence established in that literature is that large sample inference based on an approximate Gaussian pseudo likelihood can be expected to entail standard limit distribution theory; in particular, Lagrange multiplier tests on d_0 are asymptotically χ^2 distributed with classical local power properties, and estimates of d_0 are asymptotically normally distributed with the usual parametric rate (see Robinson (1995), Beran (1994), Velasco and Robinson (2000), Hualde and Robinson (2011)). This is the case whether d_0 lies in the stationary region $(0, 1/2)$ or the nonstationary one $[1/2, \infty)$ (or, also, the negative dependent region $(-\infty, 0)$).

If N is regarded as fixed while $T \to \infty$, (1) is just a multivariate fractional model, with a vector, possibly stochastic, location. But in many practical applications N is large, and even when smaller than T, is more reasonably treated as diverging in asymptotic theory if T is. In that case inference on d_0 is considerably complicated by an incidental parameters problem. In this paper we present and justify several approaches that resolve this question. In (1) the interest is in estimating d_0 (efficiently, perhaps with some *a priori* knowledge on the range of allowed values) and testing hypotheses such as the unit root, $d_0 = 1$. It would be possible to incorporate exogenous variables that vary with t, or with i and t, perhaps in a linear regression framework, but here we stay with the simple model (1) to focus on the incidental parameters problem. In order to cope with this we throughout employ asymptotic theory with respect to T diverging, where either N increases with T or stays fixed, and both cases are covered by indexing with respect to T only. We allow for cross-sectional dependence and heteroscedasticity in the y_{it} via the α_i. However, conditional on the α_i the y_{it} are cross-sectionally iid. It would be straightforward to relax this requirement in case of fixed N, such as by allowing $(\varepsilon_{1t}, \ldots, \varepsilon_{Nt})$ to have an unrestricted covariance matrix. For increasing N the issues are more challenging, and there is a choice between on the one hand leaving the variance and covariance structure unrestricted, and on the other adopting a parametric form, such as a factor model or, when there is knowledge of spatial locations or differences, a spatial model. Such cross-sectional dependence raises questions of robust inference and efficient estimation, but we focus here on the bias issues prompted by (1), which would remain the same under cross-sectional dependence. It would be more straightforward to relax our assumptions on temporal dependence. The iid requirement over t of the ε_{it} could be weakened to martingale difference and mild homogeneity assumptions as in Hualde and Robinson (2011), but for aesthetic reasons we keep the conditions simple by matching the iid assumption across i. The dynamics in (1), like that in (3), (4) and (5), is extremely simple, and could be straightforwardly generalized to allow for parametric short memory dependence in ε_{it} (see again the previous reference), but we prefer to keep the setting simple in order to focus on the main ideas. Hassler, Demetrescu and Tarcolea (2011) have recently developed tests in a panel with a more general temporal dependence structure which is allowed to vary across units, and with allowance for cross-sectional dependence, but without allowing for individual effects and keeping N fixed as $T \to \infty$.

The following section introduces four rival estimates of d_0. Section 3 contains consistency theorems. In Section 4 the estimates are shown to be asymptotically normal. Unless the restriction on the growth of N relative to T is very stringent, asymptotic biases are present here. In Section 5 we describe the implications of our results for hypothesis testing and interval estimation, numerically compare biases of our estimates, justify feasible bias correction, and present also corrections that are simpler, albeit less effective. Section 6 consists of a Monte Carlo study of finite−sample performance of our methods. Theorem proofs appear in Appendix A. These depend in part on two Propositions, stated in Sections 3 and 4 but proved in Appendix B. Our proofs also use technical lemmas on properties of the $\pi_j(d)$, stated and proved in Appendix C; we draw attention here to Lemma 3, which is the main new technical tool, and is used in the consistency proofs.

2. Estimation of d_0

We consider four different, but asymptotically equivalent and efficient, methods of estimating d_0 in (1). All these estimates are implicitly-defined and entail optimization over a compact set $D = [\underline{d}, \overline{d}]$, where

$$\underline{d} > \max\left(0, d_0 - \frac{1}{2}\right), d_0 \in D, \tag{7}$$

which implies that $d_0 > 0$ and $d > d_0 - \frac{1}{2}$ for $d \in D$. The choice of D thus implies some prior belief about the whereabouts of d_0, for example to cover the unit root possibility $d_0 = 1$, D can only include nonstationary d-values, $d > \frac{1}{2}$. On the other hand there is no upper limit on $\overline{d}$. As seen in our proofs, all estimates can be seen as approximating panel data extensions of conditional-sum-of-squares (CSS) estimates, recently treated in a general fractionally integrated setting by Hualde and Robinson (2011), where D is effectively unrestricted. There may accordingly be scope for relaxing our restrictions on D, though these restrictions appear to play a role in ensuring that the approximation errors stemming from the presence of the individual effects α_i, or from the measures we take to eliminate them, are small enough to enable our estimates to be consistent and asymptotically normally distributed. All our estimates optimize objective functions that cross-sectionally aggregate time series objective functions.

2.1 Uncorrected Estimation

Our first approach is essentially CSS estimation which ignores the α_i. Define

$$L_T^U(d) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=0}^{T}\left(\Delta_{t+1}^{d} y_{it}\right)^2$$

and

$$\widehat{d}_T^U = \arg\min_{d\in D} L_T^U(d).$$

Introduce the notation

$$\tau_t(d) = \pi_t(d-1).$$

Notice that

$$\begin{aligned}\Delta_{t+1}^{d} y_{it} &= \Delta_{t+1}^{d-d_0}\varepsilon_{it} + \Delta_{t+1}^{d}\alpha_i \\ &= \Delta_{t+1}^{d-d_0}\varepsilon_{it} + \tau_t(d)\,\alpha_i,\end{aligned} \tag{8}$$

since

$$\Delta_{t+1}^{d} 1 = \sum_{j=0}^{t} \pi_j (d) = \tau_t (d), \tag{9}$$

as proved in Lemma 1. The term $\tau_t(d)\,\alpha_i$ in (8) contributes a bias; in view of (6) this term is $O_p\left(t^{-d}\right)$ and thus decays to zero for $d > 0$, but only slowly, and its presence explains the need for asymptotic theory with $T \to \infty$, in order to achieve consistent estimation of d_0.

2.2 Fixed Effects Estimation

Instead of ignoring the α_i we now start from a CSS-type objective function that is based on fractionally differencing the $y_{it} - \alpha_i$, and then concentrates out the α_i. Define

$$L_T(d, \alpha_1, \ldots, \alpha_N) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=0}^{T} \left(\Delta_{t+1}^{d} (y_{it} - \alpha_i)\right)^2.$$

Differentiating gives

$$\frac{\partial}{\partial \alpha_i} L_T(d, \alpha_1, \ldots, \alpha_N) = \frac{-2}{NT} \sum_{t=0}^{T} \left(\Delta_{t+1}^{d} y_{it} - \Delta_{t+1}^{d} \alpha_i\right)\left(\Delta_{t+1}^{d} 1\right), \quad i = 1, \ldots, N,$$

and thence

$$\hat{\alpha}_{iT}(d) = S_{\tau\tau T}(d)^{-1} \sum_{t=0}^{T} \left(\Delta_{t+1}^{d} y_{it}\right) \tau_t(d), \quad i = 1, \ldots, N,$$

using (9) and defining

$$\begin{aligned} S_{\tau\tau T}(d) &= 1 + \boldsymbol{\tau}_T'(d)\, \boldsymbol{\tau}_T(d), \\ \boldsymbol{\tau}_T(d) &= (\tau_1(d), \ldots, \tau_T(d))', \end{aligned}$$

the prime denoting transposition. Thence introduce

$$L_T^F(d) = L_T(d, \hat{\alpha}_{1T}(d), \ldots, \hat{\alpha}_{NT}(d)) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=0}^{T} \left(\Delta_{t+1}^{d} (y_{it} - \hat{\alpha}_{iT}(d))\right)^2,$$

and

$$\widehat{d_T^F} = \arg\min_{d \in D} L_T^F(d).$$

The summands in $L_T^F(d)$ are squared fractional residuals after regression on the fractional final end effect $\left(\Delta_{t+1}^{d} 1\right) = \tau_t(d)$.

Define

$$a_{iT}(d) = \sum_{t=0}^{T} \tau_t(d) \sum_{j=0}^{t} \pi_j(d - d_0)\, \varepsilon_{i,t-j}. \tag{10}$$

Then

$$\Delta_{t+1}^{d} (y_{it} - \hat{\alpha}_i(d)) = \Delta_{t+1}^{d-d_0} \varepsilon_{it} - \frac{a_{iT}(d) \tau_t(d)}{S_{\tau\tau T}(d)}, \tag{11}$$

and by comparison with $\widehat{d_T^U}$ there is again a term contributing bias. We show that nevertheless $\widehat{d_T^F}$ is consistent though a bias correction may be desirable for statistical inference.

2.3 Difference Estimation

A standard approach to eliminating the α_i from (1) is first-differencing:

$$\Delta y_{it} = \Delta_{t+1}^{1-d_0} \varepsilon_{it}, \quad t = 1, \ldots, T.$$

We might then attempt to fully whiten the data by taking $(d-1)\, th$ differences of the Δy_{it},

$$z_{it}(d) = \Delta_t^{d-1} (\Delta y_{it}).$$

Define

$$L_T^D(d) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} z_{iT}^2(d)$$

and

$$\widehat{d}_T^D = \arg\min_{d \in D} L_T^D(d).$$

Notice that

$$\begin{aligned} z_{it}(d) &= \Delta_t^{d-1} \left(\Delta_{t+1}^{1-d_0} \varepsilon_{it} \right) \\ &= \Delta_{t+1}^{d-1} \left(\Delta_{t+1}^{1-d_0} \varepsilon_{it} \right) + \left\{ \Delta_t^{d-1} - \Delta_{t+1}^{d-1} \right\} \left(\Delta_{t+1}^{1-d_0} \varepsilon_{it} \right) \\ &= \Delta_{t+1}^{d-d_0} \varepsilon_{it} - \tau_t(d) \varepsilon_{i0}, \end{aligned} \tag{12}$$

so there is a bias contribution similar to that in the uncorrected estimate $\widehat{d}_T^U$.

It may be worth noting that if a possibly nonparametric time trend of the form $f(t/T)$ is added to the left hand side of (1), where f is a Lipschitz-continuous function on $[0,1]$, the first-differencing also eliminates this, to order $O(T^{-1})$. Thus $\widehat{d}_T^D$ may enjoy some robustness to the unanticipated presence of such a trend, though we do not explore this possibility here.

2.4 Pseudo Maximum Likelihood Estimation

The previous estimates all employ versions of the CSS principal, where the Gaussian pseudo-likelihood is approximated by ignoring potential dependence and heteroscedasticity in the approximately whitened data. Here we modify the previous, Difference, estimate by employing a pseudo likelihood for the $z_{it}(d)$. From (12),

$$Cov\left(z_{is}(d_0), z_{it}(d_0)\right) = \omega_{st}(d_0)\, \sigma^2,$$

where

$$\omega_{st}(d) = 1(s = t) + \tau_s(d)\, \tau_t(d).$$

Note however that $Cov\left(z_{is}(d), z_{it}(d)\right)$ differs from $\omega_{st}(d)\, \sigma^2$ for $d \neq d_0$. Introduce the $T \times T$ matrix

$$\Omega_T(d) = (\omega_{st}(d)) = I_T + \boldsymbol{\tau}_T(d)\, \boldsymbol{\tau}_T'(d)$$

and the $T \times 1$ vectors $\mathbf{z}_{iT}(d) = (z_{i1}(d), \ldots, z_{iT}(d))'$, $i = 1, \ldots, N$. Define the Gaussian pseudo log-likelihood

$$Q_T\left(d, \sigma^2\right) = -\sum_{i=1}^{N} \left\{ \frac{T}{2} \log(2\pi) + \frac{T}{2} \log \sigma^2 + \frac{1}{2} \log |\Omega_T(d)| + \frac{1}{2\sigma^2} \mathbf{z}_{iT}'(d)\, \Omega_T(d)^{-1}\, \mathbf{z}_{iT}(d) \right\}.$$

Differentiating,

$$\frac{\partial}{\partial \sigma^2} Q_T\left(d, \sigma^2\right) = -\sum_{i=1}^{N}\left\{\frac{T}{2\sigma^2} - \frac{1}{2\sigma^4}\mathbf{z}'_{iT}(d)\,\Omega_T(d)^{-1}\mathbf{z}_{iT}(d)\right\},$$

leading to

$$\widehat{\sigma}^2_T(d) = \frac{1}{NT}\sum_{i=1}^{N}\mathbf{z}'_{iT}(d)\,\Omega_T(d)^{-1}\mathbf{z}_{iT}(d),$$

and the concentrated function

$$\begin{aligned} Q_T\left(d, \widehat{\sigma}^2_T(d)\right) &= -\left\{\frac{NT}{2}\log(2\pi) + \frac{NT}{2}\log\widehat{\sigma}^2_T(d) + \frac{N}{2}\log|\Omega_T(d)| + \frac{NT}{2}\right\} \\ &= -\frac{NT}{2}(1+\log(2\pi)) - \frac{NT}{2}\log\widehat{\sigma}^2_T(d) - \frac{N}{2}\log|\Omega_T(d)|. \end{aligned}$$

Thus define

$$\begin{aligned} L^P_T(d) &= \exp\left\{-\frac{2}{NT}Q^P_T\left(d, \widehat{\sigma}^2(d)\right) - (1+\log(2\pi))\right\} \\ &= |\Omega_T(d)|^{\frac{1}{T}}\widehat{\sigma}^2_T(d), \end{aligned}$$

and

$$\widehat{d}^P_T = \arg\min_{d\in D} L^P_T(d).$$

For computations, note the formulae

$$\Omega_T(d)^{-1} = I_T - \frac{\boldsymbol{\tau}_T(d)\,\boldsymbol{\tau}'_T(d)}{S_{\tau\tau T}(d)},\ \ |\Omega_T(d)| = S_{\tau\tau T}(d). \tag{13}$$

3. Consistency

Consistency proofs are facilitated by noting that all four of the objective functions introduced in the previous section are approximately equal, and are of the form

$$L_T(d) = A_T(d) + B_T(d),$$

where

$$A_T(d) = \frac{1}{N}\sum_{i=1}^{N}A_{iT}(d),\quad A_{iT}(d) = \frac{1}{T}\sum_{t=0}^{T}\left(\Delta^{d-d_0}_{t+1}\varepsilon_{it}\right)^2$$

and $B_T(d)$ is a measurable function of ε_{it}, $1 \le i \le N$, $0 \le t \le T$, of smaller order of magnitude. Hualde and Robinson (2011) showed under conditions on ε_{1t}, $t = 1, 2, \ldots$, that are implied by ours, that the statistic

$$\widetilde{d}^1_T = \arg\min_D A_{1T}(d)$$

is consistent for d_0. They were thus concerned with the single time series case, but due to the identity of distribution across i, and model constancy across i, their results easily extend to establish consistency of

$$\widetilde{d}_T = \arg\min_D A_T(d).$$

We state first the following Proposition which is used to prove consistency of each of our estimates, along with Theorem 1 of Hualde and Robinson (2011). Define

$$\widehat{d}_T = \arg\min_D L_T(d).$$

Proposition 1 *Let*

$$\sup_D |B_T(d)| \to_p 0, \text{ as } T \to \infty. \tag{14}$$

Then as $T \to \infty$

$$\widehat{d}_T \to_p d_0.$$

Theorem 3.1 *If* $\alpha_i = O_p(1)$ *uniformly in* i, *as* $T \to \infty$,

$$\widehat{d}_T^U \to_p d_0.$$

Theorem 3.2 *As* $T \to \infty$,

$$\widehat{d}_T^F \to_p d_0.$$

Theorem 3.3 *As* $T \to \infty$,

$$\widehat{d}_T^D \to_p d_0.$$

Theorem 3.4 *As* $T \to \infty$,

$$\widehat{d}_T^P \to_p d_0.$$

4. Asymptotic Normality

The following Proposition is not new when $N = 1$, but we include it to demonstrate that N may increase with T. Define

$$J_{t+1}(L) = -\sum_{j=1}^{t} j^{-1} L^j.$$

Proposition 2. *As* $T \to \infty$,

$$\frac{1}{(NT)^{\frac{1}{2}}} \sum_{i=1}^{N} \sum_{t=0}^{T} (J_{t+1}(L)\varepsilon_{it})\varepsilon_{it} \to_d \mathcal{N}\left(0, 6/\pi^2\right). \tag{15}$$

Theorem 4.1 *Let* $\alpha_i = O_p(1)$ *uniformly in* i *and* $d_0 \in Int(D)$. *When* $d_0 > \frac{1}{4}$, *as* $T \to \infty$,

$$(NT)^{\frac{1}{2}} \left(\widehat{d}_T^U - d_0\right) \to_d \mathcal{N}\left(0, 6/\pi^2\right) \tag{16}$$

if, as $T \to \infty$, $NT^{1-4d_0}\log^2 T \to 0$ *when* $d_0 \in \left(\frac{1}{4}, \frac{1}{2}\right)$, $NT^{-1}\log^4 T \to 0$ *when* $d_0 = \frac{1}{2}$, *and* $NT^{-1} \to 0$ *when* $d_0 > \frac{1}{2}$.

Note that Theorem 4.1, like Theorems 3.1-3.4, allows N to grow, but a slower rate than T, and arbitrarily slowly for d_0 close enough to $\frac{1}{4}$ from above, and no central limit theorem is available when $d_0 < \frac{1}{4}$.

Define $\dot{\pi}_t(d) = (\partial/\partial d)\pi_t(d)$, $\dot{\tau}_t(d) = \dot{\pi}_t(d-1)$ and

$$\begin{aligned} \dot{\boldsymbol{\tau}}_T(d) &= (\dot{\tau}_1(d), \ldots, \dot{\tau}_T(d))', \\ S_{\tau\dot{\tau}T}(d) &= \boldsymbol{\tau}_T'(d)\,\dot{\boldsymbol{\tau}}_T(d). \end{aligned}$$

Theorem 4.2 *Let $d_0 \in Int\,(D)$. Then as $T \to \infty$,*

$$(NT)^{\frac{1}{2}} \left(\widehat{d}_T^F - d_0 - T^{-1} b_T^F (d_0) \right) \to_d \mathcal{N} \left(0, 6/\pi^2\right), \tag{17}$$

where

$$b_T^F (d) = \frac{6 S_{\tau\dot{\tau}T}(d)}{\pi^2 S_{\tau\tau T}(d)},$$

with $b_T^F (d) = O\left(\log T 1(d \leq \frac{1}{2}) + 1(d > \frac{1}{2})\right)$. Thus

$$(NT)^{\frac{1}{2}} \left(\widehat{d}_T^F - d_0 \right) \to_d \mathcal{N} \left(0, 6/\pi^2\right) \tag{18}$$

if, as $T \to \infty$, $NT^{-1} \log^2 T \to 0$ when $d_0 \leq \frac{1}{2}$, and if $NT^{-1} \to 0$ otherwise.

When $d_0 > \frac{1}{2}$, the restrictions on N for (18) are the same as those for (16) for $\widehat{d}_T^U$ but when $d_0 \leq \frac{1}{2}$ they are weaker, and do not strengthen with decreasing d_0, indeed (18), unlike (16), holds for $d_0 \in (0, \frac{1}{4}]$. Moreover, whereas Theorem 4.1, like Theorem 3.1, imposes some restriction on the α_i, this is avoided in Theorem 4.2. The recentering in (17) avoids any restrictions on N. Note that $b_T^F (d)$ is a known function of d. For $d < 1$, $\pi_t(d-1) > 0$ for all $t > 0$, whence Lemma 4 in Appendix C implies that $S_{\tau\dot{\tau}T}(d) < 0$, and thus $b_T^F (d) < 0$.

Define the $T \times 1$ vector $\mathbf{m}_T = (1, \frac{1}{2}, \ldots, \frac{1}{T})'$, and

$$S_{\tau m T}(d) = \boldsymbol{\tau}'_T (d)\, \mathbf{m}_T.$$

Theorem 4.3 *Let $d_0 \in Int\,(D)$. Then as $T \to \infty$,*

$$(NT)^{\frac{1}{2}} \left(\widehat{d}_T^D - d_0 - T^{-1} b_T^D (d_0) \right) \to_d \mathcal{N} \left(0, 6/\pi^2\right), \tag{19}$$

where

$$b_T^D (d) = -\frac{6}{\pi^2} \left(S_{\tau\dot{\tau}T}(d) + S_{\tau m T}(d) \right)$$

and $b_T^D (d) = O(T^{1-2d} \log T 1(d < \frac{1}{2}) + \log^2 T 1(d = \frac{1}{2}) + 1(d > \frac{1}{2}))$. Thus, when $d_0 > \frac{1}{4}$,

$$(NT)^{\frac{1}{2}} \left(\widehat{d}_T^D - d_0 \right) \to_d \mathcal{N} \left(0, 6/\pi^2\right) \tag{20}$$

if, as $T \to \infty$, $NT^{1-4d_0} \log^2 T \to 0$ when $d_0 \in \left(\frac{1}{4}, \frac{1}{2}\right)$, $NT^{-1} \log^4 T \to 0$ when $d_0 = \frac{1}{2}$, and $NT^{-1} \to 0$ when $d_0 > \frac{1}{2}$.

The result (20) is the same as (16) for $\widehat{d}_T^U$, except that it imposes no restrictions on the α_i. As with (17) for $\widehat{d}_T^F$, (19) avoids any restrictions on N. The bias term $b_T^D (d)$ lacks the deflating factor $S_{\tau\tau T}^{-1}(d) \leq 1$ of $\widehat{d}_T^F$, making it of larger order of magnitude than $b_T^F (d)$ when $d_0 \leq \frac{1}{2}$, and it also involves the additional term $S_{\tau m T}(d)$. This is $O(1)$ for all d (see Lemma 1) and is thus dominated asymptotically by $S_{\tau\dot{\tau}T}(d)$ when $d_0 \leq \frac{1}{2}$. For $d < 1$, $\pi_t(d-1) > 0$ for all t, and thus $S_{\tau m T}(d) > 0$, and since $S_{\tau\dot{\tau}T}(d) < 0$ as previously observed, there is some cancelation in the bias. For $d > 1$, $\sum_{t=1}^{\infty} \pi_t(d-1) = -1$, and it is readily seen that $S_{\tau m T}(d) < 0$ for all large enough T.

Theorem 4.4 *Let $d_0 \in Int\,(D)$. Then as $T \to \infty$,*

$$(NT)^{\frac{1}{2}} \left(\widehat{d}_T^P - d_0 - T^{-1} b_T^P (d_0) \right) \to_d \mathcal{N} \left(0, 6/\pi^2\right), \tag{21}$$

where

$$b_T^P (d) = -\frac{6}{\pi^2} \frac{S_{\tau\dot{\tau}T}(d) + S_{\tau m T}(d)}{S_{\tau\tau T}(d)}$$

and $b_T^P(d) = O\left(\log T 1(d \leq \frac{1}{2}) + 1(d > \frac{1}{2})\right)$. *Thus,*

$$(NT)^{1/2}\left(\widehat{d}_T^P - d_0\right) \rightarrow_d \mathcal{N}\left(0, 6/\pi^2\right) \tag{22}$$

if, as $T \to \infty$, $NT^{-1}\log^2 T \to 0$ *when* $d_0 \leq \frac{1}{2}$, *and if* $NT^{-1} \to 0$ *otherwise.*

The result (22) is identical to (18) for $\widehat{d}_T^F$, while (21) differs from (17) for $\widehat{d}_T^F$ only in that $b_T^P(d) = b_T^F(d) - (6\pi^2)S_{\tau m T}(d)S_{\tau\tau T}^{-1}(d)$, where $S_{\tau m T}(d)S_{\tau\tau T}^{-1}(d) = O(1)$ at most, for all d. Comparing to (19) for $\widehat{d}_T^D$, we find that $b_T^P(d) = b_T^D(d)\, S_{\tau\tau T}^{-1}(d)$; for all d, $\left|b_T^P(d)\right| \leq \left|b_T^D(d)\right|$ since $S_{\tau\tau T}(d) \geq 1$.

5. Statistical Inference

In the present section we discuss and develop the results of the previous section for statistical inference on d_0. We focus on the estimates $\widehat{d}_{T,}^F$, $\widehat{d}_T^D$ and $\widehat{d}_T^P$, since Theorems 4.1-4.3 indicate potential for bias correction of these, and thereby relaxing the restrictions on the rate of increase of N relative to T, whereas (as the proof of Theorem 4.1 indicates) the leading term in the bias of $\widehat{d}_T^U$ depends on the α_i, as well as d_0. We should also bear in mind our discussion in the previous section which suggests that, on theoretical grounds, there is little to choose between the estimates when $d_0 > \frac{1}{2}$, whereas overall $\widehat{d}_T^F$ and $\widehat{d}_T^P$ dominate when $d_0 \leq \frac{1}{2}$ in respect of their entailing the weakest restrictions on N in central limit theorems centered at d_0, of their biases being of smaller order, and of admitting a central limit theorem when $d_0 \leq \frac{1}{4}$. We will assume the form $6/\pi^2$ appearing as asymptotic variance in the theorems of the previous section is used, but it is the limit of $\left(\sum_{t=1}^T t^{-2}\right)^{-1}$ and could be replaced by the latter, which might perform better in finite samples.

We first discuss Wald hypothesis testing on d_0. The leading case, mentioned in the Introduction, of testing the unit root null $d_0 = 1$, turns out to be the most favourable. Since $\tau_t(1) = 0$, $1 \leq t \leq T$, it follows that $b_T^F(1) = b_T^D(1) = b_T^P(1) = 0$. Thus the results (18), (20) and (22) are respectively identical to (17), (19), and (21), and so $(NT)^{1/2}\left(\hat{d}_T^F - 1\right)$, $(NT)^{1/2}\left(\widehat{d}_T^D - 1\right)$ and $(NT)^{1/2}\left(\widehat{d}_T^P - 1\right)$ are asymptotically $\mathcal{N}\left(0, 6/\pi^2\right)$ with no restrictions on N. Another case that is sometimes of interest is the $I(2)$ hypothesis $d_0 = 2$. It is easy to see that $S_{\tau\tau T}(2) = S_{\tau\dot{\tau} T}(2) = 1$, $S_{\tau m T}(2) = -1$, so $b_T^F(2) = -6/\pi^2$, $b_T^D(2) = b_T^P(2) = 0$, and $\widehat{d}_T^F$ is simply bias-corrected, while no correction of $\widehat{d}_T^D$ or $\widehat{d}_T^P$ is needed. In general, for other null hypotheses, for example $d_0 = \frac{1}{2}$ (the boundary between the stationary and nonstationary regions), we can carry out the bias correction by evaluating $b_T^F(d_0)$, $b_T^D(d_0)$ and $b_T^P(d_0)$ at the null, which is straightforward given Lemma 4, and applying (17), (19), and (21).

Some numerical comparisons of the biases are of interest. Tables 1 and 2 present the scaled biases of $\widehat{d}_T^F$ and $\widehat{d}_T^D$ for selected values of T and d. We find that $b_T^F(d)$ decreases monotonically in d and in T, sharing the sign of d, whereas $b_T^D(d)$ is positive and increasing in $|d-1|$ (though not symmetrically) and is mostly decreasing in T (note that scaling with respect to T has already been carried out). Table 3 presents the ratio of biases of $\widehat{d}_T^P$ and $\widehat{d}_T^D$, namely the quantity $S_{\tau\tau T}^{-1}(d)$; it decreases in both $|d-1|$ and T.

For interval estimation $b_T^F(d_0)$, $b_T^D(d_0)$ and $b_T^P(d_0)$ need to be estimated, and it is natural to consider $b_T^F\left(\widehat{d}_T^F\right)$, $b_T^D\left(\widehat{d}_T^D\right)$ and $b_T^P\left(\widehat{d}_T^P\right)$. Thus we define

$$\begin{aligned}\widetilde{d}_T^F &= \widehat{d}_T^F - T^{-1}b_T^F\left(\widehat{d}_T^F\right),\\ \widetilde{d}_T^D &= \widehat{d}_T^D - T^{-1}b_T^D\left(\widehat{d}_T^D\right),\\ \widetilde{d}_T^P &= \widehat{d}_T^P - T^{-1}b_T^P\left(\widehat{d}_T^P\right).\end{aligned}$$

The following theorems indicate that these feasibly mean-corrected estimates entail stronger restrictions on N (and in some cases on d_0) than the infeasibly mean-corrected ones, but milder restrictions than the uncorrected ones. In particular, $\widetilde{d}_T^D$ and $\widetilde{d}_T^P$ require $d_0 > \frac{1}{8}$ and all estimates require N to increase slower than T^3, though as with $\widehat{d}_T^U$, the rates for $\widetilde{d}_T^D$ and $\widetilde{d}_T^P$ the rate are heavily d_0-dependent, such that N must increase slower than T when $d_0 = \frac{1}{4}$ and arbitrarily slowly as d_0 approaches $\frac{1}{8}$ from above.

We give a single theorem to cover $\widetilde{d}_T^F$ and $\widetilde{d}_T^P$ because the regularity conditions in Theorems 4.1 and 4.3 are identical, only the bias differs.

Theorem 5.1 *Let* $d_0 \in Int\,(D)$. *Then as* $T \to \infty$,

$$(NT)^{\frac{1}{2}}\left(\widetilde{d}_T^j - d_0\right) \to_d \mathcal{N}\left(0, 6/\pi^2\right), \quad j \in \{F, P\}$$

if, as $T \to \infty$, $NT^{-3}\log^6 T \to 0$ *when* $d_0 \le \frac{1}{2}$ *or* $NT^{-3} \to 0$ *when* $d_0 > \frac{1}{2}$.

Theorem 5.2 *Let* $d_0 \in Int\,(D)$. *Then when* $d_0 > \frac{1}{8}$, *as* $T \to \infty$,

$$(NT)^{\frac{1}{2}}\left(\widetilde{d}_T^D - d_0\right) \to_d \mathcal{N}\left(0, 6/\pi^2\right),$$

if $NT^{1-8d_0}\log^6 T \to 0$ *when* $d_0 \in \left(\frac{1}{8}, \frac{1}{2}\right)$, *or if* $NT^{-3}\log^{10} T \to 0$ *when* $d_0 = \frac{1}{2}$, *or if* $NT^{-3} \to 0$ *when* $d_0 > \frac{1}{2}$.

It is of some interest to note that simplified corrections are possible that improve on our original F, D and E estimates, but by less than our feasible bias-corrected ones. From Lemma 2,

$$\begin{aligned}b_T^F(d) &= -\frac{6\log T}{\pi^2} + O(1),\ d < \frac{1}{2},\\ &= -\frac{3\log T}{\pi^2} + O(1),\ d = \frac{1}{2},\\ &= \frac{6}{\pi^2}\left(\psi(2d) - \psi(d) - \frac{1}{2d-1}\right) + O(T^{1-2d}\log T),\ d > \frac{1}{2};\end{aligned}$$

$$\begin{aligned}b_T^D(d) &= \frac{6T^{1-2d}\log T}{\pi^2(1-2d)\Gamma(1-d)^2} + O(T^{1-2d}),\ d < \frac{1}{2},\\ &= \frac{3\log^2 T}{2\pi^3} + O(\log T),\ d = \frac{1}{2},\\ &= -\frac{6}{\pi^2}\left\{\frac{\psi(2d) - \psi(d) - \frac{1}{2d-1}}{(2d-1)B(d,d)}\right.\\ &\quad \left. + \int_0^1 \left(\frac{(1-x)^{d-1} - 1}{x}\right)dx\right\} + O(T^{1-2d}\log T + T^{-d}),\ d > \frac{1}{2};\end{aligned}$$

$$
\begin{aligned}
b_T^P(d) &= \frac{6\log T}{\pi^2} + O(1),\ d < \frac{1}{2}, \\
&= \frac{3\log T}{2\pi^2} + O(1),\ d = \frac{1}{2}, \\
&= -\frac{6}{\pi^2}\left\{\psi(2d) - \psi(d) - \frac{1}{2d-1}\right. \\
&\quad \left. + (2d-1)\,B(d,d)\int_0^1 \left(\frac{(1-x)^{d-1} - 1}{x}\right) dx\right\} + O(T^{1-2d}\log T + T^{-d}),\ d > \frac{1}{2}.
\end{aligned}
$$

The leading terms here could be used in simpler bias corrections. For example a simple bias-corrected Fixed effects estimate is $\widehat{d}_T^F - 6\log T/\left(\pi^2 T\right)$ for $d_0 < \frac{1}{2}$, where the correction is free of $\widehat{d}_T^F$. But which correction to use requires knowledge of whether or not we are in the stationary region, and the theoretical improvements over the original bias-uncorrected estimates are small, noting the approximation errors above and bearing in and that the effect of inserting estimates of d_0 in most of the corrections needs to be taken into account. Tables 4 and 5 illustrate the approximation for the bias of the Fixed effects and Differenced estimates, and are directly comparable with those of Tables 1 and 2, respectively. The approximations work reasonably well when d_0 is close to 1, but otherwise are less precise.

6. Simulations

In this section we conduct a simulation study of the finite sample properties of our estimates of d_0. We concentrate on the Fixed Effects, Difference and the PML estimates, in both original and feasible bias-corrected forms. We do not report results for $\widehat{d}_T^U$, since this heavily depends on the magnitude of the fixed effects α_i relative to the idiosyncratic errors ε_{it}, whereas the others are invariant to the specification of α_i. We generate the ε_{it} as standard normal, noting that the estimates are invariant to the variance of ε_{it}. We focus on different specifications of N, T and d_0. In particular we set $T = 5, 10$ and 100 as in Tables 1-5, and to consider the effect of increasing the overall sample size, we used when $T = 5, 10$ three combinations of NT (100, 200 and 400) so the range of values of N oscillates from $N = 20$ to 80 for $T = 5$ and from $N = 10$ to 40 for $T = 10$, while when $T = 100$ we took only $NT =$ 200 and 400, i.e. $N = 2$ and 4 (thus omitting the case $NT = T = 100$ since we cannot remove fixed effects with a single time series). The values of d_0 include a stationary one ($d_0 = 0.3$), which is the most problematic from the point of view of bias, a moderately non-stationary one ($d_0 = 0.6$), values around the unit root ($d_0 =$ 0.9, 1.0, 1.1), and a more nonstationary one ($d_0 = 1.4$). Optimizations were carried out using the Matlab function fminbnd with $D = [0.1, 1.5]$, and the results are based on 10,000 independent replications.

We first explore the accuracy of the asymptotic approximations for the biases in Theorems 4.2-4.4, and whether feasible bias correction produces better centering properties. In Table 6 we observe that the uncorrected Fixed Effects estimate $\widehat{d}_T^F$ has a bias in line with that predicted in Table 1 when $d_0 = 0.3$ and $T = 5$, but in general it has larger bias (in absolute value) than predicted by the magnitude of $b_T^F(d_0)/T$ for large T and small d_0. For $d_0 \geq 1.0$ the bias is small, as predicted, and the accuracy of the approximation improves with increasing N. The right panel of Table 6 shows that feasible bias correction removes a large fraction of the bias of $\widehat{d}_T^F$ when $d_0 = 0.3$, but for all the smallish d_0 the biases, while reduced, are still substantial. In some cases the biases of $\widehat{d}_T^F$ and $\widetilde{d}_T^F$ do not change monotonically with d_0 and T. For the Difference estimate $\widehat{d}_T^D$ we observe that Table 7 shows more the monotonic properties of $b_T^D(d_0)/T$ found in Table 2, even for the smaller NT, and that bias correction works in $\widetilde{d}_T^D$ quite well when $d_0 \geq 0.6$. Table 8 illustrates the far superior bias properties of the uncorrected PML estimate $\widehat{d}_T^P$, which, surprisingly, are much better than those of Table 3 in comparison with $\widehat{d}_T^D$, and much better than those of the previous bias-corrected estimates, with bias correction in this case actually worsening finite sample properties.

Tables 9-11 report (scaled) Monte Carlo square error across simulations for the three estimates in both uncorrected and feasible bias-corrected versions. For all estimates, performance improves with increasing d_0, T and NT, predominately monotonically, and with bias correction (in this last respect with the exception of $\widetilde{d}_T^P$ and low values of d_0). The asymptotic standard error, $\left(6/\pi^2\right)/(NT)$, which gives 0.61, 0.30 and 0.15 for $NT = 100$, 200 and 400, respectively, are poorly approximated for low d_0, but in a number of cases quite well approximated for larger d_0.

Tables 12-14 report empirical coverage of 95% confidence intervals for d_0 based on our central limit theorems. The uncorrected $\widehat{d}_T^P$ estimate achieves much the most accurate coverage, although the results leave something to be desired when $d_0 = 0.3$ and 0.6, but the bias-corrected $\widetilde{d}_T^F$ and $\widetilde{d}_T^D$ also generally perform reasonably, at least for the larger d_0, especially by comparison with intervals based on uncorrected estimates.

Appendix A: Proofs of Theorems

Proof of Theorem 3.1 .

From (8),

$$\begin{aligned}
L_T^U(d) &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=0}^{T}\left(\Delta_{t+1}^{d-d_0}\varepsilon_{it} + \tau_t(d)\,\alpha_i\right)^2 \\
&= L_T(d) + \frac{2}{NT}\sum_{i=1}^{N}\alpha_i a_{iT}(d) + \frac{S_{\tau\tau T}(d)}{NT}\sum_{i=1}^{N}\alpha_i^2 .
\end{aligned}$$

We check Proposition 1. From (7) ,

$$\zeta = \frac{1}{2} - (d_0 - \underline{d}) > 0. \tag{23}$$

Thus, using Lemma 3,

$$\begin{aligned}
\sup_D\left|\frac{2}{NT}\sum_{i=1}^{N}\alpha_i a_{iT}(d)\right| &= O_p\left(\max_i\ |\alpha_i|\sup_D T^{\max(d_0-d,0)+\max(\frac{1}{2}-d,0)-1}\log^2 T\right) \\
&= O_p\left(T^{-\zeta-\underline{d}}\log^2 T\right) = o_p(1),
\end{aligned} \tag{24}$$

choosing $\underline{d} < 1/2$. From Lemma 2

$$S_{\tau\tau T}(d) \leq KT^{2\max(\frac{1}{2}-d,0)}\log T. \tag{25}$$

Thus

$$\sup_D\frac{S_{\tau\tau T}(d)}{NT}\sum_{i=1}^{N}\alpha_i^2 = O_p\left(\max_i\ \alpha_i^2\sup_D T^{-2\underline{d}}\log T\right) = o_p(1), \tag{26}$$

to complete verification of (14), and thus the proof. □

Proof of Theorem 3.2

From (11) and (10),

$$\begin{aligned} L_T^F(d) &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=0}^{T}\left(\Delta_{t+1}^{d-d_0}\varepsilon_{it} - \frac{a_{iT}(d)}{S_{\tau\tau T}(d)}\right)^2 \\ &= L_T(d) - \frac{1}{S_{\tau\tau T}(d)NT}\sum_{i=1}^{N} a_{iT}^2(d), \end{aligned}$$

We again check Proposition 1. From Lemma 2,

$$S_{\tau\tau T}(d) \geq \eta T^{2\max\left(\frac{1}{2}-d,0\right)}.$$

for some $\eta > 0$. Thus from Lemma 3

$$\begin{aligned} \sup_D \frac{1}{S_{\tau\tau T}(d)NT}\sum_{i=1}^{N} a_i^2(d) &= O_p\left(\sup_D T^{2\max(d_0-d,0)-1}\log^4 T\right) \\ &= O_p\left(T^{-2\zeta}\log^4 T\right) \to_p 0. \qquad \square \end{aligned}$$

Proof of Theorem 3.3

From (12),

$$\begin{aligned} L_T^D(d) &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\Delta_{t+1}^{d-d_0}\varepsilon_{it} - \tau_t(d)\,\varepsilon_{i0}\right)^2 \\ &= L_T(d) - \frac{1}{NT}\sum_{i=1}^{N}\varepsilon_{i0}^2 - \frac{2}{N}\sum_{i=1}^{N}\varepsilon_{i0}\frac{1}{T}\sum_{t=1}^{T}\tau_t(d)\,\Delta_{t+1}^{d-d_0}\varepsilon_{it} \\ &\quad + \frac{S_{\tau\tau T}(d)-1}{NT}\sum_{i=1}^{N}\varepsilon_{i0}^2 \\ &= L_T(d) - \frac{2}{NT}\sum_{i=1}^{N}\varepsilon_{i0}a_{iT}(d) + \frac{S_{\tau\tau T}(d)-1}{NT}\sum_{i=1}^{N}\varepsilon_{i0}^2 + O_p\left(T^{-1}\right) \end{aligned}$$

uniformly in d. Then as in (24) and (25)

$$\begin{aligned} \sup_D\left|\frac{2}{NT}\sum_{i=1}^{N}\varepsilon_{i0}a_{iT}(d)\right| &= o_p(1), \\ \sup_D\frac{S_{\tau\tau T}(d)-1}{NT}\sum_{i=1}^{N}\varepsilon_{i0}^2 &= o_p(1), \end{aligned}$$

to check Proposition 1. $\square$

Proof of Theorem 3.4

We have

$$\begin{aligned} L_T^P(d) &= \widehat{\sigma}_T^2(d) + |\left(\Omega_T(d)\,|^{\frac{1}{T}} - 1\right)\widehat{\sigma}_T^2(d) \\ &= L_T^D(d) + \left\{\widehat{\sigma}_T^2(d) - L^D(d)\right\} + \left(|\Omega_T(d)|^{\frac{1}{T}} - 1\right)\widehat{\sigma}_T^2(d). \end{aligned}$$

In view of Theorem 3.3, checking Proposition 1 entails verifying that

$$\sup_D \left| L_T^D(d) - \widehat{\sigma}_T^2(d) \right| = o_p(1), \tag{27}$$

$$\sup_D \left| |\Omega_T(d)|^{\frac{1}{T}} - 1 \right| = o(1), \tag{28}$$

since (27), and $\sup_D L_T^D(d) = O_p(1)$ from Theorem 1 of Hualde and Robinson (2011), imply that $\sup_D \widehat{\sigma}_T^2(d) = O_p(1)$. From (13),

$$\begin{aligned} L_T^D(d) - \widehat{\sigma}^2(d) &= \frac{1}{NT} \sum_{i=1}^{N} \mathbf{z}_{iT}'(d) \left\{ I_T - \Omega_T(d)^{-1} \right\} \mathbf{z}_{iT}(d) \\ &= \frac{1}{NTS_{\tau\tau T}(d)} \sum_{i=1}^{N} \left\{ \mathbf{z}_{iT}'(d)\, \boldsymbol{\tau}_T(d) \right\}^2. \end{aligned} \tag{29}$$

Now

$$\begin{aligned} \mathbf{z}_{iT}'(d)\, \boldsymbol{\tau}_T(d) &= \sum_{t=1}^{T} \tau_t(d) \left(\Delta_{t+1}^{d-d_0} \varepsilon_{it} - \tau_t(d)\, \varepsilon_{i0} \right) \\ &= a_{iT}(d) - \varepsilon_{i0} S_{\tau\tau T}(d). \end{aligned}$$

Lemmas 2 and 3 imply that this is uniformly

$$O_p\left(T^{\max(\frac{1}{2}-d,0)} \left(T^{\max(d_0-d,0)} + T^{\max(\frac{1}{2}-d,0)} \right) \log^2 T \right).$$

so (29) is

$$\begin{aligned} & O_p\left(\frac{T^{2\max(\frac{1}{2}-d,0)} \left(T^{2\max(d_0-d,0)} + T^{2\max(\frac{1}{2}-d,0)} \right) \log^4 T}{T^{1+2\max(\frac{1}{2}-d,0)}} \right) \\ = \; & O_p\left(\left(T^{2\max(d_0-d,0)} + T^{2\max(\frac{1}{2}-d,0)} \right) T^{-1} \log^4 T \right) \\ = \; & O_p\left(\left(T^{-2\zeta} + T^{-2\underline{d}} \right) \log^4 T \right) = o_p(1) \end{aligned}$$

uniformly, to check (27). Finally, from (13), with K denoting a generic finite constant,

$$\begin{aligned} |\Omega_T(d)|^{\frac{1}{T}} - 1 &= S_{\tau\tau T}(d)^{\frac{1}{T}} - 1 \\ &= O\left(\frac{S_{\tau\tau T}(d) - 1}{T} \right) \\ &\leq KT^{2\max(-d,0)-1} \log T \\ &\leq KT^{-2\underline{d}} \log T \to 0, \end{aligned}$$

uniformly, to check (28). □

Proof of Theorem 4.1

By the mean value theorem

$$0 = \frac{\partial}{\partial d} L_T^U\left(\widehat{d}_T^U \right) = \frac{\partial}{\partial d} L_T^U(d_0) + \frac{\partial^2}{\partial d^2} L_T^U(d^*) \left(\widehat{d}_T^U - d_0 \right), \tag{30}$$

where $|d^* - d_0| \leq \left|\widehat{d}_T^U - d_0\right|$. We have

$$
\begin{aligned}
\frac{\partial}{\partial d} L_T^U(d) &= \frac{2}{NT} \sum_{i=1}^N \sum_{t=0}^T \left(J_{t+1}(L)\, \Delta_{t+1}^d y_{it}\right) \Delta_{t+1}^d y_{it} \\
&= \frac{2}{NT} \sum_{i=1}^N \sum_{t=0}^T \left(J_{t+1}(L)\, \Delta_{t+1}^d \alpha_i + J_{t+1}(L)\, \Delta_{t+1}^{d-d_0} \varepsilon_{it}\right) \left(\Delta_{t+1}^d \alpha_i + \Delta_{t+1}^{d-d_0} \varepsilon_{it}\right).
\end{aligned}
$$

Thus

$$
\begin{aligned}
(NT)^{\frac{1}{2}} \frac{\partial}{\partial d} L_T^U(d_0) &= \frac{2}{(NT)^{\frac{1}{2}}} \sum_{i=1}^N \sum_{t=0}^T \left(J_{t+1}(L)\, \Delta_{t+1}^{d_0} \alpha_i + J_{t+1}(L)\, \varepsilon_{it}\right) \left(\Delta_{t+1}^{d_0} \alpha_i + \varepsilon_{it}\right) \\
&= 2w_T + \frac{2}{(NT)^{\frac{1}{2}}} \sum_{i=1}^N \sum_{t=0}^T \left(J_{t+1}(L)\, \Delta_{t+1}^{d_0} \alpha_i\right) \varepsilon_{it} \\
&\quad + \frac{2}{(NT)^{\frac{1}{2}}} \sum_{i=1}^N \sum_{t=0}^T \left(J_{t+1}(L)\, \varepsilon_{it}\right) \left(\Delta_{t+1}^{d_0} \alpha_i\right) \\
&\quad + \frac{2}{(NT)^{\frac{1}{2}}} \sum_{i=1}^N \alpha_i^2 \sum_{t=0}^T \left(J_{t+1}(L)\, \Delta_{t+1}^{d_0} 1\right) \tau_t^0, \qquad (31)
\end{aligned}
$$

writing $\tau_t^0 = \tau(d_0)$ and w_T for the left side of (15). By Proposition 2, $w_T \to_d \mathcal{N}(0, \sigma^4 \pi^2/6)$. The next two terms are

$$
\begin{aligned}
& \frac{2}{(NT)^{\frac{1}{2}}} \sum_{i=1}^N \alpha_i \left\{ \sum_{t=0}^T \varepsilon_{it} \left(J_{t+1}(L)\, \tau_t^0\right) + \tau_t^0 \left(J_{t+1}(L)\, \varepsilon_{it}\right) \right\} \\
= & -\frac{2}{(NT)^{\frac{1}{2}}} \sum_{i=1}^N \alpha_i \left\{ \sum_{t=1}^T \varepsilon_{it} \sum_{j=1}^t \frac{\tau_{t-j}^0}{j} + \sum_{t=0}^{T-1} \varepsilon_{it} \sum_{j=1}^{T-t} \frac{\tau_{t+j}^0}{j} \right\} \\
= & \left(\frac{N}{T}\right)^{\frac{1}{2}} O_p\left(\left\{ \sum_{t=1}^T \left(\sum_{j=1}^t \frac{\tau_{t-j}^0}{j} \right)^2 + \sum_{t=0}^{T-1} \left(\sum_{j=1}^{T-t} \frac{\tau_{t+j}^0}{j} \right)^2 \right\}^{\frac{1}{2}} \right) \\
= & \left(\frac{N}{T}\right)^{\frac{1}{2}} O_p\left(\left\{ \sum_{t=1}^T \left(\sum_{j=1}^t \frac{(t-j)^{-d_0}}{j} \right)^2 + \sum_{t=0}^{T-1} \left(\sum_{j=1}^{T-t} \frac{(t+j)^{-d_0}}{j} \right)^2 \right\}^{\frac{1}{2}} \right). \qquad (32)
\end{aligned}
$$

Now for $d > 0$,

$$
\begin{aligned}
\sum_{j=1}^t \frac{(t-j)^{-d}}{j} &= O\left(t^{-d} \log t + t^{-1}\right),\ t > 0, \\
\sum_{j=1}^{T-t} \frac{(t+j)^{-d}}{j} &= O((t+1)^{-d} \log (t+1)),\ t \geq 0,
\end{aligned}
$$

so (32) is

$$\left(\frac{N}{T}\right)^{\frac{1}{2}} O_p\left(\left\{\sum_{t=1}^{T}\left(t^{-2d_0}\log^2 t + t^{-2}\right)\right\}^{\frac{1}{2}}\right)$$
$$= \left(\frac{N}{T}\right)^{\frac{1}{2}} O_p\left(T^{\frac{1}{2}-d_0}\log T 1(d_0 < \frac{1}{2}\right) + \log^{\frac{3}{2}} T 1(d_0 = \frac{1}{2}) + 1(d_0 > \frac{1}{2})).$$

The final term in (31) is

$$\frac{2}{(NT)^{\frac{1}{2}}}\sum_{i=1}^{N}\alpha_i^2\sum_{t=0}^{T}\left(J_{t+1}(L)\Delta_{t+1}^{d_0}1\right)\tau_t^0 = O_p\left(\left(\frac{N}{T}\right)^{\frac{1}{2}}\sum_{t=1}^{T}\left|\tau_t^0\right|\sum_{j=1}^{t}\frac{\left|\tau_{t-j}^0\right|}{j}\right)$$
$$= O_p\left(\left(\frac{N}{T}\right)^{\frac{1}{2}}\sum_{t=1}^{T}\left(t^{-2d_0}\log t + t^{-d_0-1}\right)\right),$$

which is

$$\left(\frac{N}{T}\right)^{\frac{1}{2}} O_p\left(T^{1-2d_0}\log T 1(d_0 < \frac{1}{2}) + \log^2 T 1(d_0 = \frac{1}{2}) + 1(d_0 > \frac{1}{2})\right).$$

Thus the last three terms in (31) are $o_p(1)$ under the stated conditions on N and T.

Next we have

$$\frac{\partial^2}{\partial d^2}L_T^U(d) = \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=0}^{T}\left(J_{t+1}(L)^2\Delta_{t+1}^{d}y_{it}\right)\Delta_{t+1}^{d}y_{it} + \left(J_{t+1}(L)\Delta_{t+1}^{d}y_{it}\right)\left(J_{t+1}(L)\Delta_{t+1}^{d}y_{it}\right)$$
$$= \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=0}^{T}\left\{\left(J_{t+1}(L)^2\Delta_{t+1}^{d}\alpha_i + J_{t+1}(L)^2\Delta_{t+1}^{d-d_0}\varepsilon_{it}\right)\left(\Delta_{t+1}^{d}\alpha_i + \Delta_{t+1}^{d-d_0}\varepsilon_{it}\right)\right.$$
$$\left. + \left(J_{t+1}(L)\Delta_{t+1}^{d}\alpha_i + J_{t+1}(L)\Delta_{t+1}^{d-d_0}\varepsilon_{it}\right)\left(J_{t+1}(L)\Delta_{t+1}^{d}\alpha_i + J_{t+1}(L)\Delta_{t+1}^{d-d_0}\varepsilon_{it}\right)\right\}.$$

Thus

$$\frac{\partial^2}{\partial d^2}L_T^U(d_0) = \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=0}^{T}\left(J_{t+1}(L)^2\Delta_{t+1}^{d_0}\alpha_i + J_{t+1}(L)^2\varepsilon_{it}\right)\left(\Delta_{t+1}^{d_0}\alpha_i + \varepsilon_{it}\right)$$
$$+ \left(J_{t+1}(L)\Delta_{t+1}^{d_0}\alpha_i + J_{t+1}(L)\varepsilon_{it}\right)\left(J_{t+1}(L)\Delta_{t+1}^{d_0}\alpha_i + J_{t+1}(L)\varepsilon_{it}\right).$$

By arguments similar to those previously used this differs by $o_p(1)$ from

$$\frac{2}{T}\sum_{t=0}^{T}\left\{E\left(J_{t+1}(L)\varepsilon_{it}\right)^2 + O_p\left(\left|\left(J_{t+1}(L)^2\tau_t^0\right)\tau_t^0\right| + \left(J_{t+1}(L)\tau_t^0\right)^2\right)\right\}$$
$$= \frac{2}{T}\sum_{t=0}^{T}\left\{\sigma^2\sum_{j=1}^{t}j^{-2} + O_p\left(t^{-2d_0}\log^2 t\right)\right\} \to 2\sigma^2\frac{\pi^2}{6}, \text{ as } T\to\infty,$$

while, in view also of Theorem 3.1 and the proof of Theorem 2 of Hualde and Robinson (2011), $\partial^2 L_T^U(d^*)/\partial d^2 - \partial^2 L_T^U(d_0)/\partial d^2 \to_p 0$. The proof is completed. □

Proof of Theorem 4.2 . By the mean value theorem

$$0 = \frac{\partial}{\partial d}L_T^F\left(\widehat{d}_T^F\right) = E\left\{\frac{\partial}{\partial d}L_T^F(d_0)\right\} + \left[\frac{\partial}{\partial d}L_T^F(d_0) - E\left\{\frac{\partial}{\partial d}L_T^F(d_0)\right\}\right]$$
$$+\frac{\partial^2}{\partial d^2}L_T^F(d^*)\left(\widehat{d}_F - d_0\right). \quad (33)$$

Now

$$\frac{\partial}{\partial d}L_T^F(d) = \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=0}^{T}\left(J_{t+1}(L)\,\Delta_{t+1}^{d-d_0}\varepsilon_{it} - \frac{\dot\tau_t(d)a_{iT}(d)}{S_{\tau\tau T}(d)}\right)\tilde\varepsilon_{it}(d)$$

using the orthogonality, across $t = 0, 1, \ldots, T$, of $\tilde\varepsilon_{it}(d) = \Delta_{t+1}^{d}(y_{it} - \hat\alpha_{iT}(d))$ to $\tau_t(d)$. Thus, writing $\dot\tau_t^0 = \dot\tau_t(d_0)$, $S_{\tau\tau T}^0 = S_{\tau\tau T}(d_0)$, $a_{iT}^0 = a_{iT}(d_0)$,

$$\begin{aligned}(NT)^{\frac{1}{2}}\frac{\partial}{\partial d}L_T^F(d_0) &= 2w_T - \frac{2(NT)^{-\frac{1}{2}}}{S_{\tau\tau T}^0}\sum_{i=1}^{N}\sum_{t=0}^{T}\dot\tau_t^0 a_{iT}^0\varepsilon_{it}\\ &\quad - \frac{2(NT)^{-\frac{1}{2}}}{S_{\tau\tau T}^0}\sum_{i=1}^{N}a_{iT}^0\sum_{t=0}^{T}(J_{t+1}(L)\varepsilon_{it})\tau_t^0\\ &\quad + \frac{2(NT)^{-\frac{1}{2}}}{S_{\tau\tau T}^{02}}\sum_{i=1}^{N}\sum_{t=0}^{T}\tau_t^0\dot\tau_t^0 a_{iT}^{02}, \qquad (34)\end{aligned}$$

since $\tilde\varepsilon_{it}(d_0) = \varepsilon_{it} - \tau_t^0\hat\alpha_{iT}(d_0)$. In view of Proposition 2 and (33) we need to show that the remaining three terms on the right of (34) differ by $o_p(1)$ from their expectations, which we also need to evaluate. Since $a_{iT}^0 = \sum_{t=0}^{T}\tau_t^0\varepsilon_{it}$,

$$-\frac{2(NT)^{-\frac{1}{2}}}{S_{\tau\tau T}^0}\sum_{i=1}^{N}\sum_{t=1}^{T}\dot\tau_t^0 a_{iT}^0\varepsilon_{it} \qquad (35)$$

has expectation

$$-\frac{2\sigma^2(NT)^{-\frac{1}{2}}}{S_{\tau\tau T}^0}\sum_{i=1}^{N}\sum_{t=1}^{T}\tau_t^0\dot\tau_t^0 = -2\sigma^2\left(\frac{N}{T}\right)^{\frac{1}{2}}\frac{S_{\tau\dot\tau T}^0}{S_{\tau\tau T}^0}.$$

The variance of (35) is

$$\begin{aligned}&\frac{4}{NT}S_{\tau\tau T}^{0-2}\sum_{t=1}^{T}\sum_{u=1}^{T}\dot\tau_t^0\dot\tau_u^0\sum_{r=0}^{T}\sum_{s=0}^{T}\tau_r^0\tau_s^0\sum_{i=1}^{N}E(\varepsilon_{ir}\varepsilon_{is}\varepsilon_{it}\varepsilon_{iu} - E(\varepsilon_{is}\varepsilon_{it})E(\varepsilon_{ir}\varepsilon_{iu}))\\ =\;&\frac{4\sigma^4}{T}S_{\tau\tau T}^{0-2}\sum_{t=1}^{T}\dot\tau_t^{02}\sum_{s=0}^{T}\tau_s^{02} + \frac{4\sigma^4}{T}S_{\tau\tau T}^{0-2}\left(\sum_{t=1}^{T}\dot\tau_t^0\tau_t^0\right)^2 + O(T^{-1})\\ =\;&\frac{4\sigma^4}{T}S_{\tau\tau T}^{0-1}S_{\dot\tau\dot\tau T}^0 + \frac{4\sigma^4}{T}S_{\tau\tau T}^{0-2}S_{\tau\dot\tau T}^{02} + O(T^{-1}) = o(1)\end{aligned}$$

as $T \to \infty$.

Next consider the term

$$-\frac{2(NT)^{-\frac{1}{2}}}{S_{\tau\tau T}^0}\sum_{i=1}^{N}a_{iT}^0\sum_{t=0}^{T}(J_{t+1}(L)\varepsilon_{it})\tau_t^0 = \frac{2(NT)^{-\frac{1}{2}}}{S_{20,T}(d_0)}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{t}\frac{\varepsilon_{i,t-j}}{j}\tau_t^0\sum_{s=0}^{T}\tau_s^0\varepsilon_{is}.$$

This has expectation

$$\begin{aligned}-2\left(\frac{N}{T}\right)^{\frac{1}{2}}S_{\tau\tau T}^{0-1}\sum_{t=1}^{T}\sum_{s=0}^{T}\sum_{j=1}^{t}j^{-1}E(\varepsilon_{i,t-j}\varepsilon_{is})\tau_t^0\tau_s^0 &= 2\sigma^2\left(\frac{N}{T}\right)^{\frac{1}{2}}S_{\tau\tau T}^{0-1}\sum_{t=1}^{T}\tau_t^0\sum_{s=0}^{t-1}(t-s)^{-1}\tau_s^0\\ &= -2\sigma^2\left(\frac{N}{T}\right)^{\frac{1}{2}}S_{\tau\tau T}^{0-1}\sum_{t=1}^{T}\tau_t^0\dot\tau_t^0\\ &= -2\sigma^2\left(\frac{N}{T}\right)^{\frac{1}{2}}\frac{S_{\tau\dot\tau T}^{02}}{S_{\tau\tau T}^0},\end{aligned}$$

by Lemma 4. Its variance is

$$
\begin{aligned}
& \frac{4}{NT} S_{\tau\tau T}^{0-1} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{j=1}^{t} \sum_{u=0}^{T} \sum_{r=0}^{T} \sum_{k=1}^{u} j^{-1} k^{-1} \tau_r^0 \tau_s^0 \tau_t^0 \tau_u^0 \\
& \times \sum_{i=1}^{N} \left\{ E\left(\varepsilon_{i,t-j} \varepsilon_{is} \varepsilon_{i,u-k} \varepsilon_{ir}\right) - E\left(\varepsilon_{i,t-j} \varepsilon_{is}\right) E\left(\varepsilon_{i,u-k} \varepsilon_{ir}\right) \right\} \\
= \;& \frac{4\sigma^4}{T} S_{\tau\tau T}^{0-1} \sum_{t=1}^{T} \sum_{j=1}^{t} \sum_{u=0}^{T} j^{-1} |j+t-u|^{-1} \tau_t^0 \tau_u^0 + \frac{4\sigma^4}{T} S_{\tau\tau T}^{0-1} \sum_{t=1}^{T} \sum_{j=1}^{t} \sum_{u=1}^{T} \sum_{k=1}^{u} j^{-1} k^{-1} \tau_{t-j}^0 \tau_t^0 \tau_{u-k}^0 \tau_u^0 + O\left(T^{-1}\right) \\
= \;& O\left(T^{-d_0} \log T + T^{-1} \log^2 T\right) = o(1).
\end{aligned}
$$

Finally the expectation of the remaining term is

$$
2\sigma^2 \left(\frac{N}{T}\right)^{\frac{1}{2}} S_{\tau\tau T}^{0-1} \sum_{t=1}^{T} \tau_t^0 \dot{\tau}_t^0 = 2\sigma^2 \left(\frac{N}{T}\right)^{\frac{1}{2}} \frac{S_{\tau\dot{\tau}T}^{02}}{S_{\tau\tau T}^{0}}.
$$

Its variance is

$$
\frac{4 S_{\tau\tau T}^{0-4}}{NT} \sum_{i=1}^{N} \left[E\left(a_{iT}^{04}\right) - \left\{ E\left(a_{iT}^{02}\right) \right\}^2 \right] \sum_{t=1}^{T} \sum_{s=1}^{T} \tau_t^0 \dot{\tau}_t^0 \tau_s^0 \dot{\tau}_s^0 = o(1), \; as \; T \to \infty,
$$

since, as is readily shown, $E\left(a_{iT}^{04}\right) = O\left(S_{\tau\tau T}^{02}\right)$.

Overall, we deduce that

$$
(NT)^{\frac{1}{2}} \frac{\partial}{\partial d} L_T^F(d_0) + 2\sigma^2 \left(\frac{N}{T}\right)^{\frac{1}{2}} S_{\tau\dot{\tau}T}^{0} / S_{\tau\tau T}^{0} \to_d \mathcal{N}(0, \sigma^4 \pi^2/6).
$$

Using similar techniques as before, the probability limit of the second derivative term in (33) is $2\sigma^2\pi^2/6$, and the result follows. □

Proof of Theorem 4.3. We start with an analogous development to (33). Recalling $z_{it}(d) = \Delta_t^{d-d_0} \varepsilon_{it} - \pi_t(d-1)\varepsilon_{i0}$ and using Lemma 4,

$$
\begin{aligned}
\frac{\partial}{\partial d} L_T^D(d) &= \frac{2}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(\frac{\partial}{\partial d} z_{it}(d)\right) z_{it}(d) \\
&= \frac{2}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(J_{t+1}(L) \Delta_t^{d-d_0} \varepsilon_{it} - \dot{\tau}_t^0 \varepsilon_{i0}\right) \left(\Delta_{t+1}^{d-d_0} \varepsilon_{it} - \tau_t^0 \varepsilon_{i0}\right).
\end{aligned}
$$

Thus

$$
(NT)^{\frac{1}{2}} \frac{\partial}{\partial d} L_T^D(d_0) = \frac{2}{(NT)^{\frac{1}{2}}} \sum_{i=1}^{N} \sum_{t=1}^{T} \left(J_{t+1}(L) \varepsilon_{it} - \dot{\tau}_t^0 \varepsilon_{i0}\right) \left(\varepsilon_{it} - \tau_t^0 \varepsilon_{i0}\right). \tag{36}
$$

Expanding (36) reveals the same asymptotically $\mathcal{N}(0, 4\pi^2/6)$ term w_T, while, noting that

$$
J_{t+1}(L) \varepsilon_{it} = -\sum_{j=1}^{t-1} \frac{\varepsilon_{i,t-j}}{j} - \frac{\varepsilon_{i0}}{t},
$$

and employing similar arguments to before it is readily seen that the remainder of (36) differs by $o_p(1)$ from its expectation, which is

$$2\sigma^2 \left(\frac{N}{T}\right)^{1/2} \sum_{t=1}^{T} \tau_t^0 \left(1/t + \dot{\tau}_t^0\right).$$

The probability limit of the second derivative term is obtained much as before. $\square$

Proof of Theorem 4.4. Again we start as in (33). Then,

$$\frac{\partial}{\partial d} L_T^P(d) = \widehat{\sigma}_T^2(d) \frac{\partial}{\partial d} |\Omega_T(d)|^{\frac{1}{T}} + |\Omega_T(d)|^{\frac{1}{T}} \frac{\partial}{\partial d} \widehat{\sigma}_T^2(d),$$

where

$$\begin{aligned} \frac{\partial}{\partial d} |\Omega_T(d)|^{1/T} &= \frac{1}{T} |\Omega_T(d)|^{\frac{1}{T}} \frac{\partial}{\partial d} \log |\Omega_T(d)| \\ &= \frac{1}{T} |\Omega_T(d)|^{\frac{1}{T}} trace \left\{ \Omega_T(d)^{-1} \dot{\Omega}_T(d) \right\} \end{aligned}$$

and

$$\begin{aligned} \frac{\partial}{\partial d} \widehat{\sigma}_T^2(d) &= \frac{1}{NT} \sum i = 1^N \Big\{ 2\dot{\mathbf{z}}_{iT}'(d) \Omega_T(d)^{-1} \mathbf{z}_{iT}(d) \\ &\quad - \mathbf{z}_{iT}'(d) \Omega_T(d)^{-1} \dot{\Omega}_T(d) \Omega_T(d)^{-1} \mathbf{z}_{iT}(d) \Big\}, \end{aligned}$$

where we introduce $\dot{\Omega}_T(d) = (\partial/\partial d) \Omega_T(d)$, $\dot{\mathbf{z}}_{iT}(d) = (\partial/\partial d) \mathbf{z}_{iT}(d)$. Suppressing some d arguments and using (13) and

$$\dot{\Omega}_T(d) = \dot{\boldsymbol{\tau}}_T \boldsymbol{\tau}_T' + \boldsymbol{\tau}_T \dot{\boldsymbol{\tau}}_T',$$

we have

$$\begin{aligned} trace \left\{ \Omega_T(d)^{-1} \dot{\Omega}_T(d) \right\} &= trace \left\{ \left(I_T - \frac{\boldsymbol{\tau}_T \boldsymbol{\tau}_T'}{S_{\tau\tau T}} \right) \left(\dot{\boldsymbol{\tau}}_T \boldsymbol{\tau}_T' + \boldsymbol{\tau}_T \dot{\boldsymbol{\tau}}_T' \right) \right\} \\ &= \frac{2 S_{\tau\dot{\tau}T}}{S_{\tau\tau T}}. \end{aligned}$$

and

$$\begin{aligned} \Omega_T(d)^{-1} \dot{\Omega}_T(d) \Omega_T(d)^{-1} &= \left(I_T - \frac{\boldsymbol{\tau}_T \boldsymbol{\tau}_T'}{S_{\tau\tau T}} \right) \left(\dot{\boldsymbol{\tau}}_T \boldsymbol{\tau}_T' + \boldsymbol{\tau}_T \dot{\boldsymbol{\tau}}_T' \right) \left(I_T - \frac{\boldsymbol{\tau}_T \boldsymbol{\tau}_T'}{S_{\tau\tau T}} \right) \\ &= \frac{\dot{\boldsymbol{\tau}}_T \boldsymbol{\tau}_T' + \boldsymbol{\tau}_T \dot{\boldsymbol{\tau}}_T'}{S_{\tau\tau T}} - 2 \frac{\boldsymbol{\tau}_T \dot{\boldsymbol{\tau}}_T' \boldsymbol{\tau}_T \boldsymbol{\tau}_T'}{S_{\tau\tau T}^2}. \end{aligned}$$

Writing $\boldsymbol{\varepsilon}_{iT} = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})'$, $\mathbf{J}_T(L) = diag\{J_2(L), \ldots, J_{T+1}(L)\}$, $\boldsymbol{\tau}_T^0 = \boldsymbol{\tau}_T(d_0)$, $\dot{\boldsymbol{\tau}}_T^0 = \dot{\boldsymbol{\tau}}_T(d_0)$, and from the proof of Theorem 4.3

$$\mathbf{z}_{iT}(d_0) = \boldsymbol{\varepsilon}_{iT} - \boldsymbol{\tau}_T^0 \varepsilon_{i0}, \ \dot{\mathbf{z}}_{iT}(d_0) = \mathbf{J}_T(L) \boldsymbol{\varepsilon}_{iT} - \dot{\boldsymbol{\tau}}_T^0 \varepsilon_{i0}.$$

Thus, with $S_{\tau\tau T}^0 = S_{\tau\tau T}(d_0)$,

$$\begin{aligned} & \frac{2}{NT} \sum_{i=1}^{N} \dot{\mathbf{z}}_{iT}'(d_0) \Omega_T(d_0)^{-1} \mathbf{z}_{iT}(d_0) \\ = \ & \frac{2}{NT} \sum_{i=1}^{N} \left(\mathbf{J}_T(L) \boldsymbol{\varepsilon}_{iT} - \dot{\boldsymbol{\tau}}_T^0 \varepsilon_{i0} \right)' \left(I_T - \frac{\boldsymbol{\tau}_T^0 \boldsymbol{\tau}_T^{0\prime}}{S_{\tau\tau T}^0} \right) \left(\boldsymbol{\varepsilon}_{iT} - \boldsymbol{\tau}_T^0 \varepsilon_{i0} \right) \end{aligned}$$

has expectation

$$\frac{2\sigma^2}{T}\left(\dot{\boldsymbol{\tau}}_T^0+\mathbf{m}_T\right)'\left(I_T-\frac{\boldsymbol{\tau}_T^0\boldsymbol{\tau}_T^{0\prime}}{S_{\tau\tau T}^0}\right)\boldsymbol{\tau}_T^0=\frac{2\sigma^2}{T}\frac{\left(S_{\tau\dot{\tau}T}^0+S_{\tau mT}^0\right)}{S_{\tau\tau T}^0}.$$

Next

$$\begin{aligned}&-\frac{1}{NT}\sum_{i=1}^{N}\mathbf{z}_{iT}'\left(d_0\right)\Omega_T\left(d_0\right)^{-1}\dot{\Omega}_T\left(d_0\right)\Omega_T\left(d_0\right)^{-1}\mathbf{z}_{iT}\left(d_0\right)\\ =\;&-\frac{1}{NT}\sum_{i=1}^{N}\left(\boldsymbol{\varepsilon}_{iT}-\boldsymbol{\tau}_T^0\varepsilon_{i0}\right)'\left(\frac{\left(\dot{\boldsymbol{\tau}}_T^0\boldsymbol{\tau}_T^{0\prime}+\boldsymbol{\tau}_T^0\dot{\boldsymbol{\tau}}_T^{0\prime}\right)}{S_{\tau\tau T}^0}-2\frac{S_{\tau\dot{\tau}T}^0\boldsymbol{\tau}_T^0\boldsymbol{\tau}_T^{0\prime}}{S_{\tau\tau T}^{02}}\right)\left(\boldsymbol{\varepsilon}_{iT}-\boldsymbol{\tau}_T^0\varepsilon_{i0}\right)\end{aligned}$$

has expectation

$$\begin{aligned}&-\frac{1}{T}\left(2\frac{E\left(\dot{\boldsymbol{\tau}}_T^{0\prime}\boldsymbol{\varepsilon}_{iT}\boldsymbol{\varepsilon}_{iT}'\boldsymbol{\tau}_T^0\right)}{S_{\tau\tau T}^0}-2\frac{S_{\tau\dot{\tau}T}^0E\left(\boldsymbol{\tau}_T^{0\prime}\boldsymbol{\varepsilon}_{iT}\right)^2}{S_{\tau\tau T}^{02}}\right)-\sigma^2\frac{1}{T}\boldsymbol{\tau}_T^{0\prime}\left(\frac{\left(\dot{\boldsymbol{\tau}}_T^0\boldsymbol{\tau}_T^{0\prime}+\boldsymbol{\tau}_T^0\dot{\boldsymbol{\tau}}_T^{0\prime}\right)}{S_{\tau\tau T}^0}-2\frac{S_{\tau\dot{\tau}T}^0\boldsymbol{\tau}_T^0\boldsymbol{\tau}_T^{0\prime}}{S_{\tau\tau T}^{02}}\right)\boldsymbol{\tau}_T^0\\ =\;&-\frac{2\sigma^2}{T}\frac{S_{\tau\dot{\tau}T}^0}{S_{\tau\tau T}^0}.\end{aligned}$$

Thus

$$\begin{aligned}E\frac{\partial}{\partial d}L_T^P\left(d_0\right)&=\frac{1}{T}|\Omega_T\left(d_0\right)|^{1/T}\left\{E\left(\widehat{\sigma}_T^2\left(d_0\right)\right)trace\left\{\Omega_T\left(d_0\right)^{-1}\dot{\Omega}_T\left(d_0\right)\right\}+\frac{\partial}{\partial d}\widehat{\sigma}_T^2\left(d_0\right)\right\}\\&=\frac{\sigma^2}{T}|\Omega_T\left(d_0\right)|^{1/T}\left(\frac{2S_{\tau\dot{\tau}T}^0}{S_{\tau\tau T}^0}-\frac{2S_{\tau\dot{\tau}T}^0}{S_{\tau\tau T}^0}+\frac{2\left(S_{\tau\dot{\tau}T}^0+S_{\tau mT}^0\right)}{S_{\tau\tau T}^0}\right)\\&=\frac{2\sigma^2}{T}|\Omega_T\left(d_0\right)|^{1/T}\left(\frac{S_{\tau\dot{\tau}T}^0+S_{\tau mT}^0}{S_{\tau\tau T}^0}\right)\\&=-\frac{\sigma^2}{T}b_T^P\left(d_0\right)\left(1+O\left(\frac{S_{\tau\tau T}^0}{T}\right)\right).\end{aligned}$$

By similar means to before it may be shown that

$$(NT)^{1/2}\left\{\frac{\partial}{\partial d}L_T^P\left(d_0\right)-E\frac{\partial}{\partial d}L_T^P\left(d_0\right)\right\}=w_T+o_p(1)\to_d\mathcal{N}(0,4\pi^2/6).$$

We again omit the details of the convergence of the second derivative term. □

Proof of Theorem 5.1

We give the proof for $\widetilde{d}_T^F$ only, that for $\widetilde{d}_T^P$ is almost identical. We have

$$\begin{aligned}(NT)^{\frac{1}{2}}\left(\widetilde{d}_T^F-d_0\right)&=(NT)^{\frac{1}{2}}\left(\widehat{d}_T^F-d_0-T^{-1}b_T^F\left(d_0\right)\right)\\&\quad-(N/T)^{\frac{1}{2}}\left(b_T^F\left(\widehat{d}_T^F\right)-b_T^F\left(d_0\right)\right).\end{aligned}$$

It suffices to show that the second term on the right is $o_p(1)$. By the mean value theorem,

$$b_T^F\left(\widehat{d}_T^F\right)-b_T^F\left(d_0\right)=\frac{\partial}{\partial d}b_T^F\left(d^*\right)\left(\widehat{d}_T^F-d_0\right).$$

We have

$$\begin{aligned}\frac{\partial}{\partial d}b_T^F\left(d\right)&=-\frac{6}{\pi^2}\left(S_{\tau\tau T}(d)\frac{\partial}{\partial d}S_{\tau\dot{\tau}T}(d)-S_{\tau\dot{\tau}T}(d)\frac{\partial}{\partial d}S_{\tau\tau T}(d)\right)S_{\tau\tau T}^{-2}(d)\\&=O\left(\left|\frac{\partial}{\partial d}S_{\tau\dot{\tau}T}(d)\right|S_{\tau\tau T}^{-1}(d)+\left(S_{\tau\dot{\tau}T}(d)\right)^2S_{\tau\tau T}^{-2}(d)\right)\\&=O(\log^2T1\left(d\leq\frac{1}{2}\right)+1\left(d>\frac{1}{2}\right))\end{aligned}$$

from Lemma 2. Since $\widehat{d}_T^F - d_0 = O_p(\left|b_T^F(d_0)\right|/T + (NT)^{-\frac{1}{2}})$, where $b_T^F(d) = O(\log T 1(d \le \frac{1}{2}) + 1(d > \frac{1}{2}))$,

$$\begin{aligned}(N/T)^{\frac{1}{2}}\left(b_T^F\left(\widehat{d}_T^F\right) - b_T^F(d_0)\right) &= O_p\left((N/T)^{\frac{1}{2}}\log^2 T(\log T/T + (NT)^{-\frac{1}{2}})\right)\\ &= O_p\left(N^{\frac{1}{2}}T^{-\frac{3}{2}}\log^3 T + \log^2 T/T\right),\ d_0 \le \frac{1}{2},\end{aligned}$$

$$\begin{aligned}(N/T)^{\frac{1}{2}}\left(b_T^F\left(\widehat{d}_T^F\right) - b_T^F(d_0)\right) &= O_p\left((N/T)^{\frac{1}{2}}(T^{-1} + (NT)^{-\frac{1}{2}})\right)\\ &= O_p\left(N^{\frac{1}{2}}T^{-\frac{1}{2}} + T^{-1}\right),\ d_0 > \frac{1}{2},\end{aligned}$$

and these are $o(1)$ under the stated conditions. □

Proof of Theorem 5.2

As in the previous proof

$$\begin{aligned}(NT)^{\frac{1}{2}}\left(\widetilde{d}_T^D - d_0\right) &= (NT)^{\frac{1}{2}}\left(\widehat{d}_T^D - d_0 - T^{-1}b_T^D(d_0)\right)\\ &\quad - (N/T)^{\frac{1}{2}}\left(b_T^D\left(\widehat{d}_T^D\right) - b_T^D(d_0)\right),\end{aligned}$$

with

$$b_T^D\left(\widehat{d}_T^D\right) - b_T^D(d_0) = \frac{\partial}{\partial d}b_T^D(d^*)\left(\widehat{d}_T^D - d_0\right)$$

and

$$\begin{aligned}\frac{\partial}{\partial d}b_T^D(d) &= O\left(\left|\frac{\partial}{\partial d}S_{\tau\dot{\tau}T}(d)\right| + \left|\frac{\partial}{\partial d}S_{\tau m T}(d)\right|\right)\\ &= O\left(T^{1-2d}\log^2 T\right),\ d < \frac{1}{2},\\ &= O\left(\log^3 T\right),\ d = \frac{1}{2},\\ &= O(1),\ d > \frac{1}{2},\end{aligned}$$

applying Lemma 2 again. Since $\widehat{d}_T^D - d_0 = O_p(\left|b_T^D(d_0)\right|/T + (NT)^{-\frac{1}{2}})$, where $b_T^D(d) = O(T^{1-2d}\log T 1(d < \frac{1}{2}) + \log^2 T 1(d = \frac{1}{2}) + 1(d > \frac{1}{2}))$,

$$\begin{aligned}(N/T)^{\frac{1}{2}}\left(b_T^D\left(\widehat{d}_T^D\right) - b_T^D(d_0)\right) &= O_p\left((N/T)^{\frac{1}{2}}T^{1-2d_0}\log^2 T\left(T^{-2d_0}\log T + (NT)^{-\frac{1}{2}}\right)\right)\\ &= O_p\left(N^{\frac{1}{2}}T^{\frac{1}{2}-4d_0}\log^3 T + T^{-2d_0}\log^2 T\right),\ d_0 < \frac{1}{2},\end{aligned}$$

$$\begin{aligned}(N/T)^{\frac{1}{2}}\left(b_T^D\left(\widehat{d}_T^D\right) - b_T^D(d_0)\right) &= O_p\left((N/T)^{\frac{1}{2}}\log^3 T\left(\log^2 T/T + (NT)^{-1/2}\right)\right)\\ &= O_p\left(N^{\frac{1}{2}}T^{-\frac{3}{2}}\log^5 T + \log^3 T/T\right),\ d_0 = \frac{1}{2},\end{aligned}$$

$$\begin{aligned}(N/T)^{\frac{1}{2}}\left(b_T^D\left(\widehat{d}_T^D\right) - b_T^D(d_0)\right) &= O_p\left((N/T)^{\frac{1}{2}}\left(T^{-1} + (NT)^{-1/2}\right)\right)\\ &= O_p\left(N^{\frac{1}{2}}T^{-\frac{3}{2}} + T^{-1}\right),\ d_0 > \frac{1}{2},\end{aligned}$$

which are $o(1)$ under the stated conditions. □

Proof of Theorem 5.4

The proof is almost identical to that of Theorem 5.3, and is thus omitted. □

Appendix B: Proofs of Propositions

Proof of Proposition 1 For $\eta > 0$ let $N_\eta = \{d : |d - d_0| \leq \eta\}$, $\overline{N}_\eta = \{d : d \notin N_\eta, d \in D\}$. Writing $M_T(d) = L_T(d) - L_T(d_0)$,

$$\begin{aligned} P(\bar{d} \in \overline{N}_\eta) &\leq P\left(\inf_{\overline{N}_\eta} M_T(d) \leq 0\right) \\ &\leq P\left(\sup_D |V_T(d)| \geq \inf_{\overline{N}_\eta} U(d)\right), \end{aligned}$$

where $V_T(d) = U(d) - M_T(d)$ with

$$U(d) = \sigma_0^2 \sum_{j=1}^{\infty} \nu_j^2(d)$$

and the abbreviation $\nu_j(d) = \pi_j(d - d_0)$. Because $U(d)$ is continuous, vanishes if and only if $d = d_0$, and is otherwise positive, $\inf_{\overline{N}_\eta} U(d) > 0$. It remains to show that $\sup|V_T(d)| \to_p 0$. We have

$$T_T(d) = U(d) + A_T(d_0) - A_T(d) + B_T(d_0) - B_T(d),$$

so in view of (14) it suffices to show that

$$\sup_D |A_T(d) - A_T(d_0) - U(d)| \to_p 0. \tag{37}$$

We have

$$\begin{aligned} A_T(d) - A_T(d_0) - U(d) &= \frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{T}\nu_j^2(d)\sum_{t=0}^{T-j}\left(\varepsilon_{it}^2 - \sigma_0^2\right) \\ &+ \frac{2}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{t}\sum_{k=0}^{j-1}\nu_j(d)\nu_k(d)\varepsilon_{i,t-j}\varepsilon_{i,t-k} \\ &- \frac{\sigma_0^2}{T}\sum_{j=1}^{T}(j-1)\,\nu_j^2(d) - \sigma_0^2\sum_{j=T+1}^{\infty}\nu_j(d). \end{aligned}$$

Thus it remains to show that

$$\sup_D \left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{j=1}^{T}\nu_j^2(d)\sum_{t=0}^{T-j}\left(\varepsilon_{it}^2 - \sigma_0^2\right)\right| \to_p 0, \tag{38}$$

$$\sup_D \left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{j=1}^{t}\sum_{k=0}^{j-1}\nu_j(d)\nu_k(d)\varepsilon_{i,t-j}\varepsilon_{i,t-k}\right| \to_p 0, \tag{39}$$

$$\sup_D \frac{1}{T}\sum_{j=0}^{T}(j-1)\,\nu_j^2(d) \to 0, \tag{40}$$

$$\sup_D \sum_{j=T+1}^{\infty}\nu_j^2(d) \to 0. \tag{41}$$

The left sides of (38) and (39) are bounded respectively by

$$\frac{1}{N}\sum_{i=1}^{N}\sup_{D}\left|\frac{1}{T}\sum_{j=1}^{T}\nu_j^2(d)\sum_{t=0}^{T-j}\left(\varepsilon_{it}^2-\sigma_0^2\right)\right|, \tag{42}$$

$$\frac{1}{N}\sum_{i=1}^{N}\sup_{D}\left|\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{t}\sum_{k=0}^{j-1}\nu_j(d)\nu_k(d)\varepsilon_{i,t-j}\varepsilon_{i,t-k}\right|. \tag{43}$$

It follows from the proof of Theorem 1 of Hualde and Robinson (2011), constancy of the π_j across i, and identity of distribution of ε_{it} across i, that the $i-$summands in both (42) and (43) are $o_p(1)$ uniformly in i, whence (38) and (39) are true, and thus

$$\sup_{D}\ |\nu_j(d)| = O\left(\sup_{D} j^{-(d-d_0)-1}\right) = O\left(j^{(d_0-\underline{d})-1}\right) = O\left(j^{-\frac{1}{2}-\zeta/2}\right).$$

Then (40) and (41) are straightforwardly checked, to complete the proof. □

Proof of Proposition 2 The left side of (15) can be written $T^{-\frac{1}{2}}\sum_{t=0}^{T}v_{tT}$, where

$$v_{tT} = \frac{1}{N^{\frac{1}{2}}}\sum_{i=1}^{N}\varepsilon_{it}\left(J_{t+1}(L)\,\varepsilon_{it}\right),$$

and our notation stresses the possibility that N increases with T. Denoting by F_{t-1} the σ-field of events generated by $\{\varepsilon_{is}, i\geq 1, s<t\}$ we have

$$\left\{E\left(\sum_{t=0}^{T}v_{tT}\right)^2\right\}^{-\frac{1}{2}}\sum_{t=0}^{T}v_{tT}\to_d \mathcal{N}(0,1)$$

if (see e.g. Brown (1969))

$$E\left(v_{tT}\,|F_{t-1}\right) = 0,\ t\geq 1, \tag{44}$$

$$E\left(v_{tT}^4\right) \leq K,\ t\geq 0, \tag{45}$$

$$\frac{1}{T}\sum_{t=0}^{T}\left\{E\left(v_{tT}^2\,|F_{t-1}\right)-E\left(v_{tT}^2\right)\right\}\to_p 0. \tag{46}$$

Clearly (44) holds by serial independence of the ε_{it}. Since $J_{t+1}(L)\,\varepsilon_{it} = -\sum_{j=1}^{t}\varepsilon_{i,t-j}/j$,

$$E\left(v_{tT}^4\right) = O\left(\frac{1}{N^2}\sum_{h=1}^{N}\sum_{i=1}^{N}\left(\sum_{j=1}^{t}\frac{1}{j^2}\right)^2\right)\leq K,$$

to check (45). The left side of (46) has mean zero and variance

$$\frac{1}{T^2}\sum_{t=0}^{T}\left(E\left\{E\left(v_{tT}^2\,|F_{t-1}\right)^2\right\}-\left\{E\left(v_{tT}^2\right)\right\}^2\right). \tag{47}$$

Nowso

$$
\begin{aligned}
E\left\{E\left(v_{tT}^{2}\left|F_{t-1}\right.\right)^{2}\right\} &= \frac{1}{N^{2}}\sum_{h=1}^{N}\sum_{i=1}^{N}E\left\{\left(\sum_{j=1}^{t}\frac{\varepsilon_{h,t-j}}{j}\right)^{2}\left(\sum_{k=1}^{t}\frac{\varepsilon_{i,t-k}}{k}\right)^{2}\right\} \\
&= \sigma^{4}\left\{\sum_{j=1}^{t}j^{-2}\right\}^{2}+O\left(\frac{1}{N^{2}}\sum_{i=1}^{N}\left\{\left(\sum_{j=1}^{t}\frac{1}{j^{2}}\right)^{2}+\sum_{j=1}^{t}\frac{1}{j^{4}}\right\}\right) \\
&= \left\{E\left(v_{tT}^{2}\right)\right\}^{2}+O(N^{-1}),
\end{aligned}
$$

and thus the left side of (47) is $O((NT)^{-1})$, to check (46). $\square$

Appendix C: Technical Lemmas

Lemma 1 *For all d,*

$$
\Delta_{t+1}^{d}1=\sum_{j=0}^{t}\pi_{j}\left(d\right)=\pi_{t}\left(d-1\right).
$$

Proof of Lemma 1 The first equality is immediate. To prove the second, note that $\pi_t\left(d-1\right)$ is the coefficient of L^t in the expansion of Δ^{d-1}. But also, formally,

$$
\Delta^{d-1}=\Delta^{d}\Delta^{-1}=\sum_{t=0}^{\infty}\pi_{t}\left(d\right)L^{t}\sum_{j=0}^{\infty}L^{j}=\sum_{t=0}^{\infty}\left\{\sum_{j=0}^{t}\pi_{j}\left(d-1\right)\right\}L^{t}. \quad \square
$$

Introduce the digamma function $\psi\left(x\right)=\left(\partial/\partial x\right)\log\Gamma(x)$.

Lemma 2 *As $T\rightarrow\infty$,*

$$
\begin{aligned}
S_{\tau\tau T}(d) &= \frac{T^{1-2d}}{(1-2d)\Gamma(1-d)^{2}}+O(1), \quad d<\frac{1}{2}, \\
&= \frac{\log T}{\pi}+O(1), \quad d=\frac{1}{2}, \\
&= \frac{1}{(2d-1)\,B(d,d)}+O(T^{1-2d}), \ d>\frac{1}{2};
\end{aligned}
$$

$$
\begin{aligned}
S_{\tau\dot{\tau}T}(d) &= -\frac{T^{1-2d}\log T}{(1-2d)\Gamma(1-d)^{2}}+O(T^{1-2d}), \ d<\frac{1}{2}, \\
&= \frac{\log T}{\pi}+O(1), \quad d=\frac{1}{2}, \\
&= \frac{\psi\left(2d\right)-\psi\left(d\right)-\frac{1}{2d-1}}{(2d-1)\,B\left(d,d\right)}+O(T^{1-2d}\log T), \ d>\frac{1}{2};
\end{aligned}
$$

$$
S_{\tau mT}(d)=\int_{0}^{1}\left(\frac{\left(1-x\right)^{d-1}-1}{x}\right)dx+O(T^{-d}), \ \textit{all } d>0;
$$

$$\frac{\partial}{\partial d}S_{\tau\dot{\tau}T}(d) = O(T^{1-2d}\log^2 T1\left(d<\frac{1}{2}\right) + \log^3 T1\left(d=\frac{1}{2}\right) + 1\left(d>\frac{1}{2}\right));$$

$$\frac{\partial}{\partial d}S_{\tau m T}(d) = O(1), \ \textit{all } d > 0.$$

Proof of Lemma 2

For $d \le \frac{1}{2}$ we have

$$\left|\sum_{j=1}^{T} t^{-2d} - \int_1^T x^{-2d}dx\right| \le K$$

and from (6)

$$S_{\tau\tau T}(d) = 1 + \frac{1}{\Gamma(1-d)^2}\sum_{t=1}^{T} t^{-2d}(1+O(t^{-1})).$$

Thus, since $\int_1^T x^{-2d}dx = (1-2d)^{-1}T^{1-2d} + O(1)$, $d \le \frac{1}{2}$, $\int_1^T x^{-1}dx = \log T$, the approximations of $S_{\tau\tau T}(d)$ for $d \le \frac{1}{2}$ are readily checked, whereas that for $d > \frac{1}{2}$ follows because $\sum_{j=0}^{\infty}\tau_j^2(d) = \frac{1}{(2d-1)B(d,d)}$ and (6) implies $\sum_{j=T+1}^{\infty}\tau_j^2(d) = O(T^{1-2d})$ for $d > \frac{1}{2}$. Next, since

$$\dot{\tau}_j(d) = -\tau_j(d)\left\{\log j + O(1)\right\}, \text{ as } j\to\infty, \tag{48}$$

and, for $d \le \frac{1}{2}$

$$\left|\sum_{t=1}^{T} t^{-2d}\log t - \int_1^T x^{-2d}\log x dx\right| \le K,$$

where $\int_1^T x^{-2d}\log x dx = (1-2d)^{-1}T^{1-2d}(\log T + O(1))$, $d < \frac{1}{2}$, $\int_1^T x^{-1}\log x dx = \frac{1}{2}\log^2 T$, the approximations of $S_{\tau\dot{\tau}T}(d)$ with $d \le \frac{1}{2}$ may be checked, whereas that for $d > \frac{1}{2}$ follows because $\sum_{j=0}^{\infty}\tau_j(d)\dot{\tau}_t(d) = \frac{1}{2}(\partial/\partial d)\sum_{j=0}^{\infty}\tau_j^2(d)$, $(\partial/\partial d)\log B(d,d) = 2(\psi(2d)-\psi(d))$, and (6) and (48) imply that $\sum_{j=T+1}^{\infty}\tau_j(d)\dot{\tau}_t(d) = O(T^{1-2d}\log T)$. Given the identity $\sum_{j=1}^{\infty}\tau_j(d)/j = \int_0^1\left((1-x)^{d-1}-1\right)/xdx$, the remaining results follow similarly and straightforwardly. □

Lemma 3 *Uniformly in* i and d, *as* $T\to\infty$,

$$a_{iT}(d) = O_p\left(T^{\max(d_0-d,0)+\max(\frac{1}{2}-d,0)}\log^2 T\right).$$

Proof of Lemma 3 From (10), write $a_{iT}(d)$ as

$$a_{iT}(d) = \sum_{t=0}^{T} c_t(d)\varepsilon_{it},$$

where

$$c_t(d) = \sum_{j=0}^{T-t}\tau_{j+t}(d)\pi_j(d-d_0).$$

By summation-by-parts

$$a_{iT}(d) = \sum_{t=0}^{T-1}(c_t(d)-c_{t+1}(d))\sum_{s=0}^{t}\varepsilon_{is} + c_T(d)\sum_{t=0}^{T}\varepsilon_{it}.$$

Now

$$|c_T(d)| \le K\sum_{j=0}^{T}(j+t)^{-d}j^{d_0-d-1} \le KT^{-d}(T^{d_0-d}1(d<d_0) + \log T1(d=d_0) + 1(d>d_0)).$$

Also,

$$c_t(d) - c_{t+1}(d) = \sum_{j=0}^{T-t-1} \left(\tau_{j+t}(d) - \tau_{j+t+1}(d)\right) \pi_j(d-d_0) + \tau_T(d)\, \pi_{T-t}(d-d_0),$$

so

$$\begin{aligned} |c_t(d) - c_{t+1}(d)| &\leq K \sum_{j=1}^{T-t-1} (t+j)^{-d-1} j^{d_0-d-1} + KT^{-d}(T-t)^{d_0-d-1} \\ &\leq Kt^{-d-1} \sum_{j=1}^{T-t-1} j^{d_0-d-1} + KT^{-d}(T-t)^{d_0-d-1}. \end{aligned} \tag{50}$$

For $d < d_0$ (50) is bounded by $Kt^{-d-1}(T-t)^{d_0-d} + KT^{-d}(T-t)^{d_0-d-1}$, which is bounded by $Kt^{-d-1}T^{d_0-d}$ for $t \leq T/2$, and by $KT^{-d}(T-t)^{d_0-d-1}$ for $t \geq T/2$. For $d = d_0$ (50) is bounded by $Kt^{-d_0-1}\log T$ for $t \leq T/2$, and by $KT^{-d_0-1}\log T + KT^{-d_0}(T-t)^{-1}$ for $t \geq T/2$. For $d > d_0$ (50) is bounded by Kt^{-d-1} for $t \leq T/2$, and by $KT^{-d-1} + KT^{-d}(T-t)^{d_0-d-1}$ for $t \geq T/2$. Now for all i, t, $\sum_{s=0}^{t} \varepsilon_{is} = O_p\left(t^{1/2}\right)$ (the left side having variance $\sigma^2(t+1)$).

In view of these calculations, and the fact that $\sum_{s=0}^{t} \varepsilon_{is} = O_p\left(t^{1/2}\right)$ uniformly in i for $d < d_0$

$$\begin{aligned} a_{iT}(d) = O_p &\left(T^{d_0-d} \sum_{t=1}^{[T/2]} t^{\frac{1}{2}} + T^{-d} \sum_{t=[T/2]}^{T-1} (T-t)^{d_0-d-1} t^{\frac{1}{2}} + T^{d_0-2d+\frac{1}{2}} \right) \\ = O_p &\left(T^{d_0-2d+\frac{1}{2}} 1\left(d < \frac{1}{2}\right) + T^{d_0-\frac{1}{2}} \log T 1(d = \frac{1}{2}) + T^{d_0-d} 1(d > \frac{1}{2})\right) \\ = O_p &\left(T^{d_0-d+\max(\frac{1}{2}-d,0)} 1\left(d \neq \frac{1}{2}\right) + T^{d_0-\frac{1}{2}} 1(d = \frac{1}{2}\right) \\ = O_p &\left(T^{d_0-d+\max(\frac{1}{2}-d,0)} \log T \right) \end{aligned}$$

uniformly; for $d = d_0$

$$\begin{aligned} a_{iT}(d) &= O_p\left(\log T \sum_{t=1}^{[T/2]} t^{-d_0-\frac{1}{2}} + T^{-d_0} \log T \sum_{t=[T/2]}^{T-1} (T-t)^{-1} t^{\frac{1}{2}} + KT^{-d_0+\frac{1}{2}} \log T + \log T \right) \\ &= O_p(\log T 1(d_0 > \frac{1}{2}) + \log^2 T 1(d_0 = \frac{1}{2}) + T^{\frac{1}{2}-d_0} \log T 1(d_0 < 1/2) \\ &= O_p\left(T^{\max(\frac{1}{2}-d,0)} \log T 1\left(d \neq \frac{1}{2}\right) + \log^2 T 1(d = \frac{1}{2})\right) \\ &= O_p\left(T^{\max(\frac{1}{2}-d,0)} \log^2 T \right) \end{aligned}$$

uniformly; and for $d > d_0$,

$$\begin{aligned} a_{iT}(d) &= O_p\left(\sum_{t=1}^{[T/2]} t^{-d-1/2} + T^{-d} \sum_{t=[T/2]}^{T-1} (T-t)^{d_0-d-1} t^{1/2} + T^{-d-1} T^{3/2} + T^{1/2-d} \right) \\ &= O_p\left(1(d > 1/2) + \log T 1(d = 1/2) + T^{1/2-d} 1(d < 1/2)\right) \\ &= O_p\left(T^{\max(1/2-d,0)} 1(d \neq 1/2) + \log T 1(d = 1/2)\right) \\ &= O_p\left(T^{\max(1/2-d,0)} \log T \right) \end{aligned}$$

uniformly. The claimed bound is then readily assembled. □

Lemma 4 *For all* d,

$$\sum_{j=1}^{t} \frac{\pi_{t-j}(d)}{j} = -\dot{\pi}_t(d).$$

Proof of Lemma 4 Note that $\dot{\pi}_t(d)$ is the coefficient of L^t in the expansion of $(\partial/\partial d)\,\Delta^d$. But also

$$\frac{\partial}{\partial d}\Delta^d = \Delta^d \log\Delta = -\sum_{j=0}^{\infty}\pi_j(d)\,L^j\sum_{k=1}^{\infty}\frac{L^k}{k} = -\sum_{t=1}^{\infty}\left\{\sum_{j=1}^{t}\frac{\pi_{j-k}(d)}{k}\right\}L^t. \quad \square$$

References

Beran, J., 1995. Maximum likelihood estimation of the differencing parameter for invertible short and long memory autoregressive integrated moving average models. Journal of the Royal Statistical Society, Ser. B. 57, 659–672.

Brown, B.M., 1971. Martingale central limit theorems. Annals of Mathematical Statistics 42, 59–66.

Han, C. and P. C.B. Phillips, 2010. GMM estimation for dynamic panels with fixed effects and strong instruments at unity. Econometric Theory 26, 119-151.

Hassler, U., M. Demetrescu and A.L. Tarcolea, 2011. Asymptotic normal tests for integration in panels with cross-dependent units. Advances in Statistical Analysis 95, 187-204.

Hualde, J. and P. M. Robinson, P. M., 2011. Gaussian pseudo-maximum likelihood estimation of fractional time series models. Annals of Statistics 39, 3152-3181.

Robinson, P. M., 1994. Efficient tests of nonstationary hypotheses. Journal of the American Statistical Association 89, 1420-1437.

Velasco, C. and P. M. Robinson, 2000. Whittle pseudo-maximum likelihood estimation for nonstationary time series. Journal of the American Statistical Association 95, 1229–1243.

Table 1: Scaled asymptotic bias $b_T^F(d) \times 100/T$ of Fixed Effect estimate

$T \setminus d$:	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
5	-17.77	-11.04	-2.25	0	1.76	4.77
10	-11.54	-6.64	-1.17	0	0.85	2.24
100	-2.25	-1.04	-0.13	0	0.08	0.21

Table 2: Scaled asymptotic bias $b_T^D(d) \times 100/T$ of Differenced estimate

$T \setminus d$:	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
5	27.05	5.43	0.20	0	0.14	1.17
10	28.94	4.51	0.14	0	0.08	0.63
100	18.90	1.18	0.02	0	0.01	0.06

Table 3: Relative bias of PML and Difference estimates $b_T^P(d_0)/b_T^D(d_0)$

$T \setminus d$:	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
5	0.386	0.739	0.984	-	0.986	0.845
10	0.291	0.696	0.983	-	0.986	0.846
100	0.111	0.600	0.981	-	0.986	0.845

Table 4. Approximation to asymptotic bias of Fixed Effect estimate $b_T^F(d) \times 100/T$

$T \setminus d$:	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
5	-19.57	-22.02	-2.50	0	1.62	3.56
10	-14.00	-11.01	-1.25	0	0.81	1.78
100	-2.80	-1.10	-0.13	0	0.08	0.18

Table 5. Approximation to asymptotic bias of Difference estimate $b_T^D(d) \times 100/T$

$T \setminus d$:	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
5	55.27	82.64	0.44	0	0.20	1.28
10	52.17	41.32	0.22	0	0.10	0.64
100	26.21	4.13	0.02	0	0.01	0.06

Table 6: 100× Empirical bias of Fixed Effect estimates $\widehat{d}_T^F$, $\widetilde{d}_T^F$

	Uncorrected estimates $\widehat{d}_T^F$						Bias-corrected estimates $\widetilde{d}_T^F = \widehat{d}_T^F - b_T^F(\widehat{d}_T^F)/T$					
d_0 :	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
T	$NT = 100$						$NT = 100$					
5	-19.95	-45.42	-19.43	-6.51	-0.33	2.37	-0.42	-26.69	-12.57	-4.69	-1.62	-2.48
10	-17.80	-21.13	-4.27	-1.81	-0.42	0.50	-5.11	-11.34	-2.22	-1.41	-1.11	-1.73
T	$NT = 200$						$NT = 200$					
5	-20.00	-48.28	-14.06	-2.49	1.25	3.69	-0.47	-29.01	-8.34	-1.70	-0.52	-1.26
10	-18.92	-20.23	-2.88	-0.91	0.24	1.51	-6.15	-10.32	-1.17	-0.72	-0.57	-0.76
100	-12.95	-7.99	-1.54	-0.71	-0.19	0.43	-5.05	-3.07	-0.76	-0.63	-0.57	-0.66
T	$NT = 400$						$NT = 400$					
5	-20.00	-49.62	-9.11	-0.88	1.75	4.49	-0.47	-30.13	-4.48	-0.57	-0.18	-0.53
10	-19.58	-19.32	-2.20	-0.45	0.59	2.01	-6.77	-9.37	-0.65	-0.36	-0.28	-0.29
100	-13.38	-7.31	-1.11	-0.35	0.13	0.83	-5.45	-2.44	-0.40	-0.31	-0.28	-0.26

Table 7: 100× Empirical bias of Difference estimates $\widehat{d}_T^D$, $\widetilde{d}_T^D$

	Uncorrected estimates $\widehat{d}_T^D$						Bias-corrected estimates $\widetilde{d}_T^D = \widehat{d}_T^D - b_T^D(\widehat{d}_T^D)/T$					
d_0 :	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
T	$NT = 100$						$NT = 100$					
5	21.80	5.64	-0.76	-1.05	-0.85	-0.46	11.58	1.17	-1.36	-1.28	-1.09	-1.58
10	17.91	3.63	-0.93	-1.10	-1.00	-0.93	6.04	-0.56	-1.31	-1.22	-1.13	-1.53
T	$NT = 200$						$NT = 200$					
5	22.34	6.13	-0.28	-0.56	-0.36	0.66	13.01	2.22	-0.66	-0.67	-0.54	-0.51
10	18.63	4.14	-0.44	-0.60	-0.50	-0.00	8.07	0.50	-0.69	-0.66	-0.60	-0.62
100	15.12	2.59	-0.50	-0.59	-0.54	-0.33	4.04	-0.37	-0.65	-0.62	-0.59	-0.64
T	$NT = 400$						$NT = 400$					
5	22.65	6.42	-0.00	-0.29	-0.09	1.23	13.77	2.79	-0.29	-0.34	-0.25	0.03
10	19.06	4.47	-0.15	-0.31	-0.21	0.43	9.13	1.10	-0.34	-0.34	-0.30	-0.20
100	15.71	2.94	-0.20	-0.30	-0.24	0.06	5.56	0.22	-0.32	-0.31	-0.29	-0.26

Table 8: 100× Empirical bias of PML estimates $\widehat{d}_T^P$, $\widetilde{d}_T^P$

	Uncorrected estimates $\widehat{d}_T^P$						Bias-corrected estimates $\widetilde{d}_T^P = \widehat{d}_T^P - b_T^P(\widehat{d}_T^P)/T$					
d_0 :	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
T	$NT = 100$						$NT = 100$					
5	0.18	-1.86	-1.17	-0.98	-0.84	-1.22	-9.75	-6.19	-1.70	-1.17	-1.05	-2.14
10	-0.58	-1.43	-1.09	-1.00	-0.94	-1.27	-8.75	-4.87	-1.43	-1.11	-1.06	-1.77
T	$NT = 200$						$NT = 200$					
5	-0.31	-0.98	-0.58	-0.49	-0.42	-0.47	-10.47	-5.17	-0.95	-0.58	-0.58	-1.42
10	-0.54	-0.76	-0.56	-0.51	-0.47	-0.52	-8.84	-4.07	-0.80	-0.56	-0.57	-1.03
100	-0.55	-0.65	-0.56	-0.53	-0.52	-0.57	-6.49	-2.94	-0.70	-0.56	-0.57	-0.83
T	$NT = 400$						$NT = 400$					
5	-0.31	-0.46	-0.28	-0.23	-0.20	-0.17	-10.60	-4.57	-0.56	-0.28	-0.35	-1.15
10	-0.33	-0.38	-0.28	-0.26	-0.24	-0.23	-8.69	-3.60	-0.47	-0.28	-0.33	-0.75
100	-0.28	-0.31	-0.27	-0.26	-0.25	-0.24	-6.25	-2.54	-0.38	-0.27	-0.30	-0.51

Table 9: Empirical MSE×100 of Fixed Effect estimates $\widehat{d}_T^F$, $\widetilde{d}_T^F$

	Uncorrected estimates $\widehat{d}_T^F$						Bias-corrected estimates $\widetilde{d}_T^F = \widehat{d}_T^F - b_T^F(\widehat{d}_T^F)/T$					
d_0 :	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
T	$NT = 100$						$NT = 100$					
5	4.00	22.95	15.76	6.69	2.64	0.75	0.02	8.64	8.90	4.08	1.79	0.66
10	3.51	8.09	2.05	1.37	1.10	0.68	0.55	3.98	1.38	1.06	0.94	0.67
T	$NT = 200$						$NT = 200$					
5	4.00	24.04	9.49	2.15	0.87	0.51	0.00	8.91	5.23	1.31	0.61	0.35
10	3.69	6.29	0.88	0.61	0.52	0.40	0.48	2.69	0.59	0.48	0.45	0.36
100	2.17	1.56	0.52	0.45	0.42	0.35	0.71	0.81	0.43	0.40	0.39	0.35
T	$NT = 400$						$NT = 400$					
5	4.00	24.73	4.32	0.62	0.42	0.42	0.00	9.16	2.26	0.39	0.29	0.19
10	3.86	4.94	0.2	0.29	0.26	0.25	0.49	1.76	0.27	0.23	0.22	0.20
100	2.09	0.99	0.25	0.22	0.20	0.19	0.58	0.41	0.21	0.19	0.19	0.18

Table 10: 100× Empirical MSE of Difference estimates $\widehat{d}_T^D$, $\widetilde{d}_T^D$

	Uncorrected estimates $\widehat{d}_T^D$						Bias-corrected estimates $\widetilde{d}_T^D = \widehat{d}_T^D - b_T^D(\widehat{d}_T^D)/T$					
d_0 :	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
T	$NT = 100$						$NT = 100$					
5	5.79	1.43	1.19	1.20	1.19	0.82	3.96	1.95	1.37	1.25	1.17	0.80
10	4.09	1.01	0.92	0.93	0.93	0.73	3.17	1.56	1.03	0.96	0.92	0.72
T	$NT = 200$						$NT = 200$					
5	5.51	0.93	0.59	0.60	0.59	0.47	2.93	0.97	0.67	0.61	0.57	0.44
10	3.90	0.60	0.45	0.46	0.45	0.40	1.88	0.73	0.49	0.46	0.45	0.39
100	2.72	0.45	0.39	0.39	0.39	0.36	1.59	0.60	0.41	0.40	0.39	0.36
T	$NT = 400$						$NT = 400$					
5	5.40	0.70	0.30	0.30	0.30	0.27	2.51	0.54	0.33	0.30	0.29	0.24
10	3.85	0.42	0.22	0.23	0.22	0.21	1.42	0.36	0.24	0.23	0.22	0.21
100	2.68	0.27	0.19	0.19	0.19	0.18	0.95	0.29	0.20	0.19	0.19	0.18

Table 11: 100× Empirical MSE of PML estimates $\widehat{d}_T^P$, $\widetilde{d}_T^P$

	Uncorrected estimates $\widehat{d}_T^P$						Bias-corrected estimates $\widetilde{d}_T^P = \widehat{d}_T^P - b_T^P(\widehat{d}_T^P)/T$					
d_0 :	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
T	$NT = 100$						$NT = 100$					
5	2.37	2.22	1.26	1.09	1.00	0.75	4.07	3.31	1.42	1.14	0.98	0.75
10	1.56	1.40	0.94	0.87	0.83	0.68	2.77	2.05	1.03	0.89	0.82	0.69
T	$NT = 200$						$NT = 200$					
5	1.46	1.08	0.60	0.52	0.48	0.40	3.00	1.73	0.67	0.53	0.46	0.40
10	0.90	0.67	0.45	0.42	0.40	0.36	1.92	1.06	0.49	0.42	0.39	0.36
100	0.63	0.50	0.39	0.37	0.36	0.34	1.18	0.71	0.41	0.37	0.36	0.34
T	$NT = 400$						$NT = 400$					
5	0.82	0.52	0.29	0.25	0.23	0.21	2.18	0.93	0.32	0.26	0.22	0.21
10	0.47	0.33	0.22	0.20	0.19	0.18	1.35	0.57	0.23	0.20	0.19	0.19
100	0.32	0.25	0.19	0.18	0.18	0.17	0.77	0.37	0.20	0.18	0.17	0.17

Table 12: Empirical coverage of 95% CI based on $\widehat{d}_T^F$, $\widetilde{d}_T^F$

	Uncorrected estimates $\widehat{d}_T^F$						Bias-corrected estimates $\widetilde{d}_T^F = \widehat{d}_T^F - b_T^F(\widehat{d}_T^F)/T$					
d_0 :	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
T	$NT = 100$						$NT = 100$					
5	0.17	4.18	56.17	70.08	76.82	96.08	99.90	3.94	65.28	79.29	84.15	92.36
10	14.34	36.83	78.00	83.57	86.42	95.51	98.30	53.69	84.70	87.72	89.10	93.66
T	$NT = 200$						$NT = 200$					
5	0.01	1.36	55.98	71.08	76.79	97.63	99.99	1.22	68.82	80.81	84.65	92.75
10	4.25	26.89	78.22	84.27	86.88	96.71	99.37	48.75	85.69	88.36	89.50	94.12
100	34.47	61.87	87.23	89.64	90.79	96.23	66.33	78.74	90.46	91.34	91.77	95.04
T	$NT = 400$						$NT = 400$					
5	0	0.18	54.20	71.41	75.85	67.41	100.00	0.19	70.20	81.60	85.16	93.17
10	0.39	14.63	77.75	84.57	86.91	85.98	99.88	41.74	86.03	88.66	89.75	94.59
100	16.27	52.86	87.56	90.04	91.09	91.50	59.08	77.94	90.87	91.72	92.10	92.35

Table 13: Empirical coverage of 95% CI based on $\widehat{d}_T^D$, $\widetilde{d}_T^D$

	Uncorrected estimates $\widehat{d}_T^D$						Bias-corrected estimates $\widetilde{d}_T^D = \widehat{d}_T^D - b_T^D(\widehat{d}_T^D)/T$					
d_0 :	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
T	$NT = 100$						$NT = 100$					
5	25.38	79.33	84.26	84.05	84.08	93.01	49.49	73.94	81.94	83.65	84.83	92.20
10	37.15	87.32	89.09	88.98	89.07	93.90	59.10	80.50	87.57	88.64	89.38	93.45
T	$NT = 200$						$NT = 200$					
5	5.84	71.93	83.90	83.75	83.94	94.13	36.60	72.45	81.91	83.56	84.79	92.79
10	11.91	83.00	89.31	89.16	89.14	94.75	50.94	80.10	87.96	89.01	89.54	94.01
100	24.68	89.49	91.70	91.61	91.64	95.27	59.11	84.83	90.95	91.53	91.81	94.97
T	$NT = 400$						$NT = 400$					
5	0.26	58.03	83.87	83.77	84.02	83.80	20.17	69.13	81.95	83.68	84.9	86.33
10	0.93	74.37	89.32	89.18	89.29	89.67	37.27	79.32	88.03	89.10	89.76	90.34
100	4.56	85.27	91.93	91.83	91.91	92.17	51.33	84.91	91.28	91.79	92.11	92.28

Table 14: Empirical coverage of 95% CI based on $\widehat{d}_T^P$, $\widetilde{d}_T^P$

	Uncorrected estimates $\widehat{d}_T^P$						Bias-corrected estimates $\widetilde{d}_T^P = \widehat{d}_T^P - b_T^P(\widehat{d}_T^P)/T$					
d_0 :	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**	**0.3**	**0.6**	**0.9**	**1.0**	**1.1**	**1.4**
T	$NT = 100$						$NT = 100$					
5	59.04	71.41	83.97	86.23	87.81	93.60	48.77	63.38	82.09	85.94	88.38	92.84
10	73.30	81.69	89.13	90.25	90.92	94.35	60.36	73.98	87.80	89.99	91.15	93.95
T	$NT = 200$						$NT = 200$					
5	58.88	71.45	84.41	86.74	88.21	94.10	43.86	61.83	82.59	86.61	88.88	92.76
10	73.34	81.97	89.59	90.67	91.30	94.94	53.96	72.52	88.37	90.54	91.60	94.34
100	82.29	87.63	91.91	92.42	92.69	95.41	67.29	81.06	91.21	92.30	92.85	95.09
T	$NT = 400$						$NT = 400$					
5	59.08	71.63	84.98	87.36	88.85	90.36	35.72	59.27	83.11	87.30	89.61	89.98
10	73.47	82.22	89.94	91.11	91.72	92.32	43.44	70.20	88.80	91.03	92.09	92.17
100	82.56	87.94	92.34	92.96	93.28	93.45	58.46	79.65	91.67	92.91	93.44	93.32