Truncated sum of squares estimation of fractional time series models with deterministic trends*

Javier Hualde†
Universidad Pública de Navarra

Morten Ørregaard Nielsen‡
Queen's University and CREATES

May 15, 2017

Abstract

We consider truncated (or conditional) sum of squares estimation of a parametric model composed of a fractional time series and an additive generalized polynomial trend. Both the memory parameter, which characterizes the behaviour of the stochastic component of the model, and the exponent parameter, which drives the shape of the deterministic component, are considered not only unknown real numbers, but also lying in arbitrarily large (but finite) intervals. Thus, our model captures different forms of nonstationarity and noninvertibility. As in related settings, the proof of consistency (which is a prerequisite for proving asymptotic normality) is challenging due to non-uniform convergence of the objective function over a large admissible parameter space, but, in addition, our framework is substantially more involved due to the competition between stochastic and deterministic components. We establish consistency and asymptotic normality under quite general circumstances, finding that results differ crucially depending on the relative strength of the deterministic and stochastic components.

JEL classification: C22.

Key words and phrases: Asymptotic normality, consistency, deterministic trend, fractional process, generalized polynomial trend, noninvertibility, nonstationarity, truncated sum of squares estimation.

---

*We are grateful to Peter M. Robinson and seminar participants at Universidad Carlos III, Universitat Pompeu Fabra, Universidad de Alicante, and at the 3rd CREATES Long Memory Symposium for useful comments. Javier Hualde's research is supported by the Spanish Ministerio de Economía y Competitividad through project ECO2015-64330-P. Morten Ø. Nielsen's research is supported by the Canada Research Chairs program and the Social Sciences and Humanities Research Council of Canada (SSHRC). Both authors are thankful to the Center for Research in Econometric Analysis of Time Series (CREATES, funded by the Danish National Research Foundation, DNRF78) for financial support.

†Email: `javier.hualde@unavarra.es`

‡Corresponding author. Email: `mon@econ.queensu.ca`

1 Introduction

In time series analysis the most common approach to modelling stochastic components is by stationary and invertible autoregressive moving average (ARMA) processes, but unit root nonstationary and noninvertible processes have also been considered. Additionally, supplementing the random component, the presence of a low-order polynomial term, such as a constant or a linear deterministic trend is usually assumed.

More recently, the relatively simple ARMA modeling framework has been generalized in various directions. Here, one of the main developments is that of fractionally integrated ARMA (FARIMA) models which bridge the behavioral gap between stationary and invertible ARMA, which has "memory parameter" δ_0 equal to zero, and unit root nonstationary process, where $\delta_0 = 1$. A zero-mean FARIMA(p_1, δ_0, p_2) process z_t is given by

$$z_t = \Delta_+^{-\delta_0} u_t,\ t = 0, \pm 1, \pm 2, \ldots, \tag{1}$$

$$\alpha(L) u_t = \beta(L) \varepsilon_t,\ t = 0, \pm 1, \pm 2, \ldots, \tag{2}$$

where L is the lag operator and Δ_+^{ζ} is given by $\Delta_+^{\zeta} x_t = \Delta^{\zeta} x_t \mathbb{I}(t \geq 1) = \sum_{i=0}^{t-1} \pi_i(-\zeta) x_{t-i}$ with $\pi_i(v) = 0$ for $i < 0$, $\pi_0(v) = 1$, and

$$\pi_i(v) = \frac{\Gamma(v+i)}{\Gamma(v)\Gamma(1+i)} = \frac{v(v+1)\ldots(v+i-1)}{i!},\ i \geq 1, \tag{3}$$

denoting the coefficients in the usual binomial expansion of $(1-z)^{-v}$ and $\mathbb{I}(\cdot)$ the indicator function. Additionally, $\alpha(L)$ and $\beta(L)$ are real polynomials of degrees p_1 and p_2, which share no common zeros and have all their zeros outside the unit circle in the complex plane, and ε_t is a zero-mean, serially uncorrelated and homoscedastic sequence. More precise conditions will be imposed below.

For the sake of greater generality, we retain (1) but generalize (2) to

$$u_t = \omega(L; \boldsymbol{\varphi}_0) \varepsilon_t,\ t = 0, \pm 1, \pm 2, \ldots, \tag{4}$$

where $\boldsymbol{\varphi}_0$ is an unknown $p \times 1$ vector and $\omega(s; \boldsymbol{\varphi}) = \sum_{j=0}^{\infty} \omega_j(\boldsymbol{\varphi}) s^j$ with, for all $\boldsymbol{\varphi}$, $\omega_0(\boldsymbol{\varphi}) = 1$ and $|\omega(s; \boldsymbol{\varphi})| \neq 0, |s| \leq 1$. Like α and β in (2), ω in (4) characterizes parametric short memory autocorrelation.

Although many theoretical developments have exclusively assumed a purely random process (see, e.g., the discussion in Hualde and Robinson, 2011, or Nielsen, 2015), in practice, model (1), (4) (or a semiparametric version of it, where u_t is a nonparametric invertible weakly dependent process, that is with spectrum which is bounded and bounded away from zero at all frequencies) is usually complemented with deterministic components. A simple extension of the fractional model is to allow a non-zero mean as in, for example,

$$\Delta_+^{\delta_0} x_t = \mu_0 + u_t \text{ or } x_t = \Delta_+^{-\delta_0} \mu_0 + \Delta_+^{-\delta_0} u_t, \tag{5}$$

so, noting that $\sum_{j=0}^{t-1} \pi_j(\delta_0) = \pi_{t-1}(1+\delta_0)$, introducing a non-zero mean as in (5) leads to consideration of $x_t = \mu_0 \pi_{t-1}(1+\delta_0) + z_t$. This partly motivates the more general model

$$x_t = \mu_0 \pi_{t-1}(\gamma_0) + z_t, \tag{6}$$

where μ_0 and γ_0 are both unknown real-valued parameters and z_t is generated by (1) and (4), and so in particular can be either short or long memory. For mathematical convenience, given that by Stirling's approximation $\pi_{t-1}(c)$ behaves, apart from a constant factor, like t^{c-1}, we use $\pi_{t-1}(c)$ rather than t^{c-1} in (6). The reason is the property $\Delta_+^d \pi_{t-1}(c) = \pi_{t-1}(c-d)$, which is only shared approximately by t^{c-1}, in the sense that $\Delta_+^d t^{c-1}$ does not equal a constant times t^{c-d-1}, although, as $t \to \infty$ the rate of growth of $\Delta_+^d t^{c-1}$ is t^{c-d-1}. Model (6) covers standard cases like a constant ($\gamma_0 = 1$) or linear trend ($\gamma_0 = 2$) (because $\pi_{t-1}(1) = \mathbb{I}(t \geq 1)$ and $\pi_{t-1}(2) = t\mathbb{I}(t \geq 1)$, see (3)), but, given that γ_0 is allowed to take any real value, it characterizes a wide range of situations.

In addition, although the literature stresses low-order polynomials in t, such as a constant or linear function, to capture deterministic behavior, this seems arbitrary in light of the fractional behavior of z_t. Thus, we consider in (6) the generalized polynomial trend, or power law trend, given by $\pi_{t-1}(\gamma_0)$, and in this way we allow the deterministic structure to be of a fractional order, thereby mimicking the fractional behavior of the stochastic component. This can be made more precise using the terminology of White and Granger (2011), where different definitions of trends appear. Because $Var(x_t)$ grows at rate $t^{2\delta_0-1}$ when $\delta_0 > 1/2$, whereas $E(x_t) = \mu_0 \pi_{t-1}(\gamma_0)$ evolves at rate t^{γ_0-1}, according to White and Granger (2011), if $\delta_0 > 1/2$, the process (6) has a "stochastic trend in variance", whereas if $\gamma_0 \neq 1$, it also has a "stochastic trend in mean". We note that the evolution of these two trends is governed by the parameters δ_0 and γ_0, respectively, and hence letting γ_0 be real-valued appears as natural as letting δ_0 be real-valued, it just affects another aspect of the distribution of x_t.

We note that if γ_0 were known in (6), the estimation problem is simplified greatly. In this case, one can eliminate the deterministic component by differencing and simply estimate $\delta_0 - \gamma_0$ from $\Delta_+^{\gamma_0} x_t = \mu_0 \pi_{t-1}(0) + \Delta_+^{\gamma_0} z_t = \Delta_+^{\gamma_0} z_t$ for $t \geq 2$ by standard parametric methods which assume zero mean and apply for arbitrary values of the memory parameter (even large negative values), e.g. Hualde and Robinson (2011). Having thus estimated δ_0, and denoting this estimator $\widetilde{\delta}$, one could estimate μ_0 by regression, i.e. by $\widehat{\mu}(\widetilde{\delta})$, where $\widehat{\mu}(\delta) = \sum_{t=1}^T \pi_{t-1}(\gamma_0 - \delta) \Delta_+^\delta x_t / \sum_{t=1}^T \pi_{t-1}^2(\gamma_0 - \delta)$, and the problem reduces to a question of whether $\widehat{\mu}(\widetilde{\delta})$ has identical limiting distribution to $\widehat{\mu}(\delta_0)$. This question will be covered by our more general theory, which assumes γ_0 unknown. The inclusion of deterministic trends of *known* order and the idea of differencing-and-adding-back is also considered by Velasco (1999a, 1999b) and Chen and Hurvich (2003) in combination with tapering of the periodogram in frequency domain methods. Similarly, Robinson (2005) considers M-estimation of a model like (6) (although involving more deterministic terms) with *known* γ_0 and allowing for fractional z_t.

On the other hand, several authors have considered the same type of generalized polynomial trends as in (6), with γ_0 being an unknown real-valued parameter, in similar contexts to ours. For the same type of truncated/conditional sum-of-squares estimator that we analyze in this paper, Wu (1981) noted in his Example 4 that model (6) does not satisfy his assumptions for the asymptotic analysis, even when z_t is an independent sequence, because of the asymptotic singularity of the Hessian and the requirement that the parameters μ and γ have different normalizations. The analysis of Wu (1981) was generalized by Phillips (2007) to allow such singularity of the Hessian and hence accommodate model (6), but assuming at most weakly dependent errors, z_t. In a spatial setting, Robinson (2012) considers a general version of (6) involving more deterministic terms, but with only weakly dependent errors, z_t, and explicitly excludes the situation where the deterministic component is dominated by the stochastic component, i.e. where $\gamma_0 - 1/2 < \delta_0$ in our notation. The latter situation is discussed in Johansen and Nielsen (2016), who consider truncated/conditional sum-of-squares estimation of (6) with $\gamma_0 = 1$ and $\delta_0 > 1/2$, and hence $\gamma_0 - 1/2 < \delta_0$, although with u_t being an independent sequence, and prove consistency and asymptotic normality of the standard estimator which ignores the deterministic term. Finally, Robinson and Marinucci (2000) and Robinson and Iacone (2005) consider semiparametric frequency domain estimators in models that include both a generalized polynomial trend and a nonstationary FARIMA stochastic component.

In this paper we analyze the model (6) with the stochastic term z_t given by (1) and (4), and prove consistency and asymptotic normality of the parameter estimators. Note that (6) with $\mu_0 = 0$ assumed as known corresponds to the model discussed in Hualde and Robinson (2011), where the behaviour of the observable process is entirely driven by δ_0 and the short-memory parameters $\boldsymbol{\varphi}_0$. As will be seen, the complication added by the consideration of unknown deterministic parameters is substantial, especially because we let both the memory (δ_0) and exponent (γ_0) parameters lie in arbitrarily large, but finite, intervals, so x_t can display many different behaviours. As in related works, e.g. Hualde and Robinson (2011) and Nielsen (2015), the proof of consistency (which is a prerequisite for proving asymptotic normality) is challenging due to non-uniform convergence of the objective function over a large admissible parameter space. However, in addition to this well-known complication, our framework is substantially more involved due to the competition between the stochastic and deterministic components, and this competition needs to be explicitly taken advantage of in the proof of consistency. Thus, we establish consistency and asymptotic normality under quite general circumstances, finding that results differ substantially depending on the relative strength of the deterministic and stochastic components. In particular, when $\gamma_0 - 1/2 > \delta_0$ we find that the estimators of all parameters in the model are consistent and asymptotically normally distributed. On the other hand, when $\gamma_0 - 1/2 < \delta_0$ the parameters related to the deterministic part of the model, μ_0 and γ_0, cannot be consistently estimated, but, interestingly, those related to the stochastic part of the model, i.e. δ_0 and $\boldsymbol{\varphi}_0$, are still consistently estimated and their asymptotic normal distribution is unaffected by the presence of the remaining unestimable parameters. This latter result resembles that of Heyde and Dai (1996), who provided conditions under which small trends do not affect the properties of Whittle estimators applied to short- or long-range dependent processes. Similar results were derived by Abadir, Distaso, and Giraitis (2007) and Iacone (2010) for different versions of the local Whittle estimator. Actually, in comparison to these works, our main results can be view as a step forward in the difficult task of disentangling the persistence properties of observed time series from the low-frequency effect of deterministic components. In fact, our proposed estimators of the stochastic components of the model retain their limiting properties regardless the intensity of the deterministic signal, which seems an advantage of our approach in this context. Finally, we include a small Monte Carlo simulation study which supports our theoretical results and illustrates the findings.

The next section formalizes the model and assumptions. In Section 3 we present the estimator and our main results on consistency and asymptotic normality. The results of some Monte Carlo simulations are reported in Section 4, and concluding remarks are presented in Section 5. The proofs of the main theorems are given in Appendix A, which applies some auxiliary lemmas and technical lemmas provided in Appendices B and C, respectively. The supplementary material in Hualde and Nielsen (2017) contains proofs of all lemmas.

2 Model and assumptions

As usual, we let the true values of the parameters be denoted by subscript zero. We consider the model (6), where z_t is generated by (1) and (4), and μ_0, γ_0, δ_0, and $\boldsymbol{\varphi}_0$ are unknown parameters to be estimated.

We first impose an assumption on the short memory component, ω, where $\boldsymbol{\varphi}_0$ is assumed to lie in Ψ, which is a compact and convex subset of $\mathbb{R}^p$.

A1. **(i)** for all $\boldsymbol{\varphi} \in \Psi \backslash \{\boldsymbol{\varphi}_0\}$, $|\omega(s; \boldsymbol{\varphi})| \neq |\omega(s; \boldsymbol{\varphi}_0)|$ on a set $S \subset \{s : |s| = 1\}$ of positive Lebesgue measure;

(ii) for all $\boldsymbol{\varphi} \in \Psi$, $\omega\left(e^{i\lambda}; \boldsymbol{\varphi}\right)$ is differentiable in λ with derivative in $\text{Lip}\,(\varsigma)$ for $1/2 < \varsigma \leq 1$;

(iii) for all λ, $\omega\left(e^{i\lambda}; \boldsymbol{\varphi}\right)$ is continuous in $\boldsymbol{\varphi}$;

(iv) for all $\boldsymbol{\varphi} \in \Psi$, $|\omega\,(s; \boldsymbol{\varphi})| \neq 0, |s| \leq 1$.

Assumption A1 is identical to A1 in Hualde and Robinson (2011). In particular, (i) ensures identification, (ii) and (iv) imply that u_t is an invertible weakly dependent process, while by (ii) and (iii), for all j, $\sup_{\boldsymbol{\varphi}\in\Psi} |\omega_j\,(\boldsymbol{\varphi})| = O\left(j^{-1-\varsigma}\right)$ as $j \to \infty$ (see Hualde and Robinson, 2011). Also, writing $\omega^{-1}(s; \boldsymbol{\varphi}) = \phi(s; \boldsymbol{\varphi}) = \sum_{j=0}^{\infty} \phi_j(\boldsymbol{\varphi}) s^j$, it holds that $\phi_0\,(\boldsymbol{\varphi}) = 1$ for all $\boldsymbol{\varphi}$, and (ii), (iii), and (iv) imply that

$$\sup_{\boldsymbol{\varphi}\in\Psi} |\phi_j\,(\boldsymbol{\varphi})| = O\left(j^{-1-\varsigma}\right) \text{ as } j \to \infty, \tag{7}$$

whereas (ii) also implies that

$$\inf_{|s|=1, \boldsymbol{\varphi}\in\Psi} |\phi\,(s; \boldsymbol{\varphi})| > 0. \tag{8}$$

A1 is easily satisfied in the stationary and invertible ARMA case. Another model covered by A1 is the exponential spectrum model of Bloomfield (1973), which leads to a relatively simple covariance matrix formula in the context of fractional time series models, see Robinson (1994). More generally, A1 is also similar to other conditions employed in asymptotic theory for the estimate $\widehat{\boldsymbol{\tau}} = (\widehat{\delta}, \widehat{\boldsymbol{\varphi}}')'$ below, see Hualde and Robinson (2011) and Nielsen (2015), as well as Whittle estimators that restrict to stationarity, e.g. Fox and Taqqu (1986), Dahlhaus (1989), and Giraitis and Surgailis (1990). Assumption A1 can be readily verified because ω is a known parametric function. In fact ω satisfying A1 are invariably employed by practitioners.

A2. The ε_t in (4) are stationary and ergodic with finite fourth moment, $E\left(\varepsilon_t | \mathcal{F}_{t-1}\right) = 0$, $E\left(\varepsilon_t^2 | \mathcal{F}_{t-1}\right) = \sigma_0^2$, a.s., where $\mathcal{F}_t$ is the σ-field of events generated by ε_s, $s \leq t$, and conditional (on $\mathcal{F}_{t-1}$) third and fourth moments of ε_t equal the corresponding unconditional moments.

Assumption A2 is identical to A2 in Hualde and Robinson (2011). It does not impose independence or identity of distribution of ε_t, but rules out conditional heteroskedasticity. It is standard in the time series asymptotics literature since Hannan (1973).

A3. The parameter space for $\boldsymbol{\vartheta} = (\delta, \boldsymbol{\varphi}', \gamma)'$ is given by $\Xi = [\triangledown_1, \triangledown_2] \times \Psi \times [\square_1, \square_2]$ with $\triangledown_1 < \triangledown_2$ and $\square_1 < \square_2$, where Ψ is compact and convex and $\boldsymbol{\vartheta}_0 = (\delta_0, \boldsymbol{\varphi}_0', \gamma_0)' \in \Xi$. For μ the parameter space is $\mathbb{R}$, and if $\gamma_0 - 1/2 > \delta_0$, we also assume that $\mu_0 \neq 0$.

We assume that $\mu_0 \neq 0$ when $\gamma_0 - 1/2 > \delta_0$ since otherwise γ_0 is not identified. A similar identification problem arises when $\gamma_0 - \delta_0 = 0, -1, -2, \ldots$, but this is not in fact problematic because in those cases the deterministic components cannot be consistently estimated anyway (see part (ii) of Theorems 1 and 2 below). In fact, even very small trends (small γ_0) can be identified when δ_0 is sufficiently small ($\delta_0 < \gamma_0 - 1/2$) because then, in a sense, the value of δ_0 helps the identification of the deterministic part.

Finally, note that the model where γ_0 is known, e.g., the model with a constant term or a linear trend, is a special case of our model with unknown γ_0. Hence, the asymptotic results below can easily be specialized to this situation. In general, though, $[\triangledown_1, \triangledown_2]$ and $[\square_1, \square_2]$ are allowed to be arbitrarily large.

3 Truncated sum of squares estimation

We collect the parameters for the stochastic component in $\boldsymbol{\tau} = (\delta, \boldsymbol{\varphi}')'$ with true value $\boldsymbol{\tau}_0 = (\delta_0, \boldsymbol{\varphi}_0')'$, and denote the estimator (to be defined below) by $\widehat{\boldsymbol{\tau}} = (\widehat{\delta}, \widehat{\boldsymbol{\varphi}}')'$. We also use the notation $\boldsymbol{\vartheta} = (\boldsymbol{\tau}', \gamma)'$, $\boldsymbol{\vartheta}_0 = (\boldsymbol{\tau}_0', \gamma_0)'$, and $\widehat{\boldsymbol{\vartheta}} = (\widehat{\boldsymbol{\tau}}', \widehat{\gamma})'$. The Gaussian log-likelihood, conditional on $x_t = 0$ for $t \leq 0$, is, apart from a constant, given by

$$L_T(\boldsymbol{\vartheta}, \mu, \sigma^2) = -\frac{T}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\left(\phi(L;\boldsymbol{\varphi})\Delta_+^{\delta}(x_t - \mu\pi_{t-1}(\gamma))\right)^2$$
$$= -\frac{T}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\left(\phi(L;\boldsymbol{\varphi})x_t(\delta) - \mu c_{t-1}(\gamma-\delta, \boldsymbol{\varphi})\right)^2, \tag{9}$$

where

$$x_t(\delta) = \Delta_+^{\delta} x_t$$

and we have defined

$$c_{t-1}(d, \boldsymbol{\varphi}) = \phi(L;\boldsymbol{\varphi})\pi_{t-1}(d) = \sum_{j=0}^{t-1}\phi_j(\boldsymbol{\varphi})\pi_{t-j-1}(d). \tag{10}$$

Clearly, the likelihood function (9) is quadratic in μ, so for any given $\boldsymbol{\vartheta}$ we concentrate with respect to μ and find

$$\widehat{\mu}(\boldsymbol{\vartheta}) = \frac{\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})x_t(\delta)\, c_{t-1}(\gamma-\delta,\boldsymbol{\varphi})}{\sum_{t=1}^{T}c_{t-1}^2(\gamma-\delta,\boldsymbol{\varphi})}, \tag{11}$$

and we then propose the estimator

$$\widehat{\boldsymbol{\vartheta}} = \arg\min_{\boldsymbol{\vartheta}\in\Xi} R_T(\boldsymbol{\vartheta}), \quad R_T(\boldsymbol{\vartheta}) = \frac{1}{T}\sum_{t=1}^{T}\left(\phi(L;\boldsymbol{\varphi})x_t(\delta) - \widehat{\mu}(\boldsymbol{\vartheta})\, c_{t-1}(\gamma-\delta,\boldsymbol{\varphi})\right)^2,$$

along with $\widehat{\mu} = \widehat{\mu}(\widehat{\boldsymbol{\vartheta}})$. This estimator, often termed nonlinear least squares or conditional sum-of-squares, although we prefer the term truncated sum-of-squares as suggested by Hualde and Robinson (2011), is motivated by the Gaussian likelihood function (9) and is therefore expected to be asymptotically efficient under Gaussianity (though we do not assume Gaussianity anywhere in the analysis). The estimator goes back to, at least, Box and Jenkins (1971) for estimation of nonfractional ARMA models (where δ_0 is a known integer). In the context of fractional time series, the estimator was first analyzed by Li and McLeod (1986) in stationary FARIMA models with $0 < \delta_0 < 1/2$, Beran (1995), and later by Hualde and Robinson (2011) and Nielsen (2015) for δ_0 lying in arbitrarily large compact intervals.

Theorem 1 *Let (1), (4), (6), and Assumptions A1–A3 hold.*

(i) *If* $\gamma_0 - 1/2 > \delta_0$ *then* $\widehat{\boldsymbol{\vartheta}} \to_p \boldsymbol{\vartheta}_0$ *as* $T \to \infty$.

(ii) *If* $\gamma_0 - 1/2 < \delta_0$ *then* $\widehat{\boldsymbol{\tau}} \to_p \boldsymbol{\tau}_0$ *as* $T \to \infty$.

We note that the result in part (i) of Theorem 1 includes consistency of the estimator of the parameter vector $\boldsymbol{\vartheta}_0$. Under (i) μ_0 can also be consistently estimated, but we do not report this result here: $\widehat{\mu}$ is given in explicit form and consistency is not a prerequisite to justify its limiting distribution, which we provide in Theorem 2. The result in part (ii) only includes consistency of the estimator of $\boldsymbol{\tau}_0$. In fact, γ_0 and μ_0 cannot possibly be consistently estimated in the case in part (ii), where the deterministic signal is not strong enough. This is easily seen by considering for example $\delta_0 = 1$ (a random walk) in which case the deterministic parameters cannot be estimated consistently when $\gamma_0 - 1 < 1/2$ because the deterministic signal is drowned by the stochastic noise. This is the well-known result that a level ($\gamma_0 = 1$) cannot be estimated consistently for a unit root process ($\delta_0 = 1$), whereas a linear trend ($\gamma_0 = 2$) can be consistently estimated. Another example is $\delta_0 = 0$ (short memory) in which case trends of order $\gamma_0 - 1 < -1/2$ cannot be estimated consistently. In other words, suppose the unknown deterministic component is a (weakly) decreasing trend, or more generally is dominated by a constant ($\gamma_0 = 1$), then $\widehat{\gamma}$ and $\widehat{\mu}$ are consistent as long as $\delta_0 < 1/2$, i.e. z_t is (asymptotically) stationary. Thus, a notable feature about part (ii) of Theorem 1 is that, even though γ_0 and μ_0 cannot be consistently estimated, the remaining parameters $\boldsymbol{\tau}_0$ can still be consistently estimated.

As in related work, e.g. Hualde and Robinson (2011) and Nielsen (2015), the proof of Theorem 1 is challenging due to non-uniform convergence of the objective function over a large admissible parameter space for δ. However, in addition to this well-known complication, our framework is substantially more involved due to the competition between the stochastic and deterministic components. In our proof of Theorem 1(i), this competition is used explicitly for some parts of the parameter space; for details of the proof strategy please see Section A.1.1.

Next, we discuss the asymptotic distribution of our estimators, which requires an additional regularity condition.

A4. (i) $\boldsymbol{\vartheta}_0 \in \operatorname{int}(\Xi)$;

(ii) for all λ, $\omega\left(e^{i\lambda}; \boldsymbol{\varphi}\right)$ is thrice continuously differentiable in $\boldsymbol{\varphi}$ on a closed neighbourhood $\mathcal{N}_\epsilon\left(\boldsymbol{\varphi}_0\right)$ of radius $\epsilon \in (0, 1/2)$ about $\boldsymbol{\varphi}_0$, and for all $\boldsymbol{\varphi} \in \mathcal{N}_\epsilon(\boldsymbol{\varphi}_0)$ these partial derivatives with respect to $\boldsymbol{\varphi}$ are themselves differentiable in λ with derivative in $\operatorname{Lip}(\varsigma)$ for $1/2 < \varsigma \leq 1$;

(iii) the matrix

$$\boldsymbol{A} = \begin{pmatrix} \pi^2/6 & -\sum_{j=1}^{\infty} \boldsymbol{b}_j'(\boldsymbol{\varphi}_0)/j \\ -\sum_{j=1}^{\infty} \boldsymbol{b}_j(\boldsymbol{\varphi}_0)/j & \sum_{j=1}^{\infty} \boldsymbol{b}_j(\boldsymbol{\varphi}_0)\boldsymbol{b}_j'(\boldsymbol{\varphi}_0) \end{pmatrix}$$

is nonsingular, where $\boldsymbol{b}_j(\boldsymbol{\varphi}_0) = \sum_{k=0}^{j-1} \omega_k(\boldsymbol{\varphi}_0)\,\partial\phi_{j-k}(\boldsymbol{\varphi}_0)/\partial\boldsymbol{\varphi}$.

This assumption is almost identical to A3 in Hualde and Robinson (2011), with the only difference that our A4(ii) is slightly stronger than their A3(ii) in imposing thrice instead of twice continuously differentiable $\omega(e^{i\lambda};\boldsymbol{\varphi})$ and corresponding $\mathrm{Lip}(\varsigma)$ conditions, which appear to be necessary to derive the bounds in (12) below and the corresponding bounds in Hualde and Robinson (2011, p. 3169). The main reason for strengthening the assumption in Hualde and Robinson (2011) is to obtain the bounds (12) and also that, in our proof, third derivatives of $\phi_j(\boldsymbol{\varphi})$ are involved in the proof of convergence of the Hessian matrix below. As in (7), letting φ_i denote the i-th element of $\boldsymbol{\varphi}$, A1(ii), A1(iv) and A4(ii) imply that, as $j \to \infty$,

$$\sup_{\boldsymbol{\varphi}\in\mathcal{N}_\epsilon(\boldsymbol{\varphi}_0)} \left| \frac{\partial\phi_j(\boldsymbol{\varphi})}{\partial\varphi_i} \right| = O(j^{-1-\varsigma}), \sup_{\boldsymbol{\varphi}\in\mathcal{N}_\epsilon(\boldsymbol{\varphi}_0)} \left| \frac{\partial^2\phi_j(\boldsymbol{\varphi})}{\partial\varphi_i\partial\varphi_l} \right| = O(j^{-1-\varsigma}),$$
$$\sup_{\boldsymbol{\varphi}\in\mathcal{N}_\epsilon(\boldsymbol{\varphi}_0)} \left| \frac{\partial^3\phi_j(\boldsymbol{\varphi})}{\partial\varphi_i\partial\varphi_l\partial\varphi_k} \right| = O(j^{-1-\varsigma}). \tag{12}$$

Again A4 is satisfied in the ARMA case.

Define the scaling matrices

$$\boldsymbol{P}_T = \begin{pmatrix} \boldsymbol{M}_T & 0 \\ 0 & T^{1-(\gamma_0-\delta_0)}\log T \end{pmatrix}, \boldsymbol{M}_T = \begin{pmatrix} I_{p+1} & 0 \\ 0 & T^{1-(\gamma_0-\delta_0)} \end{pmatrix},$$

and also

$$\boldsymbol{V} = \begin{pmatrix} \sigma_0^2\boldsymbol{A} & 0 \\ 0 & \frac{\mu_0^2\phi^2(1;\boldsymbol{\varphi}_0)}{\Gamma^2(\gamma_0-\delta_0)(2(\gamma_0-\delta_0)-1)^3} \end{pmatrix}.$$

Theorem 2 *Let (1), (4), (6), and Assumptions A1–A4 hold.*

(i) *If $\gamma_0 - 1/2 > \delta_0$, then, as $T \to \infty$,*

$$T^{1/2}\boldsymbol{P}_T^{-1}\begin{pmatrix} \widehat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0 \\ \widehat{\mu} - \mu_0 \end{pmatrix} \to_d \begin{pmatrix} I_{p+2} \\ 0 \;\; -\mu_0 \end{pmatrix}\boldsymbol{N}, \tag{13}$$

where $\boldsymbol{N}$ is a random variable distributed as $N\left(0, \sigma_0^2\boldsymbol{V}^{-1}\right)$.

(ii) *If $\gamma_0 - 1/2 < \delta_0$, then, as $T \to \infty$,*

$$T^{1/2}(\widehat{\boldsymbol{\tau}} - \boldsymbol{\tau}_0) \to_d N\left(0, \boldsymbol{A}^{-1}\right). \tag{14}$$

A notable feature of the results in Theorem 2 is that the asymptotic distribution of $\widehat{\boldsymbol{\tau}}$ is unaffected by the presence of the deterministic component in (6), and $\widehat{\boldsymbol{\tau}}$ has the same asymptotic distribution as in, e.g., Theorem 2.2 of Hualde and Robinson (2011). Moreover, as with the consistency result in Theorem 1, the asymptotic distribution result for $\widehat{\boldsymbol{\tau}}$ in Theorem 2 is also unaffected by the relative magnitudes of the stochastic and deterministic components. In particular, even when $\gamma_0 - 1/2 < \delta_0$, so that γ_0 and μ_0 cannot be consistently estimated, the asymptotic distribution of $\widehat{\boldsymbol{\tau}}$ is unaffected. In fact, the variance $\boldsymbol{A}^{-1}$ in the asymptotic distribution of $\widehat{\boldsymbol{\tau}}$ in both (13) and (14) is equal to the inverse Fisher information under Gaussianity, see also Dahlhaus (1989). Because the estimate $\widehat{\boldsymbol{\tau}}$ is also asymptotically independent of the remaining coefficient estimates, it therefore follows that $\widehat{\boldsymbol{\tau}}$ is asymptotically efficient under the additional assumption of Gaussianity.

We notice from (13) in Theorem 2 that $\widehat{\gamma}$ is $T^{\gamma_0-\delta_0-1/2}$-consistent whereas $\widehat{\mu}$ is only $T^{\gamma_0-\delta_0-1/2}/\log T$-consistent. In fact, if γ_0 were known, then the least squares regression estimator of μ would be $T^{\gamma_0-\delta_0-1/2}$-consistent, and hence there is a rate-of-convergence loss, albeit small, in not knowing γ_0.

The joint asymptotic distribution of $\widehat{\boldsymbol{\vartheta}}$ and $\widehat{\mu}$ given in (13) is singular, which makes testing of joint hypotheses on $\boldsymbol{\vartheta}_0$ and μ_0 impossible. However, separate inference can be conducted on $\boldsymbol{\vartheta}_0$ and μ_0. For example, it is straightforward given (13) to construct confidence intervals and test hypotheses, e.g. that $\gamma_0 = 1$ (constant).

4 Monte Carlo evidence

We investigate the finite sample performance of our estimators of γ_0 and δ_0 by means of a simple Monte Carlo experiment. We generate the observable series $x_t, t = 1, ..., T$, from (6) with $z_t = \varepsilon_t$ being an independent $N(0,1)$ sequence and $T \in \{64, 128, 256, 512\}$. Without loss of generality we fix $\delta_0 = 1$ and set $\gamma_0 - \delta_0 = -0.9 + 0.2i$ for $i = 1, ..., 13$. We computed $\widehat{\delta}$, $\widehat{\gamma}$ using the optimizing intervals $\delta \in [\delta_0 - 5, \delta_0 + 5]$, $\gamma \in [\gamma_0 - 5, \gamma_0 + 5]$, and we report Monte Carlo bias and standard deviation (SD) over 10,000 replications.

Results for Monte Carlo bias are presented in Table 1. Here, the performance of $\widehat{\delta}$ reflects the limiting theory developed in Theorems 1 and 2. The bias of $\widehat{\delta}$ is clearly decreasing in absolute value as T increases, even for the boundary case $\gamma_0 - \delta_0 = 1/2$, which is not covered by our theory. It is also noticeable that when the deterministic signal gets stronger (so $\gamma_0 - \delta_0$ is higher) results worsen. As expected, the behaviour of $\widehat{\gamma}$ is qualitatively different. When $\gamma_0 - \delta_0 \leq 1/2$, the bias of $\widehat{\gamma}$ is large (in absolute value) and in general does not decrease as T increases. On the other hand, the picture changes dramatically when $\gamma_0 - \delta_0 > 1/2$, with very small biases as $\gamma_0 - \delta_0$ gets larger, reflecting the fast convergence rates in those cases implied by Theorem 2.

The Monte Carlo SD is reported in Table 2. Results for $\widehat{\delta}$ are as expected, but now larger $\gamma_0 - \delta_0$ lead to slightly smaller SD for smaller sample sizes. Regarding $\widehat{\gamma}$, for $\gamma_0 - \delta_0 \leq 1/2$ the SD is very large and quite stable for different values of T. As anticipated from Theorem 2, for $\gamma_0 - \delta_0 > 1/2$, the SD of $\widehat{\gamma}$ decreases with T and is very small for larger $\gamma_0 - \delta_0$.

5 Concluding remarks

We have proposed and analyzed a parametric model which covers a wide range of situations characterized by general deterministic and stochastic components. These are driven by power law and memory parameters, γ_0 and δ_0, respectively, which are assumed to lie in sets which can be arbitrarily large. Our model might display many different behaviours, including "stochastic trend in mean and/or variance" and various types of dependence (antipersistency, weak dependence, long memory). Our results depend crucially on whether the deterministic

Table 1: Monte Carlo bias of $\widehat{\delta}$ and $\widehat{\gamma}$

	$\widehat{\delta}$				$\widehat{\gamma}$			
$\gamma_0 - \delta_0$	$T = 64$	$T = 128$	$T = 256$	$T = 512$	$T = 64$	$T = 128$	$T = 256$	$T = 512$
−0.7	−0.079	−0.038	−0.020	−0.010	0.379	0.217	0.137	0.087
−0.5	−0.081	−0.041	−0.020	−0.011	0.333	0.169	0.085	0.012
−0.3	−0.083	−0.041	−0.020	−0.011	0.189	0.135	−0.008	0.030
−0.1	−0.082	−0.040	−0.022	−0.010	0.157	−0.002	−0.001	−0.117
0.1	−0.084	−0.040	−0.021	−0.011	0.180	0.035	−0.134	−0.126
0.3	−0.079	−0.040	−0.021	−0.011	−0.042	−0.112	−0.100	−0.210
0.5	−0.085	−0.043	−0.021	−0.011	−0.090	−0.092	−0.173	−0.152
0.7	−0.098	−0.051	−0.026	−0.014	−0.058	0.023	0.040	0.031
0.9	−0.116	−0.058	−0.029	−0.015	0.025	0.016	0.008	0.004
1.1	−0.117	−0.057	−0.029	−0.016	0.011	0.004	0.001	0.001
1.3	−0.121	−0.059	−0.031	−0.016	0.001	0.001	0.001	0.000
1.5	−0.125	−0.062	−0.031	−0.017	0.000	0.000	0.000	0.000
1.7	−0.122	−0.060	−0.031	−0.016	0.000	0.000	0.000	0.000

Note: Based on 10,000 Monte Carlo replications.

signal is sufficiently strong. If this is the case, that is if $\gamma_0 - \delta_0 > 1/2$, all parameters can be consistently estimated and their estimators are asymptotically normal. Interestingly, the limiting results for estimators corresponding to the stochastic part of the model ($\widehat{\boldsymbol{\tau}}$) are identical to those achieved in the simpler, purely stochastic, setting of Hualde and Robinson (2011). When the deterministic signal is weak, i.e., $\gamma_0 - \delta_0 < 1/2$, γ_0 and μ_0 cannot be consistently estimated, but, nicely, $\widehat{\boldsymbol{\tau}}$ retains identical limiting properties as when $\gamma_0 - \delta_0 > 1/2$.

There are several interesting issues which have not been addressed in the present paper, but which will the object of future research. First, one could argue that the deterministic part of our model, which contains a single term, is too simplistic. However, our methods of proof should be extendable to cover a richer setting, allowing for multiple deterministic terms characterized by different power law parameters, such as

$$x_t = \sum_{j=1}^{d} \mu_{0j} \pi_{t-1} (\gamma_{0j}) + z_t,$$

where z_t is given in (1), (4), and, without loss of generality, we set $\gamma_{01} < \gamma_{02} < ... < \gamma_{0d} < \infty$. Whenever it exists, let $\dagger = \min\{j = 1, ..., d : \gamma_{0j} - 1/2 > \delta_0\}$ and $\mu_{0j} \neq 0$ for any $j \geq \dagger$. Our estimator can be extended to accommodate this greater generality in an obvious way, and we conjecture that results qualitatively identical to those in Theorem 2 apply. In particular, the estimator of $\boldsymbol{\tau}_0$ would retain identical properties irrespective of the strength of the deterministic signal(s), and whenever $\dagger$ exists, the estimators of γ_{0j} for $j \geq \dagger$ will be $T^{\gamma_{0j} - \delta_0 - 1/2}$-consistent and asymptotically normal. However, considering formally this extension would come at the cost of greater complication, and given that our present setting is already quite involved, we preferred to keep things as simple as possible at this stage,

Table 2: Monte Carlo standard deviation of $\widehat{\delta}$ and $\widehat{\gamma}$

	$\widehat{\delta}$				$\widehat{\gamma}$			
$\gamma_0 - \delta_0$	$T = 64$	$T = 128$	$T = 256$	$T = 512$	$T = 64$	$T = 128$	$T = 256$	$T = 512$
-0.7	0.148	0.088	0.057	0.038	2.958	3.003	2.955	2.973
-0.5	0.146	0.089	0.057	0.038	3.023	3.056	3.048	3.014
-0.3	0.146	0.089	0.057	0.038	3.079	3.091	3.104	3.098
-0.1	0.145	0.088	0.056	0.038	3.106	3.121	3.128	3.128
0.1	0.146	0.087	0.057	0.038	3.064	3.126	3.112	3.092
0.3	0.145	0.088	0.057	0.038	3.011	2.984	2.971	2.952
0.5	0.149	0.091	0.057	0.038	2.566	2.532	2.478	2.384
0.7	0.152	0.090	0.057	0.038	1.575	1.167	0.828	0.529
0.9	0.145	0.087	0.056	0.037	0.464	0.189	0.098	0.070
1.1	0.141	0.086	0.056	0.037	0.117	0.073	0.047	0.030
1.3	0.142	0.086	0.056	0.038	0.069	0.038	0.022	0.012
1.5	0.143	0.086	0.057	0.038	0.040	0.020	0.010	0.005
1.7	0.141	0.086	0.056	0.038	0.023	0.010	0.004	0.002

Note: Based on 10,000 Monte Carlo replications.

so the proofs present in a clear way the essence of the problem of the competition between deterministic and stochastic terms. Second, a semiparametric approach which focuses on estimating γ_0 and δ_0 without making parametric assumptions about the structure of z_t seems possible. Third, the fractional process which characterizes our model has been termed as "Type II". Nevertheless, it seems that our theory could also be developed for the so-called "Type I" fractional process.

A Proofs of theorems

A.1 Proof of Theorem 1(i): the $\gamma_0 - 1/2 > \delta_0$ case

A.1.1 Overall design of the proof

Throughout, ϵ will denote a generic arbitrarily small positive constant, and K a generic arbitrarily large positive constant. Fix $\varepsilon > 0$ and let $M_\varepsilon = \{\boldsymbol{\vartheta} \in \Xi : \|\boldsymbol{\tau} - \boldsymbol{\tau}_0\| < \varepsilon\}$, $\overline{M}_\varepsilon = \{\boldsymbol{\vartheta} \in \Xi : \|\boldsymbol{\tau} - \boldsymbol{\tau}_0\| \geq \varepsilon\}$, $N_\varepsilon = \{\boldsymbol{\vartheta} \in \Xi : |\gamma - \gamma_0| < \varepsilon\}$ and $\overline{N}_\varepsilon = \{\boldsymbol{\vartheta} \in \Xi : |\gamma - \gamma_0| \geq \varepsilon\}$. Then $\Pr(||\widehat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0|| \geq \varepsilon) \to 0$ as $T \to \infty$, is implied by

$$\Pr(\widehat{\boldsymbol{\vartheta}} \in \overline{M}_\varepsilon) \to 0 \text{ as } T \to \infty, \tag{15}$$

$$\Pr(\widehat{\boldsymbol{\vartheta}} \in \overline{N}_\varepsilon \cap M_\varepsilon) \to 0 \text{ as } T \to \infty. \tag{16}$$

Strictly, ε should be $\varepsilon/\sqrt{2}$ in (15) and (16), but since ε is arbitrary this is irrelevant and we continue without the $\sqrt{2}$ factor.

We decompose the objective function as $R_T(\boldsymbol{\vartheta}) = \frac{1}{T}\sum_{t=1}^{T}(d_t(\boldsymbol{\vartheta}) + s_t(\boldsymbol{\vartheta}))^2$ with

$$d_t(\boldsymbol{\vartheta}) = \mu_0 \left(c_{t-1}(\gamma_0 - \delta, \boldsymbol{\varphi}) - h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) \sum_{j=1}^{T} c_{j-1}(\gamma_0 - \delta, \boldsymbol{\varphi}) h_{j-1,T}(\gamma - \delta, \boldsymbol{\varphi}) \right),$$

$$s_t(\boldsymbol{\vartheta}) = \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) - h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) \sum_{j=1}^{T} \phi(L; \boldsymbol{\varphi}) u_j(\delta - \delta_0) h_{j-1,T}(\gamma - \delta, \boldsymbol{\varphi}),$$

defining also the coefficient

$$h_{t,T}(d, \boldsymbol{\varphi}) = \frac{c_t(d, \boldsymbol{\varphi})}{(\sum_{j=1}^{T} c_{j-1}^2(d, \boldsymbol{\varphi}))^{1/2}}, \tag{17}$$

which clearly satisfies $\sum_{t=1}^{T} h_{t-1,T}^2(d, \boldsymbol{\varphi}) = 1$.

The strategy of proof relies on recognizing the competition between the stochastic term $s_t(\boldsymbol{\vartheta})$ and deterministic term $d_t(\boldsymbol{\vartheta})$ in $R_T(\boldsymbol{\vartheta})$, taking into account that when considering (15), just $\boldsymbol{\tau}$ is for sure "far" from $\boldsymbol{\tau}_0$, whereas when dealing with (16), just γ is "far" from γ_0. As will be seen, an important feature of the problem is that when $\gamma = \gamma_0$ we have $d_t(\boldsymbol{\vartheta}) = 0$, which complicates the treatment of (15). In any case, as in Hualde and Robinson (2011), we need to carefully consider the cases where $R_T(\boldsymbol{\vartheta})$ shows distinct behaviours, noting that either the deterministic or the stochastic term might dominate, and below we partition the parameter space accordingly.

A.1.2 Proof of (15)

To prove (15) we use

$$\Pr(\widehat{\boldsymbol{\vartheta}} \in \overline{M}_\varepsilon) = \Pr\left(\inf_{\boldsymbol{\vartheta} \in \overline{M}_\varepsilon} R_T(\boldsymbol{\vartheta}) \le \inf_{\boldsymbol{\vartheta} \in M_\varepsilon} R_T(\boldsymbol{\vartheta})\right) \le \Pr\left(\inf_{\boldsymbol{\vartheta} \in \overline{M}_\varepsilon} S_T(\boldsymbol{\vartheta}) \le 0\right), \tag{18}$$

where $S_T(\boldsymbol{\vartheta}) = R_T(\boldsymbol{\vartheta}) - R_T(\boldsymbol{\vartheta}_0)$. Fix an arbitrarily small $\eta > 0$ such that $\eta < (\gamma_0 - \delta_0 - 1/2)/2$ and suppose that $\bigtriangledown_1 < \delta_0 - 1/2 - \eta$ and $\bigtriangledown_2 > \gamma_0 - 1 - \eta$. Our proof will cover trivially the situation where any of these conditions does not hold, in which case some of the steps below are superfluous. Let $\mathcal{I}_1 = \{\delta : \bigtriangledown_1 \le \delta \le \delta_0 - 1/2 - \eta\}$, $\mathcal{I}_2 = \{\delta : \delta_0 - 1/2 - \eta \le \delta \le \delta_0 - 1/2\}$, $\mathcal{I}_3 = \{\delta : \delta_0 - 1/2 \le \delta \le \delta_0 - 1/2 + \eta\}$, $\mathcal{I}_4 = \{\delta : \delta_0 - 1/2 + \eta \le \delta \le \gamma_0 - 1 - \eta\}$, and $\mathcal{I}_5 = \{\delta : \gamma_0 - 1 - \eta \le \delta \le \bigtriangledown_2\}$, noting that the upper bound for η guarantees that $\mathcal{I}_4$ is non-empty. Correspondingly define $\mathcal{T}_i = \mathcal{I}_i \times \Psi$ and, fixing $\xi > 0$ and $\varrho > 0$, such that $\varrho < \eta/2$, also define $\mathcal{H}_i = \{\boldsymbol{\vartheta} \in \overline{M}_\varepsilon : \boldsymbol{\tau} \in \mathcal{T}_i, |\gamma - \gamma_0| < \xi T^{-\varkappa_i}\}$, $\overline{\mathcal{H}}_i = \{\boldsymbol{\vartheta} \in \overline{M}_\varepsilon : \boldsymbol{\tau} \in \mathcal{T}_i, \xi T^{-\varkappa_i} \le |\gamma - \gamma_0| \le \varrho\}$ and $\overline{\overline{\mathcal{H}}}_i = \{\boldsymbol{\vartheta} \in \overline{M}_\varepsilon : \boldsymbol{\tau} \in \mathcal{T}_i, |\gamma - \gamma_0| \ge \varrho\}$, $i = 1, \ldots, 5$, where $\varkappa_i > 0$ will be defined subsequently, noting that $\overline{\mathcal{H}}_i$ is non-empty for any ξ, ϱ, for T large enough . Then, by (18), (15) is justified by showing

$$\Pr\left(\inf_{\overline{\overline{\mathcal{H}}}_i} S_T(\boldsymbol{\vartheta}) \le 0\right) \to 0 \text{ as } T \to \infty \text{ for } i = 1, \ldots, 5, \tag{19}$$

$$\Pr\left(\inf_{\overline{\mathcal{H}}_i} S_T(\boldsymbol{\vartheta}) \le 0\right) \to 0 \text{ as } T \to \infty \text{ for } i = 1, \ldots, 5, \tag{20}$$

$$\Pr\left(\inf_{\mathcal{H}_i} S_T(\boldsymbol{\vartheta}) \le 0\right) \to 0 \text{ as } T \to \infty \text{ for } i = 1, \ldots, 5. \tag{21}$$

We note that $\mathcal{H}_i$, $\overline{\mathcal{H}}_i$, and $\overline{\overline{\mathcal{H}}}_i$ are designed exactly such that in $\mathcal{H}_i$ the stochastic term dominates $S_T(\boldsymbol{\vartheta})$ while in $\overline{\mathcal{H}}_i \cup \overline{\overline{\mathcal{H}}}_i$ it is the deterministic term that dominates. As will be seen, the analysis on $\overline{\overline{\mathcal{H}}}_i$ is much simpler because γ is "far" from γ_0, whereas a much more delicate treatment in necessary for $\overline{\mathcal{H}}_i$. This motivates a separate analysis of (19), (20) and (21), at least for $i = 1, \ldots, 4$.

Proof of (19), (20), and (21) for $i = 5$ In this case, we give just one proof that covers the whole set $\mathcal{H}_5 \cup \overline{\mathcal{H}}_5 \cup \overline{\overline{\mathcal{H}}}_5$, where $\delta_0 - \delta \leq 1 + \delta_0 - \gamma_0 + \eta < 1/2$, so $u_t(\delta - \delta_0)$ is asymptotically stationary. Let

$$S_T(\boldsymbol{\vartheta}) = U(\boldsymbol{\tau}) - r_T(\boldsymbol{\vartheta}), \tag{22}$$

where $U(\boldsymbol{\tau}) = E((\phi(L;\boldsymbol{\varphi})\Delta^{\delta-\delta_0}u_t)^2) - \sigma_0^2$ and

$$\begin{aligned} r_T(\boldsymbol{\vartheta}) = & \frac{1}{T}\sum_{t=1}^{T}\left((\phi(L;\boldsymbol{\varphi}_0)\{u_t\mathbb{I}(t>0)\})^2 - \sigma_0^2\right) \\ & - \frac{1}{T}\sum_{t=1}^{T}\left((\phi(L;\boldsymbol{\varphi})u_t(\delta-\delta_0))^2 - E\left((\phi(L;\boldsymbol{\varphi})\Delta^{\delta-\delta_0}u_t)^2\right)\right) \\ & - \frac{1}{T}\left(\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi}_0)(u_t\mathbb{I}(t>0))h_{t-1,T}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)\right)^2 \\ & + \frac{1}{T}\left(\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})u_t(\delta-\delta_0)h_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi})\right)^2 \\ & - \frac{2}{T}\sum_{t=1}^{T}d_t(\boldsymbol{\vartheta})s_t(\boldsymbol{\vartheta}) - \frac{1}{T}\sum_{t=1}^{T}d_t^2(\boldsymbol{\vartheta}). \end{aligned}$$

It follows that (19), (20), and (21) for $i = 5$ hold if we show that

$$\inf_{\|\boldsymbol{\tau}-\boldsymbol{\tau}_0\|\geq\varepsilon, \boldsymbol{\tau}\in\mathcal{T}_5} U(\boldsymbol{\tau}) > \epsilon, \tag{23}$$

$$\frac{1}{T}\sum_{t=1}^{T}\left((\phi(L;\boldsymbol{\varphi}_0)\{u_t\mathbb{I}(t>0)\})^2 - \sigma_0^2\right) = o_p(1), \tag{24}$$

$$\sup_{\|\boldsymbol{\tau}-\boldsymbol{\tau}_0\|\geq\varepsilon, \boldsymbol{\tau}\in\mathcal{T}_5}\frac{1}{T}\sum_{t=1}^{T}\left((\phi(L;\boldsymbol{\varphi})u_t(\delta-\delta_0))^2 - E\left((\phi(L;\boldsymbol{\varphi})\Delta^{\delta-\delta_0}u_t)^2\right)\right) = o_p(1), \tag{25}$$

$$\sup_{\mathcal{H}_5\cup\overline{\mathcal{H}}_5\cup\overline{\overline{\mathcal{H}}}_5}\frac{1}{T}\left|\sum_{t=1}^{T}d_t(\boldsymbol{\vartheta})s_t(\boldsymbol{\vartheta})\right| = o_p(1), \tag{26}$$

$$\sup_{\mathcal{H}_5\cup\overline{\mathcal{H}}_5\cup\overline{\overline{\mathcal{H}}}_5}\frac{1}{T}\left(\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})u_t(\delta-\delta_0)h_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi})\right)^2 = o_p(1). \tag{27}$$

First, (23), (24), and (25) follow by identical arguments to those in the proofs of (2.8) and (2.9) in Hualde and Robinson (2011). Next, by (139) of Lemma 17 with $\gamma_0 - \delta \leq 1 + \eta$ and $\delta_0 - \delta \leq \delta_0 - \gamma_0 + 1 + \eta$, the left-hand side of (26) is $O_p(T^{\max\{\theta,1+\delta_0-\gamma_0+\eta\}+2\theta-1/2+\eta})$, and by (134) of Lemma 16, the left-hand side of (27) is $O_p(T^{2\max\{\theta,1+\delta_0-\gamma_0+\eta\}-1})$. Both are $o_p(1)$ for θ and η sufficiently small, to conclude the proof of (19), (20), and (21) for $i = 5$.

Proof of (19) for $i = 1, \ldots, 4$ First we show (19) which, in view of Lemma 1 and that $d_t(\boldsymbol{\vartheta}_0) = 0$, holds if, for $i = 1, \ldots, 4$,

$$\Pr\left(\inf_{\overline{\overline{\mathcal{H}}}_i} \frac{1}{T}\sum_{t=1}^{T}\left(d_t(\boldsymbol{\vartheta}) + s_t(\boldsymbol{\vartheta})\right)^2 \leq \sigma_0^2 + \epsilon\right) \to 0 \text{ as } T \to \infty.$$

For $\delta \in \cup_{i=1}^4 \mathcal{I}_i$ it holds that $\gamma_0 - \delta \geq 1 + \eta$, so the probability above is bounded by

$$\begin{aligned}
&\Pr\left(\inf_{\overline{\overline{\mathcal{H}}}_i} \frac{T^{2(\gamma_0-\delta)-1}}{T} \inf_{\overline{\overline{\mathcal{H}}}_i} \frac{1}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}\left(d_t(\boldsymbol{\vartheta}) + s_t(\boldsymbol{\vartheta})\right)^2 \leq \sigma_0^2 + \epsilon\right)\\
&= \Pr\left(\inf_{\overline{\overline{\mathcal{H}}}_i} \frac{1}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}\left(d_t(\boldsymbol{\vartheta}) + s_t(\boldsymbol{\vartheta})\right)^2 \leq \frac{\sigma_0^2 + \epsilon}{T^{2\eta}}\right)\\
&\leq \Pr\left(\inf_{\overline{\overline{\mathcal{H}}}_i} \frac{1}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T} d_t^2(\boldsymbol{\vartheta}) - \sup_{\overline{\overline{\mathcal{H}}}_i} \frac{2}{T^{2(\gamma_0-\delta)-1}}\left|\sum_{t=1}^{T} d_t(\boldsymbol{\vartheta})\, s_t(\boldsymbol{\vartheta})\right| \leq \frac{\sigma_0^2 + \epsilon}{T^{2\eta}}\right).
\end{aligned}$$

Thus, (19) for $i = 1, \ldots, 4$ follows for θ small enough by (141) of Lemma 17, noting also that when $\delta \in \cup_{i=1}^4 \mathcal{I}_i$, $\delta_0 - \delta \geq \delta_0 - \gamma_0 + 1 + \eta$, and by Lemma 2.

Proof of (20) and (21) for $i = 4$ Fix ζ such that $0 < \zeta < \eta$ and let $\varkappa_4 = \gamma_0 - \delta - 1 - \zeta$, noting that $\varkappa_4 \geq \eta - \zeta > 0$ when $\delta \in \mathcal{I}_4$. Then, because $d_t(\boldsymbol{\vartheta}_0) = 0$, (20) holds if

$$\Pr\left(\inf_{\overline{\mathcal{H}}_4} \frac{T^{2\varkappa_4}}{T^{2(\gamma_0-\delta)-1}}\left(\sum_{t=1}^{T} d_t^2(\boldsymbol{\vartheta}) - 2\left|\sum_{t=1}^{T} d_t(\boldsymbol{\vartheta})\, s_t(\boldsymbol{\vartheta})\right| - \sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}_0)\right) \leq 0\right) \to 0 \tag{28}$$

as $T \to \infty$, noting the change in the normalization from (20) to (28), which is justified because the right-hand side of the inequality inside the probability in (20) is 0, so multiplying the left- and right-hand sides of the inequality by the same positive number does not alter the probability. Because $\sum_{t=1}^{T} d_t(\boldsymbol{\vartheta})\, c_{t-1}(\gamma - \delta, \boldsymbol{\varphi}) = 0$, it holds that

$$\sum_{t=1}^{T} d_t(\boldsymbol{\vartheta})\, s_t(\boldsymbol{\vartheta}) = \sum_{t=1}^{T} d_t(\boldsymbol{\vartheta})\, \phi(L;\boldsymbol{\varphi}) u_t(\delta - \delta_0), \tag{29}$$

where $\phi(L;\boldsymbol{\varphi}) u_t(\delta - \delta_0) = \sum_{j=0}^{t-1} c_j(\delta_0 - \delta, \boldsymbol{\varphi})\, u_{t-j}$. By the Cauchy-Schwarz inequality and (29), the probability in (28) is bounded by

$$\Pr\left(\inf_{\overline{\mathcal{H}}_4} \frac{T^{2\varkappa_4}}{T^{2(\gamma_0-\delta)-1}}\left(\sum_{t=1}^{T} d_t^2(\boldsymbol{\vartheta})\left(1 - 2v_T(\boldsymbol{\vartheta})\right) - \sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}_0)\right) \leq 0\right), \tag{30}$$

where $v_T(\boldsymbol{\vartheta}) = \left(\sum_{t=1}^{T}\left(\phi(L;\boldsymbol{\varphi}) u_t(\delta - \delta_0)\right)^2 / \sum_{t=1}^{T} d_t^2(\boldsymbol{\vartheta})\right)^{1/2}$. Then (28) holds if

$$\sup_{\overline{\mathcal{H}}_4} \frac{T^{2\varkappa_4}}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}_0) = o_p(1), \tag{31}$$

$$\Pr\left(\inf_{\overline{\mathcal{H}}_4} \frac{T^{2\varkappa_4}}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T} d_t^2(\boldsymbol{\vartheta})\left(1 - 2v_T(\boldsymbol{\vartheta})\right) \leq \epsilon\right) \to 0 \text{ as } T \to \infty. \tag{32}$$

First, given that $T^{2\varkappa_4-2(\gamma_0-\delta)+1}=T^{-1-2\zeta}$, (31) follows immediately by Lemma 1. Next, fixing c such that $0<c<1/2$, the probability in (32) equals

$$\begin{aligned}
&\Pr\left(\inf_{\overline{\mathcal{H}}_4}\frac{T^{2\varkappa_4}}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}d_t^2\left(\boldsymbol{\vartheta}\right)\left(1-2v_T\left(\boldsymbol{\vartheta}\right)\right)\leq\epsilon,\sup_{\overline{\mathcal{H}}_4}v_T\left(\boldsymbol{\vartheta}\right)\leq c\right)\\
&+\Pr\left(\inf_{\overline{\mathcal{H}}_4}\frac{T^{2\varkappa_4}}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}d_t^2\left(\boldsymbol{\vartheta}\right)\left(1-2v_T\left(\boldsymbol{\vartheta}\right)\right)\leq\epsilon,\sup_{\overline{\mathcal{H}}_4}v_T\left(\boldsymbol{\vartheta}\right)>c\right)\\
&\leq\Pr\left(\inf_{\overline{\mathcal{H}}_4}\frac{T^{2\varkappa_4}}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}d_t^2\left(\boldsymbol{\vartheta}\right)(1-2c)\leq\epsilon\right)+\Pr\left(\sup_{\overline{\mathcal{H}}_4}v_T\left(\boldsymbol{\vartheta}\right)>c\right),
\end{aligned}\tag{33}$$

so (32) holds on showing

$$\lim_{T\to\infty}\inf_{\overline{\mathcal{H}}_4}\frac{T^{2\varkappa_4}}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}d_t^2\left(\boldsymbol{\vartheta}\right)>\epsilon,\tag{34}$$

$$\sup_{\overline{\mathcal{H}}_4}v_T\left(\boldsymbol{\vartheta}\right)=o_p\left(1\right).\tag{35}$$

By the Cauchy-Schwarz inequality, $\sum_{t=1}^{T}d_t^2\left(\boldsymbol{\vartheta}\right)\geq T^{-1}\overline{d}_T^2\left(\boldsymbol{\vartheta}\right)$, where $\overline{d}_T\left(\boldsymbol{\vartheta}\right)=\sum_{t=1}^{T}d_t\left(\boldsymbol{\vartheta}\right)$, so that (34) holds by (114) of Lemma 3. To show (35), note that

$$\sup_{\overline{\mathcal{H}}_4}v_T\left(\boldsymbol{\vartheta}\right)\leq\left(\frac{\sup_{\overline{\mathcal{H}}_4}T^{-1-2\zeta}\sum_{t=1}^{T}\left(\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)\right)^2}{\inf_{\overline{\mathcal{H}}_4}T^{2\varkappa_4-(2(\gamma_0-\delta)-1)}\sum_{t=1}^{T}d_t^2\left(\boldsymbol{\vartheta}\right)}\right)^{1/2}\tag{36}$$

using $\varkappa_4=\gamma_0-\delta-1-\zeta$, where $\sup_{\overline{\mathcal{H}}_4}T^{-1-2\zeta}\sum_{t=1}^{T}\left(\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)\right)^2=o_p\left(1\right)$ by Lemma 15 because $\delta_0-\delta\leq1/2-\eta$. Then (36) is $o_p(1)$ by (34), which concludes the proof of (20) for $i=4$.

Next we show (21) for $i=4$. A potential problem here is that $\gamma=\gamma_0$ is admissible, so we cannot directly exploit the lower bound for the normalized $\sum_{t=1}^{T}d_t^2\left(\boldsymbol{\vartheta}\right)$ as in (34) because $d_t\left(\boldsymbol{\vartheta}\right)=0$ when $\gamma=\gamma_0$. However, we can instead take advantage of $|\gamma-\gamma_0|\leq\xi T^{-\varkappa_4}$ in $\mathcal{H}_4$ and apply the mean value theorem. First note that $\delta\in\mathcal{I}_4$ implies that $\delta_0-\delta\leq1/2-\eta$ and $\gamma_0-\delta\geq1+\eta$, so that $u_t\left(\delta-\delta_0\right)$ is asymptotically stationary as in the proof for $i=5$. Then, given (22), the result follows by (23), (24), (25) (whose proofs apply also for $\delta\in\mathcal{I}_4$), and showing also that

$$\sup_{\mathcal{H}_4}\frac{1}{T}\left(\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right)^2=o_p\left(1\right),\tag{37}$$

$$\sup_{\mathcal{H}_4}\frac{1}{T}\left|\sum_{t=1}^{T}d_t\left(\boldsymbol{\vartheta}\right)s_t\left(\boldsymbol{\vartheta}\right)\right|=o_p\left(1\right).\tag{38}$$

From (134) of Lemma 16, the left-hand side of (37) is $O_p(T^{-2\eta})=o_p\left(1\right)$ by choosing $\theta<1/2-\eta$. Next, because $|\gamma-\gamma_0|<\xi T^{-\varkappa_4}$ in $\mathcal{H}_4$, by (138) and (140) of Lemma 17 the left-hand side of (38) is $O_p\left(T^{\zeta-\eta+2\theta}\right)=o_p(1)$ for θ small enough because $\zeta<\eta$.

Proof of (20) and (21) for $i = 3$ Fix $\varkappa_3 = \gamma_0 - \delta - 1$, so noting that $\delta \in \mathcal{I}_3$, $\varkappa_3 \geq \gamma_0 - \delta_0 - 1/2 - \eta > 0$. Then, by the Cauchy-Schwarz inequality,

$$\Pr\left(\inf_{\overline{\mathcal{H}}_3} S_T(\boldsymbol{\vartheta}) \leq 0\right) \leq \Pr\left(\inf_{\overline{\mathcal{H}}_3} \frac{1}{T^2}\left(\overline{d}_T(\boldsymbol{\vartheta}) + \overline{s}_T(\boldsymbol{\vartheta})\right)^2 - \frac{1}{T}\sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}_0) \leq 0\right), \tag{39}$$

where $\overline{d}_T(\boldsymbol{\vartheta}) = \sum_{t=1}^{T} d_t(\boldsymbol{\vartheta})$ and

$$\begin{aligned}\overline{s}_T(\boldsymbol{\vartheta}) = \sum_{t=1}^{T} s_t(\boldsymbol{\vartheta}) =& \phi(L;\boldsymbol{\varphi})u_T(\delta - \delta_0 - 1) \\ &- \sum_{t=1}^{T} h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) \sum_{j=1}^{T} \phi(L;\boldsymbol{\varphi})u_j(\delta - \delta_0)\, h_{j-1,T}(\gamma - \delta, \boldsymbol{\varphi}).\end{aligned} \tag{40}$$

The right-hand side of (39) is thus bounded by

$$\Pr\left(\inf_{\overline{\mathcal{H}}_3} \frac{1}{T^2}\overline{d}_T^2(\boldsymbol{\vartheta})\left(1 - 2\left|\overline{v}_T(\boldsymbol{\vartheta})\right|\right) - \frac{1}{T}\sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}_0) \leq 0\right), \tag{41}$$

where $\overline{v}_T(\boldsymbol{\vartheta}) = \overline{s}_T(\boldsymbol{\vartheta})/\overline{d}_T(\boldsymbol{\vartheta})$. Applying Lemma 1, (20) for $i = 3$ would then hold if

$$\Pr\left(\inf_{\overline{\mathcal{H}}_3} \frac{1}{T^2}\overline{d}_T^2(\boldsymbol{\vartheta})\left(1 - 2\left|\overline{v}_T(\boldsymbol{\vartheta})\right|\right) \leq K\right) \to 0 \text{ as } T \to \infty, \tag{42}$$

for an arbitrarily large K. As in (33), fixing c such that $0 < c < 1/2$, the probability in (42) is bounded by

$$\Pr\left(\inf_{\overline{\mathcal{H}}_3} \frac{1}{T^2}\overline{d}_T^2(\boldsymbol{\vartheta})(1 - 2c) \leq K\right) + \Pr\left(\sup_{\overline{\mathcal{H}}_3}\left|\overline{v}_T(\boldsymbol{\vartheta})\right| > c\right), \tag{43}$$

so, as in (36), (42) holds if

$$\sup_{\overline{\mathcal{H}}_3} \frac{1}{T}\left|\overline{s}_T(\boldsymbol{\vartheta})\right| = O_p(1), \tag{44}$$

$$\lim_{T\to\infty} \inf_{\overline{\mathcal{H}}_3} \frac{1}{T^2}\overline{d}_T^2(\boldsymbol{\vartheta}) > K. \tag{45}$$

For $\delta \in \mathcal{I}_3$ it holds that $\delta_0 - \delta \leq 1/2$, so in view of (40) the proof of (44) is immediate using (131) in Lemma 14 together with Lemmas 15 and 16 with $\theta < 1/2$. Finally, given that $T^{\varkappa_3 - (\gamma_0 - \delta)} = T^{-1}$, the proof of (45) follows by Lemma 3, to conclude the proof of (20) for $i = 3$.

Next we show (21) for $i = 3$, which holds if

$$\Pr\left(\inf_{\mathcal{H}_3} \frac{1}{T}\left(\sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}) - 2\left|\sum_{t=1}^{T} d_t(\boldsymbol{\vartheta})\, s_t(\boldsymbol{\vartheta})\right| - \sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}_0)\right) \leq 0\right) \to 0 \text{ as } T \to \infty,$$

where

$$\sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}) = \sum_{t=1}^{T}\left(\phi(L;\boldsymbol{\varphi})\, u_t(\delta - \delta_0)\right)^2 - \left(\sum_{t=1}^{T} \phi(L;\boldsymbol{\varphi})\, u_t(\delta - \delta_0)\, h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi})\right)^2. \tag{46}$$

In the proof of their (2.7) for $i=3$, Hualde and Robinson (2011) showed that

$$\Pr\left(\inf_{\|\boldsymbol{\tau}-\boldsymbol{\tau}_0\|\geq\varepsilon,\boldsymbol{\tau}\in\mathcal{T}_3}\frac{1}{T}\sum_{t=1}^{T}\left(\phi\left(L;\boldsymbol{\varphi}\right)u_t\left(\delta-\delta_0\right)\right)^2>K\right)\to 1 \text{ as } T\to\infty \tag{47}$$

for any arbitrarily large fixed constant K (for small enough η). Thus, noting (46), (21) for $i=3$ holds by (47) and Lemma 1 on showing

$$\sup_{\mathcal{H}_3}\frac{1}{T}\left(\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right)^2=O_p\left(1\right), \tag{48}$$

$$\sup_{\mathcal{H}_3}\frac{1}{T}\left|\sum_{t=1}^{T}d_t\left(\boldsymbol{\vartheta}\right)s_t\left(\boldsymbol{\vartheta}\right)\right|=\xi O_p\left(1\right), \tag{49}$$

because, even if ξ had to be set large enough in the proof of (45) (see the proof of Lemma 3), this can be dominated by the constant K fixed in (47), which can be chosen arbitrarily large by setting η small enough. First, by (134) of Lemma 16, the left-hand side of (48) holds by choosing $\theta<1/2$ because $\delta_0-\delta\leq 1/2$ when $\delta\in\mathcal{I}_3$. Next, noting that $\sup_{\mathcal{H}_3}|\gamma-\gamma_0|\leq\xi T^{-\varkappa_3}$ and that $\delta\in\mathcal{I}_3$ implies $\gamma_0-\delta\geq\gamma_0-\delta_0+1/2-\eta>1$ and $\delta_0-\delta\leq 1/2$, it follows by (138) and (142) of Lemma 17 that the left-hand side of (49) is $\xi O_p\left(1\right)$ by choosing $\theta<1/2$.

Proof of (20) and (21) for $i=2$ Fix $\varkappa_2=\gamma_0-\delta_0-1/2>0$. Changing the normalization ($T^{2(\delta_0-\delta)}$ instead of T), by the Cauchy-Schwarz inequality as in (39), and proceeding as in (41), the left-hand side of (20) is bounded by

$$\Pr\left(\inf_{\overline{\mathcal{H}}_2}\frac{1}{T^{2(\delta_0-\delta)+1}}\overline{d}_T^2\left(\boldsymbol{\vartheta}\right)\left(1-2\left|\overline{v}_T\left(\boldsymbol{\vartheta}\right)\right|\right)-\sup_{\overline{\mathcal{H}}_2}\frac{1}{T^{2(\delta_0-\delta)}}\sum_{t=1}^{T}s_t^2\left(\boldsymbol{\vartheta}_0\right)\leq 0\right).$$

Then, given Lemma 1,

$$\sup_{\overline{\mathcal{H}}_2}\frac{1}{T^{2(\delta_0-\delta)}}\sum_{t=1}^{T}s_t^2\left(\boldsymbol{\vartheta}_0\right)=\sup_{\overline{\mathcal{H}}_2}\frac{T}{T^{2(\delta_0-\delta)}}\frac{1}{T}\sum_{t=1}^{T}s_t^2\left(\boldsymbol{\vartheta}_0\right)=O_p\left(1\right), \tag{50}$$

because when $\delta\in\mathcal{I}_2$, $\delta_0-\delta\geq 1/2$. Thus (20) for $i=2$ would hold if

$$\Pr\left(\inf_{\overline{\mathcal{H}}_2}\frac{1}{T^{2(\delta_0-\delta)+1}}\overline{d}_T^2\left(\boldsymbol{\vartheta}\right)\left(1-2\left|\overline{v}_T\left(\boldsymbol{\vartheta}\right)\right|\right)\leq K\right)\to 0 \text{ as } T\to\infty \tag{51}$$

for an arbitrarily large K, which, as in (43), follows if

$$\sup_{\overline{\mathcal{H}}_2}\frac{1}{T^{\delta_0-\delta+1/2}}\left|\overline{s}_T\left(\boldsymbol{\vartheta}\right)\right|=O_p\left(1\right), \tag{52}$$

$$\lim_{T\to\infty}\inf_{\overline{\mathcal{H}}_2}\frac{1}{T^{2(\delta_0-\delta)+1}}\overline{d}_T^2\left(\boldsymbol{\vartheta}\right)>K. \tag{53}$$

The proof of (52) is almost identical to that of (44), again applying Lemmas 14, 15, and 16. Also, given that $T^{\varkappa_2-(\gamma_0-\delta)}=T^{-(\delta_0-\delta)-1/2}$, (53) follows by Lemma 3, to conclude the proof of (20) for $i=2$.

Next we show (21) for $i=2$, which holds if

$$\Pr\left(\inf_{\mathcal{H}_2}\frac{1}{T^{2(\delta_0-\delta)}}\left(\sum_{t=1}^{T}s_t^2(\boldsymbol{\vartheta})-2\left|\sum_{t=1}^{T}d_t(\boldsymbol{\vartheta})\,s_t(\boldsymbol{\vartheta})\right|-\sum_{t=1}^{T}s_t^2(\boldsymbol{\vartheta}_0)\right)\leq 0\right)\to 0 \tag{54}$$

as $T\to\infty$. In the proof of their (2.7) for $i=2$, Hualde and Robinson (2011) showed that

$$\Pr\left(\inf_{\|\boldsymbol{\tau}-\boldsymbol{\tau}_0\|\geq\varepsilon,\boldsymbol{\tau}\in\mathcal{T}_2}\frac{1}{T^{2(\delta_0-\delta)}}\sum_{t=1}^{T}\left(\phi(L;\boldsymbol{\varphi})\,u_t(\delta-\delta_0)\right)^2>K\right)\to 1 \tag{55}$$

as $T\to\infty$ for any arbitrarily large fixed constant K (for small enough η). Thus, in view of (46), (50), (55), and Lemma 16 with $\theta<1/2$, it follows that (54) holds if

$$\sup_{\mathcal{H}_2}\frac{1}{T^{2(\delta_0-\delta)}}\left|\sum_{t=1}^{T}d_t(\boldsymbol{\vartheta})\,s_t(\boldsymbol{\vartheta})\right|=\xi O_p(1)\,. \tag{56}$$

Again, noting that $\sup_{\mathcal{H}_2}|\gamma-\gamma_0|\leq\xi T^{-\varkappa_2}$ and that $\delta\in\mathcal{I}_2$ implies $\gamma_0-\delta\geq\gamma_0-\delta_0+1/2>1$ and $\delta_0-\delta\geq 1/2$, (56) follows from (138) and (143) of Lemma 17 setting $\theta<1/2$.

Proof of (20) and (21) for $i=1$ Fix $\varkappa_1=\gamma_0-\delta_0-1/2>0$. As in the treatment of (28), (20) for $i=1$ holds if

$$\sup_{\mathcal{I}_1}\frac{T^{2\varkappa_1}}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}s_t^2(\boldsymbol{\vartheta}_0)=o_p(1)\,, \tag{57}$$

$$\Pr\left(\inf_{\overline{\mathcal{H}}_1}\frac{T^{2\varkappa_1}}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}d_t^2(\boldsymbol{\vartheta})\left(1-2v_T(\boldsymbol{\vartheta})\right)\leq\epsilon\right)\to 0 \text{ as } T\to\infty, \tag{58}$$

for an arbitrarily small ϵ. First, (57) holds by Lemma 1, noting that $2\varkappa_1-2(\gamma_0-\delta)+1=2(\delta-\delta_0)$ and $\sup_{\mathcal{I}_1}2(\delta-\delta_0)=-1-2\eta<-1$. Next, as in the proof of (51), see also (32) and (36), (58) follows if

$$\sup_{\overline{\mathcal{H}}_1}\frac{1}{T^{2(\delta_0-\delta)}}\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})\,u_t(\delta-\delta_0)^2=O_p(1)\,, \tag{59}$$

$$\lim_{T\to\infty}\inf_{\overline{\mathcal{H}}_1}\frac{1}{T^{2(\delta_0-\delta)}}\sum_{t=1}^{T}d_t^2(\boldsymbol{\vartheta})>K. \tag{60}$$

First, (59) follows immediately from (133) of Lemma 15, noting that $\delta_0-\delta\geq 1/2+\eta$. Next, by the Cauchy-Schwarz inequality, noting that $T^{\varkappa_1-(\gamma_0-\delta)}=T^{-(\delta_0-\delta)-1/2}$, (60) follows by Lemma 3.

Finally we show (21) for $i=1$, which holds if

$$\Pr\left(\inf_{\mathcal{H}_1}\frac{1}{T^{2(\delta_0-\delta)}}\left(\sum_{t=1}^{T}\left(d_t(\boldsymbol{\vartheta})+s_t(\boldsymbol{\vartheta})\right)^2-\sum_{t=1}^{T}s_t^2(\boldsymbol{\vartheta}_0)\right)\leq 0\right)\to 0 \tag{61}$$

as $T\to\infty$. By the Cauchy-Schwarz inequality the probability in (61) is bounded by

$$\Pr\left(\inf_{\mathcal{H}_1}\frac{1}{T^{2(\delta_0-\delta)+1}}\left(\overline{d}_T(\boldsymbol{\vartheta})+\overline{s}_T(\boldsymbol{\vartheta})\right)^2-\sup_{\mathcal{H}_1}\frac{1}{T^{2(\delta_0-\delta)}}\sum_{t=1}^{T}s_t^2(\boldsymbol{\vartheta}_0)\leq 0\right).$$

By (57) and the mean value theorem, (21) for $i = 1$ holds by showing

$$\Pr\left(\inf_{\mathcal{H}_1} \frac{1}{T^{2(\delta_0-\delta)+1}} \left(T^{\varkappa_1}(\gamma-\gamma_0)\frac{1}{T^{\varkappa_1}}\frac{\partial \overline{d}_T(\boldsymbol{\tau},\overline{\gamma})}{\partial\gamma} + \overline{s}_T(\boldsymbol{\vartheta})\right)^2 > \epsilon\right) \to 1 \tag{62}$$

as $T \to \infty$, where $|\overline{\gamma}-\gamma_0| \le |\gamma-\gamma_0|$. Note that in $\mathcal{H}_1$, $\gamma_0 - \delta \ge 1+\eta$, so there exists $\alpha > 0$ such that for T sufficiently large $\gamma - \delta \ge 1 + \alpha$. Let $g = T^{\varkappa_1}(\gamma-\gamma_0)$ and $\overline{s}_T(\boldsymbol{\tau},\gamma) = \overline{s}_T(\boldsymbol{\vartheta})$. Define the set $\mathcal{G}_1 = \{g, \boldsymbol{\tau}, \gamma_1, \gamma_2 : |g| < \xi, \boldsymbol{\tau} \in \mathcal{T}_1, \gamma_i - \delta \ge 1+\alpha, \gamma_i \in [\Box_1, \Box_2], i = 1, 2\}$. Then, noting that two different values of γ appear within the probability in (62), (21) for $i = 1$ holds if

$$\Pr\left(\inf_{\mathcal{G}_1}\left(g\frac{1}{T^{\varkappa_1+\delta_0-\delta+1/2}}\frac{\partial \overline{d}_T(\boldsymbol{\tau},\gamma_1)}{\partial\gamma} + \frac{1}{T^{\delta_0-\delta+1/2}}\overline{s}_T(\boldsymbol{\tau},\gamma_2)\right)^2 > \epsilon\right) \to 1 \tag{63}$$

as $T \to \infty$.

First, by (113) in Lemma 3, noting that $\varkappa_1 + \delta_0 - \delta + 1/2 = \gamma_0 - \delta$,

$$\frac{1}{T^{\gamma_0-\delta}}\frac{\partial \overline{d}_T(\boldsymbol{\tau},\gamma_1)}{\partial\gamma} = \mu_0\phi(1;\boldsymbol{\varphi})D(\delta,\gamma_1) + m_{1,T}(\boldsymbol{\tau},\gamma_1), \tag{64}$$

where

$$D(\delta,\gamma_1) = \frac{2(\gamma_1-\delta)^2 - 2(\gamma_1-\delta) + 1 - (\gamma_0-\delta)}{\Gamma(\gamma_0-\delta)(\gamma_1-\delta)^2(\gamma_0+\gamma_1-2\delta-1)^2}$$

and $\sup_{\mathcal{G}_1}|m_{1,T}(\boldsymbol{\tau},\gamma_1)| = o_p(1)$. Next, we analyze the behavior of $\overline{s}_T(\boldsymbol{\tau},\gamma_2)$ in (63). Note that for any $k \ge 1$ it holds that $\sum_{l=1}^{k}\phi(L;\boldsymbol{\varphi})u_l(\delta-\delta_0) = \phi(L;\boldsymbol{\varphi})u_k(\delta-\delta_0-1)$ and $\pi_{j+1}(d) - \pi_j(d) = \pi_{j+1}(d-1)$, and therefore

$$c_t(d,\boldsymbol{\varphi}) - c_{t-1}(d,\boldsymbol{\varphi}) = c_t(d-1,\boldsymbol{\varphi}). \tag{65}$$

By summation by parts and (65) we find

$$\begin{aligned}\sum_{t=1}^{T} c_{t-1}(\gamma-\delta,\boldsymbol{\varphi})\phi(L;\boldsymbol{\varphi})u_t(\delta-\delta_0) =& c_{T-1}(\gamma-\delta,\boldsymbol{\varphi})\phi(L;\boldsymbol{\varphi})u_T(\delta-\delta_0-1)\\ &- \sum_{t=1}^{T-1} c_t(\gamma-\delta-1,\boldsymbol{\varphi})\phi(L;\boldsymbol{\varphi})u_t(\delta-\delta_0-1).\end{aligned} \tag{66}$$

Also, by summation by parts on (10),

$$\begin{aligned}c_{t-1}(d,\boldsymbol{\varphi}) &= \phi(1;\boldsymbol{\varphi})\pi_{t-1}(d) - \pi_{t-1}(d)\sum_{k=t}^{\infty}\phi_k(\boldsymbol{\varphi}) - \sum_{k=0}^{t-2}\pi_{k+1}(d-1)\sum_{l=0}^{k}\phi_{t-1-l}(\boldsymbol{\varphi})\\ &= \phi(1;\boldsymbol{\varphi})\pi_{t-1}(d) + c_{2,t-1}(d,\boldsymbol{\varphi}) + c_{3,t-1}(d,\boldsymbol{\varphi}),\end{aligned} \tag{67}$$

where $c_{2,t-1}(d,\boldsymbol{\varphi}) = -\pi_{t-1}(d)\sum_{k=t}^{\infty}\phi_k(\boldsymbol{\varphi})$ and $c_{3,t-1}(d,\boldsymbol{\varphi}) = -\sum_{k=0}^{t-2}\pi_{k+1}(d-1)\sum_{l=0}^{k}\phi_{t-1-l}(\boldsymbol{\varphi})$. Thus,

$$\begin{aligned}\phi(L;\boldsymbol{\varphi})u_t(\delta-\delta_0-1) &= \sum_{j=0}^{t-1} c_j(\delta_0-\delta+1,\boldsymbol{\varphi})u_{t-j}\\ &= \phi(1;\boldsymbol{\varphi})\omega(1;\boldsymbol{\varphi}_0)\varepsilon_t(\delta-\delta_0-1) + p_t(\boldsymbol{\tau}),\end{aligned} \tag{68}$$

where

$$p_t(\boldsymbol{\tau}) = \phi(1;\boldsymbol{\varphi})\left(u_t(\delta-\delta_0-1) - \omega(1;\boldsymbol{\varphi}_0)\,\varepsilon_t(\delta-\delta_0-1)\right) + \sum_{j=0}^{t-1} c_{2,j}(\delta_0-\delta+1,\boldsymbol{\varphi})\,u_{t-j} + \sum_{j=0}^{t-1} c_{3,j}(\delta_0-\delta+1,\boldsymbol{\varphi})\,u_{t-j}. \tag{69}$$

Substituting (66)–(69) into $\overline{s}_T(\boldsymbol{\tau},\gamma_2)$, see (40), we get

$$\begin{aligned}\frac{1}{T^{\delta_0-\delta+1/2}}\overline{s}_T(\boldsymbol{\tau},\gamma_2) =& \frac{1}{T^{\delta_0-\delta+1/2}}\phi(L;\boldsymbol{\varphi})\,u_T(\delta-\delta_0-1)\\ &- \frac{c_{T-1}(\gamma_2-\delta,\boldsymbol{\varphi})\,\phi(L;\boldsymbol{\varphi})\,u_T(\delta-\delta_0-1)\sum_{t=1}^{T} c_{t-1}(\gamma_2-\delta,\boldsymbol{\varphi})}{T^{\delta_0-\delta+1/2}\sum_{j=1}^{T} c_{j-1}^2(\gamma_2-\delta,\boldsymbol{\varphi})}\\ &+ \frac{\sum_{t=1}^{T} c_{t-1}(\gamma_2-\delta,\boldsymbol{\varphi})\sum_{j=1}^{T-1} c_j(\gamma_2-\delta-1,\boldsymbol{\varphi})\,\phi(L;\boldsymbol{\varphi})\,u_j(\delta-\delta_0-1)}{T^{\delta_0-\delta+1/2}\sum_{j=1}^{T} c_{j-1}^2(\gamma_2-\delta,\boldsymbol{\varphi})}\\ =& \phi(1;\boldsymbol{\varphi})\,V_T(\delta,\gamma_2) + m_{2,T}(\boldsymbol{\tau},\gamma_2),\end{aligned} \tag{70}$$

where

$$\begin{aligned}V_T(\delta,\gamma_2) =& \frac{\omega(1;\boldsymbol{\varphi}_0)}{T^{\delta_0-\delta+1/2}}\left(1 - \frac{\pi_{T-1}(\gamma_2-\delta)\sum_{t=1}^{T}\pi_{t-1}(\gamma_2-\delta)}{\sum_{j=1}^{T}\pi_{j-1}^2(\gamma_2-\delta)}\right)\varepsilon_T(\delta-\delta_0-1)\\ &+ \frac{\omega(1;\boldsymbol{\varphi}_0)}{T^{\delta_0-\delta+1/2}}\frac{\sum_{t=1}^{T}\pi_{t-1}(\gamma_2-\delta)\sum_{j=1}^{T-1}\pi_j(\gamma_2-\delta-1)\,\varepsilon_j(\delta-\delta_0-1)}{\sum_{j=1}^{T}\pi_{j-1}^2(\gamma_2-\delta)}.\end{aligned}$$

and $m_{2,T}(\boldsymbol{\tau},\gamma_2)$ collects remainder terms arising from (67) and (69). Then, by relatively straightforward arguments, it can be shown that

$$\sup_{\mathcal{G}_1}|m_{2,T}(\boldsymbol{\tau},\gamma_2)| = o_p(1). \tag{71}$$

The proof of (71) involves several results. First, in order to deal with the first term on the right-hand side of (69), note that $u_t = \omega(1;\boldsymbol{\varphi}_0)\,\varepsilon_t + \widetilde{\varepsilon}_{t-1} - \widetilde{\varepsilon}_t$, where $\widetilde{\varepsilon}_t = \sum_{j=0}^{\infty}\widetilde{\omega}_j(\boldsymbol{\varphi}_0)\,\varepsilon_{t-j}$, $\widetilde{\omega}_j(\boldsymbol{\varphi}_0) = \sum_{k=j+1}^{\infty}\omega_k(\boldsymbol{\varphi}_0)$. By Assumptions A1 and A2, $\widetilde{\varepsilon}_t$ is well defined in the mean-square sense and $|\widetilde{\omega}_j(\boldsymbol{\varphi}_0)| = O(j^{-\varsigma})$. We also apply Lemma 11, which immediately leads to the result that for any $c>0$, $\sup_{d\geq 1/2+c}\pi_t(d)/T^d = O(t^{-1/2+c}/T^{1/2+c})$, as well as Lemmas 14, 15, and 9.

Thus, noting (8), (64), (70), Assumption A3, and defining $\mathcal{L}_1 = \{g,\delta,\gamma_1,\gamma_2 : |g|\leq\xi, \delta\in\mathcal{I}_1, \gamma_i-\delta\geq 1+\alpha, \gamma_i\in[\Box_1,\Box_2], i=1,2\}$, (63) holds if

$$\Pr\left(\inf_{\mathcal{L}_1}\left(\mu_0 g D(\delta,\gamma_1) + V_T(\delta,\gamma_2)\right)^2 > \epsilon\right) \to 1 \text{ as } T\to\infty. \tag{72}$$

Considering $V_T(\delta,\gamma)$ as a continuous process indexed by (δ,γ), we next show that

$$V_T(\delta,\gamma) \Rightarrow V(\delta,\gamma), \tag{73}$$

where $\Rightarrow$ means weak convergence in the space of continuous functions on $\mathcal{L} = \{\delta, \gamma : \delta \in \mathcal{I}_1, \gamma - \delta \geq 1 + \alpha, \gamma \in [\Box_1, \Box_2]\}$ endowed with the uniform topology and

$$V(\delta,\gamma) = \omega(1;\boldsymbol{\varphi}_0)\frac{1-(\gamma-\delta)}{\gamma-\delta}W(1;1+\delta_0-\delta) + \omega(1;\boldsymbol{\varphi}_0)\frac{2(\gamma-\delta)-1}{\gamma-\delta}\int_0^1 r^{\gamma-\delta-2}W(r;1+\delta_0-\delta)\,dr,$$

with $W(r;d) = \Gamma(d)^{-1}\int_0^r (1-s)^{d-1}\,dB(s)$ and $B(s)$ denoting fractional (type II) and regular scalar Brownian motions, respectively, both with variance σ_0^2. Convergence of the finite-dimensional distributions follows by Theorem 1 of Hosoya (2005) (noting that our Assumption A2 implies conditions A(i), A(ii) and A(iii) in Hosoya (2005)) and the continuous mapping theorem (see (A.1) in Robinson and Iacone (2005), for a very similar derivation). Tightness of the process $V_T(\delta,\gamma)$ on the compact set $\mathcal{L} \subseteq \mathbb{R}^2$ follows from Lemmas A.2 and C.3 of Johansen and Nielsen (2010) noting, in particular, that $V_T(\delta,\gamma)$ is continuously differentiable for $\gamma - \delta \geq 1 + \alpha$, which proves (73).

It follows from (73) and the continuous mapping theorem that, as $T \to \infty$,

$$\inf_{\mathcal{L}_1}(\mu_0 gD(\delta,\gamma_1) + V_T(\delta,\gamma_2))^2 \to_d \inf_{\mathcal{L}_1}(\mu_0 gD(\delta,\gamma_1) + V(\delta,\gamma_2))^2,$$

where the right-hand side is a.s. positive because the quantity whose infimum is taken is the square of a Gaussian random variable. Thus, as $T \to \infty$,

$$\Pr\left(\inf_{\mathcal{L}_1}(\mu_0 gD(\delta,\gamma_1) + V_T(\delta,\gamma_2))^2 > \epsilon\right) \to \Pr\left(\inf_{\mathcal{L}_1}(\mu_0 gD(\delta,\gamma_1) + V(\delta,\gamma_2))^2 > \epsilon\right),$$

and (72) follows because ϵ is arbitrarily small, to conclude the proof of (21) for $i = 1$ and therefore that for (15).

A.1.3 Proof of (16)

Here, let $R_T(\boldsymbol{\tau},\gamma) = R_T(\boldsymbol{\vartheta})$, $d_t(\boldsymbol{\tau},\gamma) = d_t(\boldsymbol{\vartheta})$, and $s_t(\boldsymbol{\tau},\gamma) = s_t(\boldsymbol{\vartheta}) = s_{1t}(\boldsymbol{\tau}) - s_{2t}(\boldsymbol{\vartheta})$ with $s_{1t}(\boldsymbol{\tau}) = \phi(L;\boldsymbol{\varphi})u_t(\delta-\delta_0)$ and $s_{2t}(\boldsymbol{\tau},\gamma) = s_{2t}(\boldsymbol{\vartheta}) = h_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi})\sum_{j=1}^T s_{1j}(\boldsymbol{\tau})h_{j-1,T}(\gamma-\delta,\boldsymbol{\varphi})$, so that, noting $\sum_{t=1}^T h_{t-1,T}^2(\gamma-\delta,\boldsymbol{\varphi}) = 1$,

$$\sum_{t=1}^T s_{2t}^2(\boldsymbol{\vartheta}) = \sum_{t=1}^T s_{1t}(\boldsymbol{\tau})s_{2t}(\boldsymbol{\vartheta}) = \left(\sum_{j=1}^T s_{1j}(\boldsymbol{\tau})h_{j-1,T}(\gamma-\delta,\boldsymbol{\varphi})\right)^2.$$

Noting also (29),

$$R_T(\boldsymbol{\vartheta}) = \frac{1}{T}\sum_{t=1}^T d_t^2(\boldsymbol{\vartheta}) + \frac{1}{T}\sum_{t=1}^T s_{1t}^2(\boldsymbol{\tau}) - \frac{1}{T}\left(\sum_{j=1}^T s_{1j}(\boldsymbol{\tau})h_{j-1,T}(\gamma-\delta,\boldsymbol{\varphi})\right)^2 + \frac{2}{T}\sum_{t=1}^T d_t(\boldsymbol{\vartheta})s_{1t}(\boldsymbol{\vartheta}). \tag{74}$$

Clearly, if $\widehat{\boldsymbol{\vartheta}} \in \overline{N}_\varepsilon \cap M_\varepsilon$, then $\inf_{\overline{N}_\varepsilon \cap M_\varepsilon} R_T(\widehat{\boldsymbol{\tau}}, \gamma) \leq R_T(\widehat{\boldsymbol{\tau}}, \gamma_0)$, so that

$$\Pr(\widehat{\boldsymbol{\vartheta}} \in \overline{N}_\varepsilon \cap M_\varepsilon) \leq \Pr\left(\widehat{\boldsymbol{\vartheta}} \in \overline{N}_\varepsilon \cap M_\varepsilon, \inf_{\overline{N}_\varepsilon \cap M_\varepsilon} R_T(\widehat{\boldsymbol{\tau}}, \gamma) - R_T(\widehat{\boldsymbol{\tau}}, \gamma_0) \leq 0\right). \tag{75}$$

Recalling that $d_t(\boldsymbol{\tau}, \gamma_0) = 0$, $R_T(\widehat{\boldsymbol{\tau}}, \gamma_0) = T^{-1}\sum_{t=1}^{T} s_{1t}(\widehat{\boldsymbol{\tau}})$ and this cancels with the corresponding term in $R_T(\widehat{\boldsymbol{\tau}}, \gamma)$, see (74). Thus, (16) holds if

$$\lim_{T\to\infty} \inf_{\boldsymbol{\vartheta}\in \overline{N}_\varepsilon \cap M_\varepsilon} \frac{1}{T^{2(\gamma_0-\delta)-1}} \sum_{t=1}^{T} d_t^2(\boldsymbol{\vartheta}) > \epsilon, \tag{76}$$

$$\sup_{\boldsymbol{\vartheta}\in \overline{N}_\varepsilon \cap M_\varepsilon} \frac{1}{T^{2(\gamma_0-\delta)-1}} \left|\sum_{t=1}^{T} d_t(\boldsymbol{\vartheta}) s_{1t}(\boldsymbol{\tau})\right| = o_p(1), \tag{77}$$

$$\sup_{\boldsymbol{\vartheta}\in \overline{N}_\varepsilon \cap M_\varepsilon} \frac{1}{T^{2(\gamma_0-\delta)-1}} \left(\sum_{j=1}^{T} s_{1j}(\boldsymbol{\tau}) h_{j-1,T}(\gamma-\delta, \boldsymbol{\varphi})\right)^2 = o_p(1), \tag{78}$$

noting the change in the normalization compared with (75) ($T^{2(\gamma_0-\delta_0)-1}$ instead of T), which is justified because the right-hand side of the inequality inside the probability in (75) is 0, so multiplying the left- and right-hand sides of the inequality by a positive number does not alter the probability.

First, (76) follows from Lemma 2, noting that in $\overline{N}_\varepsilon \cap M_\varepsilon$, $\gamma_0 - \delta \geq \gamma_0 - \delta_0 - \varepsilon > 1/2$ setting ε small enough. Next, letting both ε and θ be sufficiently small and noting that in $\overline{N}_\varepsilon \cap M_\varepsilon$, $\delta_0 - \delta \geq -\varepsilon$, by (141) of Lemma 17 the left-hand side of (77) is $O_p(T^{1/2+\delta_0-\gamma_0+3\theta+\varepsilon}) = o_p(1)$. Finally, by (135) of Lemma 16 the left-hand side of (78) is $O_p(T^{-2(\gamma_0-\delta_0-1/2-\theta-\varepsilon)}) = o_p(1)$, to conclude the proof of (16) and therefore that of consistency of $\widehat{\boldsymbol{\vartheta}}$.

A.2 Proof of Theorem 1(ii): the $\gamma_0 - 1/2 < \delta_0$ case

Clearly

$$\Pr\left(\|\widehat{\boldsymbol{\tau}} - \boldsymbol{\tau}_0\| \geq \varepsilon\right) = \Pr\left(\inf_{\boldsymbol{\vartheta}\in \overline{M}_\varepsilon} R_T(\boldsymbol{\vartheta}) \leq \inf_{\boldsymbol{\vartheta}\in M_\varepsilon} R_T(\boldsymbol{\vartheta})\right),$$

so, as in the proof for $\gamma_0 - 1/2 > \delta_0$, the result follows by showing that the right-hand side of (18) is $o(1)$, which, in view of Lemma 1, holds if

$$\Pr\left(\inf_{\boldsymbol{\vartheta}\in \overline{M}_\varepsilon} R_T(\boldsymbol{\vartheta}) \leq \sigma_0^2 + \epsilon\right) \to 0 \text{ as } T \to \infty.$$

This result is given in Lemma 4, whose proof uses very similar techniques to those employed in the proof of (15). This completes the proof of Theorem 1.

A.3 Proof of Theorem 2(i): the $\gamma_0 - 1/2 > \delta_0$ case

We first show that

$$T^{1/2}\boldsymbol{M}_T^{-1}(\widehat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0) \to_d N\left(0, \sigma_0^2 \boldsymbol{V}^{-1}\right). \tag{79}$$

By the mean value theorem,

$$\widehat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0 = -\left(\frac{\partial^2 R_T(\overline{\boldsymbol{\vartheta}})}{\partial\boldsymbol{\vartheta}\partial\boldsymbol{\vartheta}'}\right)^{-1} \frac{\partial R_T(\boldsymbol{\vartheta}_0)}{\partial\boldsymbol{\vartheta}}, \tag{80}$$

where $\overline{\boldsymbol{\vartheta}}$ represents an intermediate point which is allowed to vary across the different rows of $\partial^2 R_T(\cdot)/\partial\boldsymbol{\vartheta}\partial\boldsymbol{\vartheta}'$.

We first analyze the score in (80). It can be easily seen that $\partial d_t(\boldsymbol{\vartheta}_0)/\partial\boldsymbol{\tau} = 0$ and $\partial s_{1t}(\boldsymbol{\tau})/\partial\gamma = 0$, so, recalling that $d_t(\boldsymbol{\vartheta}_0) = 0$ and the decomposition (74),

$$\frac{\partial R_T(\boldsymbol{\vartheta}_0)}{\partial\boldsymbol{\vartheta}} = \frac{2}{T}\sum_{t=1}^{T} s_t(\boldsymbol{\vartheta}_0)\left[\begin{pmatrix} \frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial\boldsymbol{\tau}} \\ \frac{\partial d_t(\boldsymbol{\vartheta}_0)}{\partial\gamma} \end{pmatrix} - \frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial\boldsymbol{\vartheta}}\right].$$

Then, by Lemma 5(a) it holds that

$$\frac{T^{1/2}}{2}\boldsymbol{M}_T\frac{\partial R_T(\boldsymbol{\vartheta}_0)}{\partial\boldsymbol{\vartheta}} = \begin{pmatrix} T^{-1/2}I_{p+1} & 0 \\ 0 & T^{1/2-(\gamma_0-\delta_0)} \end{pmatrix}\sum_{t=1}^{T}\varepsilon_t\begin{pmatrix} \frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial\boldsymbol{\tau}} \\ \frac{\partial d_t(\boldsymbol{\vartheta}_0)}{\partial\gamma} \end{pmatrix} + o_p(1). \qquad (81)$$

Next, as in (2.54) of Hualde and Robinson (2011),

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}\varepsilon_t\frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial\boldsymbol{\tau}} = \frac{1}{T^{1/2}}\sum_{t=2}^{T}\varepsilon_t\sum_{j=1}^{\infty}\boldsymbol{m}_j(\boldsymbol{\varphi}_0)\varepsilon_{t-j} + o_p(1),$$

where $\boldsymbol{m}_j(\boldsymbol{\varphi}_0) = (-j^{-1}, \boldsymbol{b}_j'(\boldsymbol{\varphi}_0))'$. Also,

$$\begin{aligned}\frac{\partial d_t(\boldsymbol{\vartheta}_0)}{\partial\gamma} =& -\mu_0 c_{t-1}^{(1)}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0) \\ &+ \mu_0 c_{t-1}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)\frac{\sum_{j=1}^{T} c_{j-1}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)\, c_{j-1}^{(1)}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)}{\sum_{j=1}^{T} c_{j-1}^2(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)},\end{aligned}$$

where $c_t^{(1)}(\cdot,\cdot)$ is the derivative of $c_t(\cdot,\cdot)$ with respect to the first argument, so that

$$\begin{aligned}&\frac{1}{T^{\gamma_0-\delta_0-1/2}}\sum_{t=1}^{T}\varepsilon_t\frac{\partial d_t(\boldsymbol{\vartheta}_0)}{\partial\gamma} \\ &=\frac{\mu_0}{T^{\gamma_0-\delta_0-1/2}}\sum_{t=2}^{T}\varepsilon_t c_{t-1}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)\frac{\sum_{j=1}^{T} c_{j-1}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)\, c_{j-1}^{(1)}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)}{\sum_{j=1}^{T} c_{j-1}^2(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)} \\ &\quad - \frac{\mu_0}{T^{\gamma_0-\delta_0-1/2}}\sum_{t=2}^{T}\varepsilon_t c_{t-1}^{(1)}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0). \qquad (82)\end{aligned}$$

By (10), $c_{t-1}^{(1)}(d,\boldsymbol{\varphi}) = \sum_{j=0}^{t-1}\phi_j(\boldsymbol{\varphi})\pi_{t-j-1}^{(1)}(d)$, where $\pi_j^{(1)}(\cdot)$ is the first derivative of $\pi_j(\cdot)$ and

$$\pi_j^{(1)}(d) = (\psi(d+j) - \psi(d))\,\pi_j(d), \qquad (83)$$

with $\psi(\cdot)$ denoting the digamma function. Then, noting that $\gamma_0 - \delta_0 > 1/2$, by a similar analysis to that in the proof of Lemma 13, the right-hand side of (82) equals

$$\begin{aligned}&\frac{\mu_0\phi(1;\boldsymbol{\varphi}_0)}{T^{\gamma_0-\delta_0-1/2}}\frac{\sum_{t=2}^{T}\varepsilon_t\pi_{t-1}(\gamma_0-\delta_0)\sum_{j=1}^{T}\pi_{j-1}(\gamma_0-\delta_0)\,\pi_{j-1}^{(1)}(\gamma_0-\delta_0)}{\sum_{j=1}^{T}\pi_{j-1}^2(\gamma_0-\delta_0)} \\ &- \frac{\mu_0\phi(1;\boldsymbol{\varphi}_0)}{T^{\gamma_0-\delta_0-1/2}}\sum_{t=2}^{T}\varepsilon_t\pi_{t-1}^{(1)}(\gamma_0-\delta_0) + o_p(1). \qquad (84)\end{aligned}$$

Substituting (83) (evaluated at $\gamma_0-\delta_0$) into (84), the first two terms of (84) become

$$\frac{\mu_0\phi\left(1;\boldsymbol{\varphi}_0\right)}{T^{\gamma_0-\delta_0-1/2}}\frac{\sum_{t=2}^{T}\varepsilon_t\pi_{t-1}\left(\gamma_0-\delta_0\right)\sum_{j=1}^{T}\psi\left(\gamma_0-\delta_0+j-1\right)\pi_{j-1}^2\left(\gamma_0-\delta_0\right)}{\sum_{j=1}^{T}\pi_{j-1}^2\left(\gamma_0-\delta_0\right)}$$
$$-\frac{\mu_0\phi\left(1;\boldsymbol{\varphi}_0\right)}{T^{\gamma_0-\delta_0-1/2}}\sum_{t=2}^{T}\varepsilon_t\psi\left(\gamma_0-\delta_0+t-1\right)\pi_{t-1}\left(\gamma_0-\delta_0\right). \tag{85}$$

By the properties of the digamma function it can be shown that (85) is

$$\frac{\mu_0\phi\left(1;\boldsymbol{\varphi}_0\right)}{T^{\gamma_0-\delta_0-1/2}}\frac{\sum_{t=2}^{T}\varepsilon_t\pi_{t-1}\left(\gamma_0-\delta_0\right)\sum_{j=1}^{T}\log\left(\gamma_0-\delta_0+j-1\right)\pi_{j-1}^2\left(\gamma_0-\delta_0\right)}{\sum_{j=1}^{T}\pi_{j-1}^2\left(\gamma_0-\delta_0\right)}$$
$$-\frac{\mu_0\phi\left(1;\boldsymbol{\varphi}_0\right)}{T^{\gamma_0-\delta_0-1/2}}\sum_{t=2}^{T}\varepsilon_t\log\left(\gamma_0-\delta_0+t-1\right)\pi_{t-1}\left(\gamma_0-\delta_0\right)+o_p\left(1\right), \tag{86}$$

where the first two terms in (86) equal $\mu_0\phi\left(1;\boldsymbol{\varphi}_0\right)T^{1/2+\delta_0-\gamma_0}\sum_{t=2}^{T}\varepsilon_t g_{t,T}\left(\gamma_0-\delta_0\right)$ with

$$g_{t,T}\left(d\right)=\frac{\pi_{t-1}\left(d\right)\sum_{j=1}^{T}\log(\frac{d+j-1}{T})\pi_{j-1}^2\left(d\right)-\log(\frac{d+t-1}{T})\pi_{t-1}\left(d\right)\sum_{j=1}^{T}\pi_{j-1}^2\left(d\right)}{\sum_{j=1}^{T}\pi_{j-1}^2\left(d\right)}.$$

Collecting these terms shows that

$$\frac{T^{1/2}}{2}\boldsymbol{M}_T\frac{\partial R_T\left(\boldsymbol{\vartheta}_0\right)}{\partial\boldsymbol{\vartheta}}=\sum_{t=2}^{T}\varepsilon_t\boldsymbol{\eta}_{t,T}+o_p\left(1\right), \tag{87}$$

where

$$\boldsymbol{\eta}_{t,T}=\begin{pmatrix}\frac{1}{T^{1/2}}\sum_{j=1}^{\infty}\boldsymbol{m}_j\left(\boldsymbol{\varphi}_0\right)\varepsilon_{t-j}\\ \frac{1}{T^{\gamma_0-\delta_0-1/2}}\mu_0\phi\left(1;\boldsymbol{\varphi}_0\right)g_{t,T}\left(\gamma_0-\delta_0\right)\end{pmatrix}.$$

Defining $\mathcal{F}_{t,T}=\mathcal{F}_t$ for any $1\leq t\leq T$, Assumption A2 implies that $\{\varepsilon_t\boldsymbol{\eta}_{t,T},\mathcal{F}_{t,T},1\leq t\leq T,T\geq 1\}$ is a martingale difference array. For any $(p+2)$-dimensional vector $\boldsymbol{\xi}$, define $\xi_{t,T}=\varepsilon_t\boldsymbol{\xi}'\boldsymbol{\eta}_{t,T}/\sigma_0(\boldsymbol{\xi}'\boldsymbol{V}\boldsymbol{\xi})^{1/2}$ and $B_T^2=\sum_{t=2}^{T}E(\xi_{t,T}^2|\mathcal{F}_{t-1,T})$. Then, by Corollary 3.1 of Hall and Heyde (1980), if

$$B_T^2\to_p 1, \tag{88}$$

and, for all $\epsilon>0$,

$$\sum_{t=2}^{T}E\left(\xi_{t,T}^2\mathbb{I}\left(\left|\xi_{t,T}\right|>\epsilon\right)|\mathcal{F}_{t-1,T}\right)\to_p 0, \tag{89}$$

it holds that $\sum_{t=2}^{T}\xi_{t,T}\to_d N\left(0,1\right)$, and hence

$$\sum_{t=2}^{T}\varepsilon_t\boldsymbol{\eta}_{t,T}\to_d N\left(0,\sigma_0^2\boldsymbol{V}\right) \tag{90}$$

by direct application of the Cramer-Wold device. First we note that

$$E\left(\xi_{t,T}^2|\mathcal{F}_{t-1,T}\right)=\frac{\boldsymbol{\xi}'\boldsymbol{\eta}_{t,T}\boldsymbol{\eta}_{t,T}'\boldsymbol{\xi}}{\boldsymbol{\xi}'\boldsymbol{V}\boldsymbol{\xi}},$$

so that (88) holds if $\sum_{t=2}^{T} \boldsymbol{\eta}_{t,T}\boldsymbol{\eta}_{t,T}' \to_p \boldsymbol{V}$. However, this follows straightforwardly by the same arguments as in the proof of (2.55) of Hualde and Robinson (2011) and Lemma 10 because $\sum_{l=1}^{t}\sum_{j=1}^{\infty} \boldsymbol{m}_j(\boldsymbol{\varphi}_0)\varepsilon_{l-j} = O_p(t^{1/2})$, which implies, by summation by parts, that

$$\frac{1}{T^{\gamma_0-\delta_0}}\sum_{t=2}^{T} g_{t,T}(\gamma_0-\delta_0)\sum_{j=1}^{\infty}\boldsymbol{m}_j(\boldsymbol{\varphi}_0)\varepsilon_{t-j} = o_p(1). \tag{91}$$

Now (89) holds if, e.g., $\sum_{t=2}^{T} E(\xi_{t,T}^4|\mathcal{F}_{t-1,T}) \to_p 0$, which, given that the fourth moment of ε_t is finite, holds if $\sum_{t=2}^{T}(\boldsymbol{\xi}'\boldsymbol{\eta}_{t,T}\boldsymbol{\eta}_{t,T}'\boldsymbol{\xi})^2 \to_p 0$, and this can be easily justified by previous arguments. This completes the proof of (90).

Next, noting (80), (87), and (90), the proof of (79) is completed by showing

$$\boldsymbol{M}_T\left(\frac{\partial^2 R_T(\overline{\boldsymbol{\vartheta}})}{\partial\boldsymbol{\vartheta}\partial\boldsymbol{\vartheta}'} - \frac{\partial^2 R_T(\boldsymbol{\vartheta}_0)}{\partial\boldsymbol{\vartheta}\partial\boldsymbol{\vartheta}'}\right)\boldsymbol{M}_T = o_p(1) \tag{92}$$

and

$$\frac{1}{2}\boldsymbol{M}_T\frac{\partial^2 R_T(\boldsymbol{\vartheta}_0)}{\partial\boldsymbol{\vartheta}\partial\boldsymbol{\vartheta}'}\boldsymbol{M}_T \to_p \boldsymbol{V}. \tag{93}$$

By Lemma 6 it holds that, for some fixed $\varkappa > 0$, $T^{\varkappa}(\widehat{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0) \to_p 0$, and in light of this the proof of (92) is relatively straightforward. It consists of deriving all terms in $\partial^2 R_T(\overline{\boldsymbol{\vartheta}})/\partial\boldsymbol{\vartheta}\partial\boldsymbol{\vartheta}'$ and checking that the differences with respect the corresponding ones in $\partial^2 R_T(\boldsymbol{\vartheta}_0)/\partial\boldsymbol{\vartheta}\partial\boldsymbol{\vartheta}'$ satisfy (92). This requires the use of the mean value theorem and Assumption A4(ii), where typically the derivatives involve additional $\log T$ factors which are compensated by the factor $T^{-\varkappa}$ that arises because $\overline{\boldsymbol{\vartheta}} - \boldsymbol{\vartheta}_0 = O_p(T^{-\varkappa})$.

Now we show (93). Recalling $d_t(\boldsymbol{\vartheta}_0) = 0$ and noting $\partial^2 d_t(\boldsymbol{\vartheta}_0)/\partial\boldsymbol{\tau}\partial\boldsymbol{\tau}' = 0$, Lemma 5(b) implies that

$$\frac{1}{2}\boldsymbol{M}_T\frac{\partial^2 R_T(\boldsymbol{\vartheta}_0)}{\partial\boldsymbol{\vartheta}\partial\boldsymbol{\vartheta}'}\boldsymbol{M}_T = \boldsymbol{M}_T\frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix} \frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial\boldsymbol{\tau}}\frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial\boldsymbol{\tau}'} & 0 \\ 0 & \left(\frac{\partial d_t(\boldsymbol{\vartheta}_0)}{\partial\gamma}\right)^2 \end{pmatrix}\boldsymbol{M}_T + o_p(1), \tag{94}$$

so (93) holds by showing that, as $T \to \infty$,

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial\boldsymbol{\tau}}\frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial\boldsymbol{\tau}'} \to_p \sigma_0^2\boldsymbol{A}, \tag{95}$$

$$\frac{1}{T^{2(\gamma_0-\delta_0)-1}}\sum_{t=1}^{T}\left(\frac{\partial d_t(\boldsymbol{\vartheta}_0)}{\partial\gamma}\right)^2 \to \frac{\mu_0^2\phi^2(1;\boldsymbol{\varphi}_0)}{\Gamma^2(\gamma_0-\delta_0)(2(\gamma_0-\delta_0)-1)^3}. \tag{96}$$

Here, (95) follows from (2.53) of Hualde and Robinson (2011) and (96) follows by arguments used in the proof of (88).

Next, given (79), the remaining part of (13) is justified as follows. Noting

$$\phi(L;\boldsymbol{\varphi})x_t(\delta) = \mu_0 c_{t-1}(\gamma_0-\delta,\boldsymbol{\varphi}) + \phi(L;\boldsymbol{\varphi})u_t(\delta-\delta_0), \tag{97}$$

it follows from (11) that

$$\widehat{\mu} = \widehat{\mu}(\widehat{\boldsymbol{\vartheta}}) = \mu_0\sum_{t=1}^{T} c_{t-1}(\gamma_0-\widehat{\delta},\widehat{\boldsymbol{\varphi}})k_{t-1,T}(\widehat{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}}) + \sum_{t=1}^{T}\phi(L;\widehat{\boldsymbol{\varphi}})u_t(\widehat{\delta}-\delta_0)k_{t-1,T}(\widehat{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}}),$$

where $k_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi}) = c_{t-1}(\gamma-\delta,\boldsymbol{\varphi}) / \sum_{t=1}^{T} c_{t-1}^{2}(\gamma-\delta,\boldsymbol{\varphi})$. By straightforward application of the mean value theorem,

$$k_{t-1,T}(\widehat{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}}) = k_{t-1,T}(\gamma_0-\widehat{\delta},\widehat{\boldsymbol{\varphi}}) + k_{t-1,T}^{(1)}(\overline{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}})(\widehat{\gamma}-\gamma_0),$$

where $k_{t,T}^{(1)}(\cdot,\cdot)$ is the derivative of $k_{t,T}(\cdot,\cdot)$ with respect to the first argument and $|\overline{\gamma}-\gamma_0| \leq |\widehat{\gamma}-\gamma_0|$. Thus,

$$\widehat{\mu} = \mu_0 + \mu_0(\widehat{\gamma}-\gamma_0)\sum_{t=1}^{T} c_{t-1}(\gamma_0-\widehat{\delta},\widehat{\boldsymbol{\varphi}})k_{t-1,T}^{(1)}(\overline{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}}) + \sum_{t=1}^{T}\phi(L;\widehat{\boldsymbol{\varphi}})\,u_t(\widehat{\delta}-\delta_0)k_{t-1,T}(\widehat{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}}),$$

which implies that

$$\begin{aligned}\frac{T^{\gamma_0-\delta_0-1/2}}{\log T}(\widehat{\mu}-\mu_0) &= \mu_0 T^{\gamma_0-\delta_0-1/2}(\widehat{\gamma}-\gamma_0)\frac{1}{\log T}\sum_{t=1}^{T} c_{t-1}(\gamma_0-\widehat{\delta},\widehat{\boldsymbol{\varphi}})k_{t-1,T}^{(1)}(\overline{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}}) \\ &+ \frac{T^{\gamma_0-\delta_0-1/2}}{\log T}\sum_{t=1}^{T}\phi(L;\widehat{\boldsymbol{\varphi}})\,u_t(\widehat{\delta}-\delta_0)k_{t-1,T}(\widehat{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}}).\end{aligned}$$

Then, the remaining part of (13) holds on showing

$$\frac{1}{\log T}\sum_{t=1}^{T} c_{t-1}(\gamma_0-\widehat{\delta},\widehat{\boldsymbol{\varphi}})k_{t-1,T}^{(1)}(\overline{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}}) \to_p -1, \tag{98}$$

$$\frac{T^{\gamma_0-\delta_0-1/2}}{\log T}\sum_{t=1}^{T}\phi(L;\widehat{\boldsymbol{\varphi}})\,u_t(\widehat{\delta}-\delta_0)k_{t-1,T}(\widehat{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}}) = o_p(1). \tag{99}$$

Clearly (98) follows if

$$\sum_{t=1}^{T} c_{t-1}(\gamma_0-\widehat{\delta},\widehat{\boldsymbol{\varphi}})k_{t-1,T}^{(1)}(\overline{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}}) - \sum_{t=1}^{T} c_{t-1}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)k_{t-1,T}^{(1)}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0) = o_p(\log T), \tag{100}$$

$$\frac{1}{\log T}\sum_{t=1}^{T} c_{t-1}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)k_{t-1,T}^{(1)}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0) \to -1. \tag{101}$$

First, (100) can be easily justified by applying Lemmas 12 and 7, noting that

$$\begin{aligned}k_{t-1,T}^{(1)}(\gamma-\delta,\boldsymbol{\varphi}) =& \frac{c_{t-1}^{(1)}(\gamma-\delta,\boldsymbol{\varphi})}{\sum_{t=1}^{T} c_{t-1}^{2}(\gamma-\delta,\boldsymbol{\varphi})} \\ &- \frac{2c_{t-1}(\gamma-\delta,\boldsymbol{\varphi})\sum_{t=1}^{T} c_{t-1}^{(1)}(\gamma-\delta,\boldsymbol{\varphi})\,c_{t-1}(\gamma-\delta,\boldsymbol{\varphi})}{\left(\sum_{t=1}^{T} c_{t-1}^{2}(\gamma-\delta,\boldsymbol{\varphi})\right)^2}.\end{aligned}$$

Next, the left-hand side of (101) is

$$\begin{aligned}&-\frac{1}{\log T}\frac{\sum_{t=1}^{T} c_{t-1}^{(1)}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)\,c_{t-1}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)}{\sum_{t=1}^{T} c_{t-1}^{2}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)} \\ &\quad = -\frac{1}{\log T}\frac{\sum_{t=1}^{T}\pi_{t-1}^{(1)}(\gamma_0-\delta_0)\,\pi_{t-1}(\gamma_0-\delta_0)}{\sum_{t=1}^{T}\pi_{t-1}^{2}(\gamma_0-\delta_0)} + o(1),\end{aligned}$$

as in (82) and (84). Thus (101) follows immediately, noting (83), (86) and that, by Lemmas 10 and 11, for $d > 1/2$,

$$\frac{1}{T^{2d-1}\log T}\sum_{t=1}^{T}\log\left(d+t-1\right)\pi_{t-1}^{2}\left(d\right)\to\frac{1}{\Gamma^{2}\left(d\right)\left(2d-1\right)}.$$

Next, the left-hand side of (99) is

$$\frac{T^{\gamma_0-\delta_0-1/2}}{\log T}\sum_{t=1}^{T}\phi\left(L;\boldsymbol{\varphi}_0\right)\left\{u_t\mathbb{I}(t>0)\right\}k_{t-1,T}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0) \tag{102}$$

$$+\frac{T^{\gamma_0-\delta_0-1/2}}{\log T}\sum_{t=1}^{T}(\phi\left(L;\widehat{\boldsymbol{\varphi}}\right)u_t(\widehat{\delta}-\delta_0)-\phi\left(L;\boldsymbol{\varphi}_0\right)\left\{u_t\mathbb{I}(t>0)\right\})k_{t-1,T}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0) \tag{103}$$

$$+\frac{T^{\gamma_0-\delta_0-1/2}}{\log T}\sum_{t=1}^{T}\phi\left(L;\boldsymbol{\varphi}_0\right)\left\{u_t\mathbb{I}(t>0)\right\}\left(k_{t-1,T}(\widehat{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}})-k_{t-1,T}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)\right) \tag{104}$$

$$+\frac{T^{\gamma_0-\delta_0-1/2}}{\log T}\sum_{t=1}^{T}(\phi\left(L;\widehat{\boldsymbol{\varphi}}\right)u_t(\widehat{\delta}-\delta_0)-\phi\left(L;\boldsymbol{\varphi}_0\right)\left\{u_t\mathbb{I}(t>0)\right\})$$
$$\times\left(k_{t-1,T}(\widehat{\gamma}-\widehat{\delta},\widehat{\boldsymbol{\varphi}})-k_{t-1,T}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)\right). \tag{105}$$

Note that

$$\phi(L;\boldsymbol{\varphi}_0)\left\{u_t\mathbb{I}(t>0)\right\}=\varepsilon_t-\sum_{j=t}^{\infty}\phi_j\left(\boldsymbol{\varphi}_0\right)u_{t-j}, \tag{106}$$

where, by Assumptions A1 and A2, it can be easily shown that

$$\sum_{j=t}^{\infty}\phi_j\left(\boldsymbol{\varphi}_0\right)u_{t-j}=O_p(t^{-1/2-\varsigma}). \tag{107}$$

Using summation by parts, (106), (107) and noting that, as in Lemmas 12 and 13,

$$k_{t-1,T}(\gamma_0-\delta_0,\boldsymbol{\varphi}_0)=O_p(t^{\gamma_0-\delta_0-1}T^{1-2(\gamma_0-\delta_0)}), \tag{108}$$

it follows that (102) is $O_p(\log^{-1}T)$. Next, by (108) and Lemma 7 with $\varkappa = 1/2$ (because $\widehat{\boldsymbol{\tau}}$ is $T^{1/2}$-consistent), (103) is $O_p(T^{1/2-\varkappa}\log^{-1}T) = O_p(\log^{-1}T)$. Next, by summation by parts, (106), (107), the mean value theorem and Lemmas 14 and 7, it can be easily shown that (104) is $O_p(T^{\theta+1/2-(\gamma_0-\delta_0)}) = o_p(1)$, setting $\theta < \gamma_0 - \delta_0 - 1/2$. Finally, combining the arguments for (103) and (104), it is straightforward to show that (105) is $o_p(1)$, to conclude the proof of (99).

A.4 Proof of Theorem 2(ii): the $\gamma_0 - 1/2 < \delta_0$ case

First, noting (74), the loss function $R_T(\boldsymbol{\vartheta})$ can be decomposed in the sum of two terms, $R_T(\boldsymbol{\vartheta}) = Q_T(\boldsymbol{\tau}) + S_T(\boldsymbol{\vartheta})$, where $Q_T(\boldsymbol{\tau}) = T^{-1}\sum_{t=1}^{T} s_{1t}^2(\boldsymbol{\tau})$ and

$$S_T\left(\boldsymbol{\vartheta}\right)=\frac{1}{T}\sum_{t=1}^{T}\left(d_t\left(\boldsymbol{\vartheta}\right)-s_{2t}\left(\boldsymbol{\vartheta}\right)\right)^2+\frac{2}{T}\sum_{t=1}^{T}s_{1t}\left(\boldsymbol{\tau}\right)\left(d_t\left(\boldsymbol{\vartheta}\right)-s_{2t}\left(\boldsymbol{\vartheta}\right)\right). \tag{109}$$

Thus, $Q_T(\boldsymbol{\tau})$ is the loss function in Hualde and Robinson (2011). Now

$$\frac{\partial R_T(\widehat{\boldsymbol{\vartheta}})}{\partial \boldsymbol{\tau}} = 0 = \frac{\partial Q_T(\widehat{\boldsymbol{\tau}})}{\partial \boldsymbol{\tau}} + \frac{\partial S_T(\widehat{\boldsymbol{\vartheta}})}{\partial \boldsymbol{\tau}}, \tag{110}$$

and by the mean value theorem

$$\frac{\partial Q_T(\widehat{\boldsymbol{\tau}})}{\partial \boldsymbol{\tau}} = \frac{\partial Q_T(\boldsymbol{\tau}_0)}{\partial \boldsymbol{\tau}} + \frac{\partial^2 Q_T(\overline{\boldsymbol{\tau}})}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}'} \left(\widehat{\boldsymbol{\tau}} - \boldsymbol{\tau}_0\right), \tag{111}$$

where $\overline{\boldsymbol{\tau}}$ represents an intermediate point between $\widehat{\boldsymbol{\tau}}$ and $\boldsymbol{\tau}_0$ which is allowed to vary in the different rows of $\partial^2 Q_T(\cdot)/\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}'$. Inserting (111) in (110) we then find

$$T^{1/2}\left(\widehat{\boldsymbol{\tau}} - \boldsymbol{\tau}_0\right) = -\left(\frac{\partial^2 Q_T(\overline{\boldsymbol{\tau}})}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}'}\right)^{-1} T^{1/2} \frac{\partial Q_T(\boldsymbol{\tau}_0)}{\partial \boldsymbol{\tau}} - \left(\frac{\partial^2 Q_T(\overline{\boldsymbol{\tau}})}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}'}\right)^{-1} T^{1/2} \frac{\partial S_T(\widehat{\boldsymbol{\vartheta}})}{\partial \boldsymbol{\tau}}. \tag{112}$$

Now, by Hualde and Robinson (2011) (see the proof of their Theorem 2.2), the first term on the right-hand side of (112) has a $N\left(0, \boldsymbol{A}^{-1}\right)$ limiting distribution, and $\partial^2 Q_T(\overline{\boldsymbol{\tau}})/\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}'$ converges in probability to a nonsingular matrix. Thus, in view of (112), Theorem 2(ii) follows because $T^{1/2} \partial S_T(\widehat{\boldsymbol{\vartheta}})/\partial \boldsymbol{\tau} = o_p(1)$ by Lemma 8.

B Auxiliary lemmas

Lemma 1 *Under Assumptions A1–A3,* $T^{-1}\sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}_0) \to_p \sigma_0^2$.

Lemma 2 *Under Assumptions A1 and A3, for any* $g > 0$,

$$\lim_{T\to\infty} \inf_{\gamma_0 - \delta \geq 1/2 + g, |\gamma - \gamma_0| \geq g, \boldsymbol{\varphi} \in \Psi} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_t^2(\boldsymbol{\vartheta}) > \epsilon.$$

Lemma 3 *Under Assumptions A1 and A3, for* $i = 1, ..., 4$,

$$\frac{1}{T^{\gamma_0 - \delta}} \frac{\partial \overline{d}_T(\boldsymbol{\vartheta})}{\partial \gamma} = \frac{\mu_0 \phi(1; \boldsymbol{\varphi})}{\Gamma(\gamma_0 - \delta)} \frac{2(\gamma - \delta)^2 - 2(\gamma - \delta) + 1 - (\gamma_0 - \delta)}{(\gamma - \delta)^2 (\gamma_0 + \gamma - 2\delta - 1)^2} + g_T(\boldsymbol{\vartheta}), \tag{113}$$

where $\sup_{\overline{\mathcal{H}}_i} |g_T(\boldsymbol{\vartheta})| = o(1)$, *and for an arbitrarily large* K *(setting* ξ *large enough),*

$$\lim_{T\to\infty} \inf_{\overline{\mathcal{H}}_i} \frac{T^{2\varkappa_i}}{T^{2(\gamma_0 - \delta)}} \overline{d}_T^2(\boldsymbol{\vartheta}) > K. \tag{114}$$

Lemma 4 *Under the conditions of Theorem 1(ii) it holds that*

$$\Pr\left(\inf_{\boldsymbol{\vartheta} \in \overline{M}_\varepsilon} R_T(\boldsymbol{\vartheta}) \leq \sigma_0^2 + \epsilon\right) \to 0 \text{ as } T \to \infty.$$

Lemma 5 *Under the conditions of Theorem 2(i) it holds that:*

(a) The first-order derivatives satisfy

$$\frac{1}{T^{1/2}} \sum_{t=1}^{T} \left(s_t(\boldsymbol{\vartheta}_0) - \varepsilon_t\right) \frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial \boldsymbol{\tau}} = o_p(1), \frac{1}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=1}^{T} \left(s_t(\boldsymbol{\vartheta}_0) - \varepsilon_t\right) \frac{\partial d_t(\boldsymbol{\vartheta}_0)}{\partial \gamma} = o_p(1), \tag{115}$$

$$\frac{1}{T^{1/2}} \sum_{t=1}^{T} s_t(\boldsymbol{\vartheta}_0) \frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\tau}} = o_p(1), \frac{1}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=1}^{T} s_t(\boldsymbol{\vartheta}_0) \frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \gamma} = o_p(1). \tag{116}$$

(b) The second-order derivatives satisfy

$$\frac{1}{T}\sum_{t=1}^{T}\frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial \boldsymbol{\tau}}\frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\tau}'}=o_p(1), \frac{1}{T^{\gamma_0-\delta_0}}\sum_{t=1}^{T}\frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial \boldsymbol{\tau}}\frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \gamma}=o_p(1),$$

$$\frac{1}{T^{\gamma_0-\delta_0}}\sum_{t=1}^{T}\frac{\partial s_{1t}(\boldsymbol{\tau}_0)}{\partial \boldsymbol{\tau}}\frac{\partial d_t(\boldsymbol{\vartheta}_0)}{\partial \gamma}=o_p(1), \frac{1}{T}\sum_{t=1}^{T}\frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\tau}}\frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\tau}'}=o_p(1),$$

$$\frac{1}{T^{\gamma_0-\delta_0}}\sum_{t=1}^{T}\frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\tau}}\frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \gamma}=o_p(1), \frac{1}{T^{\gamma_0-\delta_0}}\sum_{t=1}^{T}\frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\tau}}\frac{\partial d_t(\boldsymbol{\vartheta}_0)}{\partial \gamma}=o_p(1),$$

$$\frac{1}{T^{2(\gamma_0-\delta_0)-1}}\sum_{t=1}^{T}\left(\frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \gamma}\right)^2=o_p(1), \frac{1}{T^{2(\gamma_0-\delta_0)-1}}\sum_{t=1}^{T}\frac{\partial s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \gamma}\frac{\partial d_t(\boldsymbol{\vartheta}_0)}{\partial \gamma}=o_p(1),$$

$$\frac{1}{T}\sum_{t=1}^{T}s_t(\boldsymbol{\vartheta}_0)\frac{\partial^2 s_t(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\tau}\partial \boldsymbol{\tau}'}=o_p(1), \frac{1}{T^{\gamma_0-\delta_0}}\sum_{t=1}^{T}s_t(\boldsymbol{\vartheta}_0)\frac{\partial^2 s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\tau}\partial \gamma}=o_p(1),$$

$$\frac{1}{T^{2(\gamma_0-\delta_0)-1}}\sum_{t=1}^{T}s_t(\boldsymbol{\vartheta}_0)\frac{\partial^2 s_{2t}(\boldsymbol{\vartheta}_0)}{\partial \gamma^2}=o_p(1), \frac{1}{T^{\gamma_0-\delta_0}}\sum_{t=1}^{T}s_t(\boldsymbol{\vartheta}_0)\frac{\partial^2 d_t(\boldsymbol{\vartheta}_0)}{\partial \boldsymbol{\tau}\partial \gamma}=o_p(1),$$

$$\frac{1}{T^{2(\gamma_0-\delta_0)-1}}\sum_{t=1}^{T}s_t(\boldsymbol{\vartheta}_0)\frac{\partial^2 d_t(\boldsymbol{\vartheta}_0)}{\partial \gamma^2}=o_p(1).$$

Lemma 6 *Under the conditions of Theorem 2(i), for some fixed* $\varkappa > 0$, $T^{\varkappa}(\widehat{\boldsymbol{\vartheta}}-\boldsymbol{\vartheta}_0) \to_p 0$.

Lemma 7 *Let* $\widehat{\boldsymbol{\tau}}-\boldsymbol{\tau}_0 = O_p(T^{-\varkappa})$ *for* $\varkappa > 0$. *Then, under Assumptions A1–A4,*

$$c_t(\widehat{\delta},\widehat{\boldsymbol{\varphi}}) = c_t(\delta_0,\boldsymbol{\varphi}_0) + O_p\left(T^{-\varkappa}t^{\max\{\delta_0-1,-1-\varsigma\}}\log^2 t\right), \tag{117}$$

$$c_t^{(1)}(\widehat{\delta},\widehat{\boldsymbol{\varphi}}) = c_t^{(1)}(\delta_0,\boldsymbol{\varphi}_0) + O_p\left(T^{-\varkappa}t^{\max\{\delta_0-1,-1-\varsigma\}}\log^3 t\right), \tag{118}$$

and, uniformly in $t = 1, ..., T$,

$$\phi(L;\widehat{\boldsymbol{\varphi}})\,u_t(\widehat{\delta}-\delta_0) = \sum_{j=0}^{t-1}\phi_j(\widehat{\boldsymbol{\varphi}})\,u_{t-j}(\widehat{\delta}-\delta_0) = \sum_{j=0}^{t-1}\phi_j(\boldsymbol{\varphi}_0)\,u_{t-j} + O_p\left(T^{-\varkappa}\right). \tag{119}$$

Lemma 8 *Under the conditions of Theorem 2(ii),* $T^{1/2}\partial S_T(\widehat{\boldsymbol{\vartheta}})/\partial \boldsymbol{\tau} = o_p(1)$.

C Technical lemmas

Lemma 9 *Uniformly for* $\max\{|\alpha|,|\beta|\} \le a_0$, $\sum_{j=1}^{t-1} j^{\alpha-1}(t-j)^{\beta-1} \le K(\log t)t^{\max\{\alpha+\beta-1,\alpha-1,\beta-1\}}$.

Lemma 10 *For any* $d > 0$ *and any fixed* $a \ge 0$, *as* $T \to \infty$,

$$\frac{1}{T^d}\sum_{t=1}^{T}t^{d-1} \to \frac{1}{d}, \quad \frac{1}{T^d}\sum_{t=1}^{T}\log\left(\frac{t+a}{T}\right)t^{d-1} \to -\frac{1}{d^2}, \quad \frac{1}{T^d}\sum_{t=1}^{T}\log^2\left(\frac{t+a}{T}\right)t^{d-1} \to \frac{2}{d^3}.$$

Lemma 11 *Let $j \geq 1$ and $\mathbb{K}$ denote any compact subset of $\mathbb{R}\backslash\mathbb{N}_0$. Then*

$$\pi_j(-v) = \frac{1}{\Gamma(-v)} j^{-v-1}(1 + \epsilon_j(v)), \tag{120}$$

where $\max_{v\in\mathbb{K}} |\epsilon_j(v)| \to 0$ as $j \to \infty$. Thus, uniformly in $j \geq 1, m \geq 0$,

(i) $\pi_j(-v) \geq Kj^{-v-1}$ *uniformly in* $v \in \mathbb{K}$,

(ii) $|\frac{\partial^m}{\partial u^m}\pi_j(u)| \leq K(1 + \log j)^m j^{u-1}$ *uniformly in* $|u| \leq u_0$,

(iii) $|\frac{\partial^m}{\partial u^m}T^{-u}\pi_j(u)| \leq KT^{-u}(1 + |\log(j/T)|)^m j^{u-1}$ *uniformly in* $|u| \leq u_0$.

Lemma 12 *Under Assumptions A1, A3, uniformly in $t = 1, \ldots, T$ and $T \geq 1$, for $m \geq 0$,*

$$\sup_{d\leq g,\boldsymbol{\varphi}\in\Psi} \left|\frac{\partial^m c_t(d,\boldsymbol{\varphi})}{\partial d^m}\right| = O(t^{\max\{g-1,-1-\varsigma\}} (\log t)^m), \tag{121}$$

$$\sup_{d\geq g,\boldsymbol{\varphi}\in\Psi} \left|\frac{\partial^m}{\partial d^m} T^{-d} c_t(d,\boldsymbol{\varphi})\right| = O(T^{-g} t^{\max\{g-1,-1-\varsigma\}}(1 + |\log(t/T)|)^m). \tag{122}$$

Lemma 13 *Under Assumptions A1 and A3,*

$$\frac{1}{T^{2d-1}} \sum_{t=1}^{T} c_{t-1}^2(d,\boldsymbol{\varphi}) \geq \frac{\phi^2(1;\boldsymbol{\varphi})}{T^{2d-1}} \sum_{t=1}^{T} \pi_{t-1}^2(d) - |r_{1,T}(d,\boldsymbol{\varphi})|, \tag{123}$$

$$\frac{1}{T^{2d-1}} \sum_{t=1}^{T} c_{t-1}^2(d,\boldsymbol{\varphi}) \leq \frac{\phi^2(1;\boldsymbol{\varphi})}{T^{2d-1}} \sum_{t=1}^{T} \pi_{t-1}^2(d) + |r_{2,T}(d,\boldsymbol{\varphi})|, \tag{124}$$

where, for any $\eta > 0$, $\sup_{d\geq 1/2+\eta,\boldsymbol{\varphi}\in\Psi} |r_{i,T}(d,\boldsymbol{\varphi})| = o(1), i = 1, 2$. Furthermore, for any $d_1 \leq d_2$ and any α such that $0 < \alpha < (\varsigma - 1/2)/3$,

$$\inf_{d_1\leq d\leq d_2,\boldsymbol{\varphi}\in\Psi} \sum\nolimits_{t=1}^{T} c_{t-1}^2(d,\boldsymbol{\varphi}) \geq 1, \tag{125}$$

$$\inf_{d\geq 1/2-\alpha,\boldsymbol{\varphi}\in\Psi} \frac{1}{T^{2d-1}} \sum\nolimits_{t=1}^{T} c_{t-1}^2(d,\boldsymbol{\varphi}) \geq \frac{\epsilon}{\alpha} + o(1), \tag{126}$$

for some $\epsilon > 0$, which does not depend on α or T.

Lemma 14 *Let θ be an arbitrary number such that $0 < \theta < \varsigma - 1/2$. Then, under Assumptions A1 and A3, for any real numbers $d_1 < 1/2 - \theta$ and $d_2 > 1/2 + \theta$, $m = 0, 1$, uniformly in $t = 1, \ldots, T$ and $T \geq 1$,*

$$\sup_{d\in[d_1,d_2],\boldsymbol{\varphi}\in\Psi} \left|\frac{\partial^m}{\partial d^m} h_{t-1,T}(d,\boldsymbol{\varphi})\right| = O(t^{-1/2-\theta} T^{\theta+2\theta m}), \tag{127}$$

$$\sup_{d\geq 1/2+\theta,\boldsymbol{\varphi}\in\Psi} \left|\frac{\partial}{\partial d} h_{t-1,T}(d,\boldsymbol{\varphi})\right| = O(t^{-1/2+\theta} T^{-\theta}(1 + |\log(t/T)|)), \tag{128}$$

$$\sup_{d\in[d_1,d_2],\boldsymbol{\varphi}\in\Psi} \left|\frac{\partial^m}{\partial d^m}(h_{t,T}(d,\boldsymbol{\varphi}) - h_{t-1,T}(d,\boldsymbol{\varphi}))\right| = O(t^{-3/2-\theta} T^{\theta+2\theta m}), \tag{129}$$

$$\sup_{d\geq 1/2+\theta,\boldsymbol{\varphi}\in\Psi} \left|\frac{\partial}{\partial d}(h_{t,T}(d,\boldsymbol{\varphi}) - h_{t-1,T}(d,\boldsymbol{\varphi}))\right| = O(t^{-3/2+\theta} T^{-\theta}(1 + |\log(t/T)|)), \tag{130}$$

$$\sup_{d\in[d_1,d_2],\boldsymbol{\varphi}\in\Psi} \left|\sum_{t=1}^{T} h_{t-1,T}(d,\boldsymbol{\varphi})\right| = O(T^{1/2}). \tag{131}$$

Lemma 15 *Under Assumptions A1–A3, uniformly in* $t = 1, \ldots, T$, $T \geq 1$, *and* $\boldsymbol{\varphi} \in \Psi$,

$$\sup_{d \leq g} |\phi(L; \boldsymbol{\varphi}) u_t(-d)| = O_p(t^{g-1/2} + \log t \mathbb{I}(g = 1/2) + \mathbb{I}(g < 1/2)), \tag{132}$$

$$\sup_{d \geq g} |T^{-d} \phi(L; \boldsymbol{\varphi}) u_t(-d)| = O_p(T^{-g}(t^{g-1/2} + \log t \mathbb{I}(g = 1/2) + \mathbb{I}(g < 1/2))). \tag{133}$$

Lemma 16 *Let* θ *be an arbitrary number such that* $0 < \theta < \varsigma - 1/2$. *Then, under Assumptions A1–A3, for* $m = 0, 1$, *and uniformly in* $\boldsymbol{\vartheta} \in \Xi$,

$$\sup_{\delta_0 - \delta \leq g} \left| \sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) \frac{\partial^m}{\partial \gamma^m} h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) \right| = O_p(T^{\max\{\theta, g\} + 2\theta m}), \tag{134}$$

$$\sup_{\delta_0 - \delta \geq g} \frac{1}{T^{\delta_0 - \delta}} \left| \sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) \frac{\partial^m}{\partial \gamma^m} h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) \right| = O_p(T^{\max\{\theta, g\} - g + 2\theta m}), \tag{135}$$

$$\sup_{\delta_0 - \delta \leq g, \gamma - \delta \geq 1/2 + \theta} \left| \sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) \frac{\partial}{\partial \gamma} h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) \right| = O_p(T^{\max\{\theta, g\}}), \tag{136}$$

$$\sup_{\delta_0 - \delta \geq g, \gamma - \delta \geq 1/2 + \theta} \frac{1}{T^{\delta_0 - \delta}} \left| \sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) \frac{\partial}{\partial \gamma} h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) \right| = O_p(T^{\max\{\theta, g\} - g}). \tag{137}$$

Lemma 17 *Under Assumptions A1–A3, for any* $g_2 > 1/2$ *and for any arbitrary* θ *such that* $0 < \theta < \min\{\varsigma - 1/2, g_2 - 1/2\}$,

$$\left| \sum_{t=1}^{T} d_t(\boldsymbol{\vartheta}) s_t(\boldsymbol{\vartheta}) \right| \leq |\gamma - \gamma_0| |M_T(\boldsymbol{\vartheta})|, \tag{138}$$

where, uniformly in $\boldsymbol{\vartheta} \in \Xi$,

$$\sup_{\delta_0 - \delta \leq g_1, \gamma_0 - \delta \leq g_2} |M_T(\boldsymbol{\vartheta})| = O_p(T^{\max\{\theta, g_1\} + 2\theta + g_2 - 1/2}), \tag{139}$$

$$\sup_{\delta_0 - \delta \leq g_1, \gamma_0 - \delta \geq g_2} T^{\delta - \gamma_0} |M_T(\boldsymbol{\vartheta})| = O_p(T^{\max\{\theta, g_1\} + 2\theta - 1/2}), \tag{140}$$

$$\sup_{\delta_0 - \delta \geq g_1, \gamma_0 - \delta \geq g_2} T^{2\delta - \delta_0 - \gamma_0} |M_T(\boldsymbol{\vartheta})| = O_p(T^{\max\{\theta, g_1\} + 2\theta - g_1 - 1/2}), \tag{141}$$

$$\sup_{\delta_0 - \delta \leq g_1, \gamma_0 - \delta \geq g_2, \gamma - \delta \geq 1/2 + \theta} T^{\delta - \gamma_0} |M_T(\boldsymbol{\vartheta})| = O_p(T^{\max\{\theta, g_1\} - 1/2}), \tag{142}$$

$$\sup_{\delta_0 - \delta \geq g_1, \gamma_0 - \delta \geq g_2, \gamma - \delta \geq 1/2 + \theta} T^{2\delta - \delta_0 - \gamma_0} |M_T(\boldsymbol{\vartheta})| = O_p(T^{\max\{\theta, g_1\} - g_1 - 1/2}). \tag{143}$$

References

1. Abadir, K.M., Distaso, W. and Giraitis, L. (2007). Nonstationarity-extended local Whittle estimation. *Journal of Econometrics* **141**, 1353–1384.

2. 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, Series B* **57**, 659–672.
3. Bloomfield, P. (1973). An exponential model for the spectrum of a scalar time series. *Biometrika* **60**, 217–226.
4. Box, G.E.P. and Jenkins, G.M. (1971). *Time Series Analysis, Forecasting and Control.* Holden-Day, San Fracisco.
5. Chen, W.W. and Hurvich, C.M. (2003). Estimating fractional cointegration in the presence of polynomial trends. *Journal of Econometrics* **117**, 95–121.
6. Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Annals of Statistics* **17**, 1749–1766.
7. Fox, R. and Taqqu, M.S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of Statistics* **14**, 517–532.
8. Giraitis, L. and Surgailis, D. (1990). A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotical normality of Whittle's estimate. *Probability Theory and Related Fields* **86**, 87–104.
9. Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and Its Application*, Academic Press, New York.
10. Hannan, E.J. (1973). The asymptotic theory of linear time series models. *Journal of Applied Probability* **10**, 130–145.
11. Heyde, C.C. and Dai, W. (1996). On the robustness to small trends of estimation based on the smoothed periodogram. *Journal of Time Series Analysis* **17**, 141–150.
12. Hualde, J. and Nielsen, M.Ø. (2017). Supplementary appendix to "Truncated sum of squares estimation of fractional time series models with deterministic trends".
13. Hualde, J. and Robinson, P.M. (2011). Gaussian pseudo-maximum likelihood estimation of fractional time series models. *Annals of Statistics* **39**, 3152–3181.
14. Iacone, F. (2010). Local Whittle estimation of the memory parameter in presence of deterministic components. *Journal of Time Series Analysis* **31**, 37–49.
15. Johansen, S. and Nielsen, M.Ø. (2016). The role of initial values in conditional sum-of-squares estimation of nonstationary fractional time series models. *Econometric Theory* **32**, 1095–1139.
16. Li, W.K. and McLeod, A.I. (1986). Fractional time series modelling. *Biometrika* **73**, 217–221.
17. Nielsen, M.Ø. (2015). Asymptotics for the conditional-sum-of-squares estimator in multivariate fractional time series models. *Journal of Time Series Analysis* **36**, 154–188.
18. Phillips, P.C.B. (2007). Regression with slowly varying regressors and nonlinear trends. *Econometric Theory* **23**, 557–614.
19. Robinson, P.M. (1994). Efficient tests of nonstationary hypotheses. *Journal of the American Statistical Association* **89**, 1420–1437.
20. Robinson, P.M. (2005). Efficiency improvements in inference on stationary and nonstationary fractional time series. *Annals of Statistics* **33**, 1800–1842.

21. Robinson, P.M. (2012). Inference on power law spatial trends. *Bernoulli* **18**, 644–677.
22. Robinson, P.M. and Iacone, F. (2005). Cointegration in fractional systems with deterministic trends. *Journal of Econometrics* **129**, 263–298.
23. Robinson, P.M. and Marinucci, D. (2000). The averaged periodogram for nonstationary vector time series. *Statistical Inference for Stochastic Processes* **3**, 149–160.
24. Velasco, C. (1999a). Non-stationary log-periodogram regression. *Journal of Econometrics* **91**, 325–371.
25. Velasco, C. (1999b). Gaussian semiparametric estimation of non-stationary time series. *Journal of Time Series Analysis* **20**, 87–127.
26. White, H. and Granger, C.W.J. (2011). Consideration of trends in time series. *Journal of Time Series Econometrics* **3**, issue 1, article 2.
27. Wu, C.-F. (1981). Asymptotic theory of nonlinear least squares estimation. *Annals of Statistics* **9**, 501–513.

Supplementary Appendix

to

Truncated sum of squares estimation of fractional time series models with deterministic trends

by

Javier Hualde
Universidad Pública de Navarra

and

Morten Ørregaard Nielsen
Queen's University and CREATES

Date: May 15, 2017

S.1 Introduction

This supplements Hualde and Nielsen (2017) by providing proofs of all lemmas. Equation references (S.n) for $n \geq 1$ refer to equations in this supplement and other equation references are to the main paper, Hualde and Nielsen (2017). Note that the proofs of the auxiliary lemmas rely on the technical lemmas, but not vice versa. The proofs of the lemmas in Hualde and Nielsen (2017) apply two additional technical lemmas, which are provided at the end of this supplement along with their proofs.

S.2 Proofs of auxiliary lemmas

S.2.1 Proof of Lemma 1

Clearly

$$\frac{1}{T}\sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}_0) = \frac{1}{T}\sum_{t=1}^{T}\left(\phi(L;\boldsymbol{\varphi}_0)\left\{u_t\mathbb{I}(t>0)\right\}\right)^2 - \frac{1}{T}\left(\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi}_0)\left\{u_t\mathbb{I}(t>0)\right\}h_{t-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right)\right)^2. \quad \text{(S.1)}$$

In view of (106), (107), by Assumptions A1 and A2 and simple application of Lemma 14, the second term on the right-hand side of (S.1) is $O_p\left(T^{2\theta-1}\right) = o_p(1)$ by choosing $\theta < 1/2$. Then the required result holds by (24).

S.2.2 Proof of Lemma 2

Letting $\alpha > 0$ be arbitrarily small (in particular $\alpha < (\varsigma - 1/2)/3$, which implies $\alpha < 1/2$) and defining $\Phi_1 = \{\boldsymbol{\vartheta} \in \Xi : \gamma - \delta \leq 1/2 - \alpha\}$, $\Phi_2 = \{\boldsymbol{\vartheta} \in \Xi : 1/2 - \alpha \leq \gamma - \delta \leq 1/2 + \alpha\}$, and $\Phi_3 = \{\boldsymbol{\vartheta} \in \Xi : \gamma - \delta \geq 1/2 + \alpha\}$, the result holds on showing

$$\lim_{T\to\infty}\inf_{\{\gamma_0-\delta\geq 1/2+g,|\gamma-\gamma_0|\geq g\}\cap\Phi_j}\frac{1}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}d_t^2(\boldsymbol{\vartheta}) > \epsilon, \quad \text{(S.2)}$$

for $j = 1, 2, 3$. We first deal with $j = 1, 2$. Clearly

$$\frac{1}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}d_t^2(\boldsymbol{\vartheta}) = \frac{\mu_0^2}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}c_{t-1}^2\left(\gamma_0-\delta,\boldsymbol{\varphi}\right) - \frac{\mu_0^2}{T^{2(\gamma_0-\delta)-1}}\left(\sum\nolimits_{t=1}^{T}c_{t-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right)^2,$$

so because $|\gamma-\gamma_0| \geq g$ and by application of Lemma 13, noting that $\gamma_0 - \delta \geq 1/2 + g > 1/2$, $\mu_0 \neq 0$, and (8), (S.2) for $j = 1, 2$ holds on showing

$$\lim_{T\to\infty}\inf_{\gamma_0-\delta\geq 1/2+g}\frac{1}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}\pi_{t-1}^2\left(\gamma_0-\delta\right) > \epsilon, \quad \text{(S.3)}$$

$$\sup_{\{\gamma_0-\delta\geq 1/2+g\}\cap\Phi_j}\frac{1}{T^{2(\gamma_0-\delta)-1}}\left(\sum\nolimits_{t=1}^{T}c_{t-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right)^2 = o(1). \quad \text{(S.4)}$$

First, we show (S.3). By (120) in Lemma 11,

$$\inf_{\gamma_0-\delta\geq 1/2+g}\frac{1}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}\pi_{t-1}^2\left(\gamma_0-\delta\right)\geq\epsilon\inf_{\gamma_0-\delta\geq 1/2+g}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{t}{T}\right)^{2(\gamma_0-\delta)-2}=\epsilon\frac{1}{2g}+o(1),\quad\text{(S.5)}$$

so that (S.3) holds by taking limits as $T\to\infty$. Next, we show (S.4) for $j=1$. By (125) of Lemma 13, the left-hand side of (S.4) is bounded by

$$\sup_{\{\gamma_0-\delta\geq 1/2+g\}\cap\Phi_1}\left(T^{1/2-(\gamma_0-\delta)}\sum\nolimits_{t=1}^{T}c_{t-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)c_{t-1}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right)^2.\quad\text{(S.6)}$$

By Lemma 12, (S.6) is $O\left(T^{-\alpha}\right)=o\left(1\right)$ to conclude the proof of (S.4), and therefore that of (S.2), for $j=1$. Regarding $j=2$, the left-hand side of (S.4) is bounded by

$$\frac{\left(\sup_{\{\gamma_0-\delta\geq 1/2+g\}\cap\Phi_2}T^{-(\gamma_0-2\delta+\gamma-1)}\sum\nolimits_{t=1}^{T}c_{t-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)c_{t-1}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right)^2}{\inf_{\Phi_2}T^{-2(\gamma-\delta)+1}\sum\nolimits_{t=1}^{T}c_{t-1}^2\left(\gamma-\delta,\boldsymbol{\varphi}\right)},\quad\text{(S.7)}$$

where the denominator can be made arbitrarily large by setting α close enough to zero, see (126) of Lemma 13. By (122) of Lemma 12 the square-root of the numerator of (S.7) is $O(T^{-1/2-\eta+\alpha}\sum_{t=1}^{T}t^{\eta-1/2-\alpha})=O\left(1\right)$. This completes the proof of (S.4), and hence that of (S.2), for $j=2$.

Finally we show (S.2) for $j=3$. By very similar steps to those in the proof of Lemma 13, noting that $\gamma_0-\delta\geq 1/2+g$, it is straightforward to show that

$$\frac{1}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}d_t^2\left(\boldsymbol{\vartheta}\right)=\frac{\mu_0^2\phi^2(1;\boldsymbol{\varphi})}{T^{2(\gamma_0-\delta)-1}}\left(\sum_{t=1}^{T}\pi_{t-1}^2\left(\gamma_0-\delta\right)-\frac{\left(\sum_{t=1}^{T}\pi_{t-1}\left(\gamma_0-\delta\right)\pi_{t-1}\left(\gamma-\delta\right)\right)^2}{\sum_{t=1}^{T}\pi_{t-1}^2\left(\gamma-\delta\right)}\right)+q_T\left(\gamma_0,\gamma,\delta,\boldsymbol{\varphi}\right),$$

where $\sup_{\{\gamma_0-\delta\geq 1/2+g\}\cap\Phi_3}\left|q_T\left(\gamma_0,\gamma,\delta,\boldsymbol{\varphi}\right)\right|=o\left(1\right)$. Next, using Lemma 11 and approximating sums by integrals, see Lemma 10,

$$\begin{aligned}&\inf_{\{\gamma_0-\delta\geq 1/2+g\}\cap\Phi_3}\frac{1}{T^{2(\gamma_0-\delta)-1}}\left[\sum_{t=1}^{T}\pi_{t-1}^2\left(\gamma_0-\delta\right)-\frac{\left(\sum_{t=1}^{T}\pi_{t-1}\left(\gamma_0-\delta\right)\pi_{t-1}\left(\gamma-\delta\right)\right)^2}{\sum_{t=1}^{T}\pi_{t-1}^2\left(\gamma-\delta\right)}\right]\\
&\geq\epsilon\inf_{\{\gamma_0-\delta\geq 1/2+g\}\cap\Phi_3}\frac{1}{\Gamma^2\left(\gamma_0-\delta\right)}\left[\frac{1}{T^{2(\gamma_0-\delta)-1}}\sum_{t=1}^{T}t^{2(\gamma_0-\delta-1)}-\frac{\left(\frac{1}{T^{\gamma_0+\gamma-2\delta-1}}\sum_{t=1}^{T}t^{\gamma_0+\gamma-2\delta-2}\right)^2}{\frac{1}{T^{2(\gamma-\delta)-1}}\sum_{t=1}^{T}t^{2(\gamma-\delta-1)}}\right]\\
&=\epsilon\inf_{\{\gamma_0-\delta\geq 1/2+g\}\cap\Phi_3}\frac{\left(\gamma_0-\gamma\right)^2}{\Gamma^2\left(\gamma_0-\delta\right)\left(2\left(\gamma_0-\delta\right)-1\right)\left(\gamma_0+\gamma-2\delta-1\right)^2}-o\left(1\right)\\
&\geq\epsilon\inf_{\gamma_0-\delta\geq 1/2+g}\frac{g^2}{\Gamma^2\left(\gamma_0-\delta\right)2g\left(\alpha+g\right)^2}-o\left(1\right),\end{aligned}$$

which is positive and bounded away from zero, to complete the proof of (S.2) for $j=3$.

S.2.3 Proof of Lemma 3

First, $\partial\overline{d}_T(\boldsymbol{\vartheta})/\partial\gamma$ equals

$$-\frac{\mu_0\sum_{j=1}^{T}c_{j-1}^{(1)}(\gamma-\delta,\boldsymbol{\varphi})\sum_{j=1}^{T}c_{j-1}(\gamma_0-\delta,\boldsymbol{\varphi})c_{j-1}(\gamma-\delta,\boldsymbol{\varphi})}{\sum_{j=1}^{T}c_{j-1}^2(\gamma-\delta,\boldsymbol{\varphi})}$$
$$-\frac{\mu_0\sum_{j=1}^{T}c_{j-1}(\gamma-\delta,\boldsymbol{\varphi})\sum_{j=1}^{T}c_{j-1}(\gamma_0-\delta,\boldsymbol{\varphi})c_{j-1}^{(1)}(\gamma-\delta,\boldsymbol{\varphi})}{\sum_{j=1}^{T}c_{j-1}^2(\gamma-\delta,\boldsymbol{\varphi})}$$
$$+\frac{2\mu_0\sum_{j=1}^{T}c_{j-1}(\gamma-\delta,\boldsymbol{\varphi})\sum_{j=1}^{T}c_{j-1}(\gamma_0-\delta,\boldsymbol{\varphi})c_{j-1}(\gamma-\delta,\boldsymbol{\varphi})\sum_{j=1}^{T}c_{j-1}(\gamma-\delta,\boldsymbol{\varphi})c_{j-1}^{(1)}(\gamma-\delta,\boldsymbol{\varphi})}{\left(\sum_{j=1}^{T}c_{j-1}^2(\gamma-\delta,\boldsymbol{\varphi})\right)^2}.$$

Noting that in $\cup_{i=1}^{4}\overline{\mathcal{H}}_i$, $\gamma_0-\delta\geq 1+\eta$ and $\gamma-\delta\geq 1+\eta-\varrho$, proceeding as in the proof of Lemma 13,

$$\begin{aligned}\frac{1}{T^{\gamma_0-\delta}}\frac{\partial\overline{d}_T(\boldsymbol{\tau},\gamma)}{\partial\gamma} &=\\ &-\frac{\mu_0\phi(1;\boldsymbol{\varphi})}{T^{\gamma_0-\delta}}\Bigg(\frac{\sum_{j=1}^{T}\pi_{j-1}^{(1)}(\gamma-\delta)\sum_{j=1}^{T}\pi_{j-1}(\gamma_0-\delta)\pi_{j-1}(\gamma-\delta)}{\sum_{j=1}^{T}\pi_{j-1}^2(\gamma-\delta)}\\ &+\frac{\sum_{j=1}^{T}\pi_{j-1}(\gamma-\delta)\sum_{j=1}^{T}\pi_{j-1}(\gamma_0-\delta)\pi_{j-1}^{(1)}(\gamma-\delta)}{\sum_{j=1}^{T}\pi_{j-1}^2(\gamma-\delta)}\\ &-\frac{2\sum_{j=1}^{T}\pi_{j-1}(\gamma-\delta)\sum_{j=1}^{T}\pi_{j-1}(\gamma_0-\delta)\pi_{j-1}(\gamma-\delta)\sum_{j=1}^{T}\pi_{j-1}(\gamma-\delta)\pi_{j-1}^{(1)}(\gamma-\delta)}{(\sum_{j=1}^{T}\pi_{j-1}^2(\gamma-\delta))^2}\Bigg)\\ &+g_{1,T}(\boldsymbol{\vartheta}),\end{aligned}\tag{S.8}$$

where $\sup_{\overline{\mathcal{H}}_i}|g_{1,T}(\boldsymbol{\vartheta})|=o(1)$. Now, substituting (83) into (S.8) (noting that the contribution of $\psi(d)$ cancels), approximating $\psi(d+j)$ by $\log j$, introducing $\log T$ terms (which cancel) and using (120) in Lemma 11, it can be shown that

$$\begin{aligned}\frac{1}{T^{\gamma_0-\delta}}\frac{\partial\overline{d}_T(\boldsymbol{\tau},\gamma)}{\partial\gamma}=-\frac{\mu_0\phi(1;\boldsymbol{\varphi})}{\Gamma(\gamma_0-\delta)T^{\gamma_0-\delta}}&\Bigg(\frac{\sum_{j=1}^{T}\log(j/T)j^{\gamma-\delta-1}\sum_{j=1}^{T}j^{\gamma_0+\gamma-2\delta-2}}{\sum_{j=1}^{T}j^{2\gamma-2\delta-2}}\\ &+\frac{\sum_{j=1}^{T}j^{\gamma-\delta-1}\sum_{j=1}^{T}\log(j/T)j^{\gamma_0+\gamma-2\delta-2}}{\sum_{j=1}^{T}j^{2\gamma-2\delta-2}}\\ &-\frac{2\sum_{j=1}^{T}j^{\gamma-\delta-1}\sum_{j=1}^{T}j^{\gamma_0+\gamma-2\delta-2}\sum_{j=1}^{T}\log(j/T)j^{2\gamma-2\delta-2}}{(\sum_{j=1}^{T}j^{2\gamma-2\delta-2})^2}\Bigg)+g_{2,T}(\boldsymbol{\vartheta}),\end{aligned}\tag{S.9}$$

where $\sup_{\overline{\mathcal{H}}_i}|g_{2,T}(\boldsymbol{\vartheta})|=o(1)$. Finally, (113) then follows by approximating sums by integrals, see Lemma 10.

Define $\overline{d}_T(\boldsymbol{\tau},\gamma)=\overline{d}_T(\boldsymbol{\vartheta})$. Because $\overline{d}_T(\boldsymbol{\tau},\gamma_0)=0$, the mean value theorem yields $\overline{d}_T(\boldsymbol{\tau},\gamma)=(\gamma-\gamma_0)\partial\overline{d}_T(\boldsymbol{\tau},\overline{\gamma})/\partial\gamma$, where $|\overline{\gamma}-\gamma_0|\leq|\gamma-\gamma_0|$, so the left-hand side of (114) can be bounded from below by

$$\lim_{T\to\infty} \inf_{\overline{\mathcal{H}}_i} T^{2\varkappa_i} (\gamma-\gamma_0)^2 \left(\frac{1}{T^{\gamma_0-\delta}} \frac{\partial \overline{d}_T(\boldsymbol{\tau},\overline{\gamma})}{\partial\gamma}\right)^2 \geq \lim_{T\to\infty} \inf_{\overline{\mathcal{H}}_i} T^{2\varkappa_i} (\gamma-\gamma_0)^2 \inf_{\overline{\mathcal{H}}_i} \left(\frac{1}{T^{\gamma_0-\delta}} \frac{\partial \overline{d}_T(\boldsymbol{\tau},\gamma)}{\partial\gamma}\right)^2$$
$$= \xi^2 \lim_{T\to\infty} \inf_{\overline{\mathcal{H}}_i} \left(\frac{1}{T^{\gamma_0-\delta}} \frac{\partial \overline{d}_T(\boldsymbol{\tau},\gamma)}{\partial\gamma}\right)^2.$$

Thus, setting ξ large enough, (114) follows if

$$\lim_{T\to\infty} \inf_{\overline{\mathcal{H}}_i} \left(\frac{1}{T^{\gamma_0-\delta}} \frac{\partial \overline{d}_T(\boldsymbol{\tau},\gamma)}{\partial\gamma}\right)^2 > \epsilon, \tag{S.10}$$

which, noting that $\gamma_0-\delta \geq 1+\eta$, is a consequence of (8) and (113) because

$$\inf_{\overline{\mathcal{H}}_i} \left(2(\gamma-\delta)^2 - 2(\gamma-\delta) + 1 - (\gamma_0-\delta)\right) = \inf_{\overline{\mathcal{H}}_i} \left(2(\gamma-\delta)^2 - 3(\gamma-\delta) + 1 - (\gamma_0-\gamma)\right)$$
$$\geq 2(\eta-\varrho)^2 + \eta - 2\varrho > 0.$$

S.2.4 Proof of Lemma 4

Recall the intervals $\mathcal{I}_i$ and define $\mathcal{W}_i = \{\boldsymbol{\vartheta} \in \overline{M}_\varepsilon : \delta \in \mathcal{I}_i\}$ for $i=1,2,3$, and $\mathcal{W}_4 = \{\boldsymbol{\vartheta} \in \overline{M}_\varepsilon : \delta \in \mathcal{I}_4 \cup \mathcal{I}_5\}$. Then the result follows on showing

$$\Pr\left(\inf_{\mathcal{W}_i} R_T(\boldsymbol{\vartheta}) \leq \sigma_0^2 + \epsilon\right) \to 0 \text{ as } T\to\infty \tag{S.11}$$

for $i=1,\ldots,4$, noting that

$$R_T(\boldsymbol{\vartheta}) = \frac{1}{T}\left(\sum_{t=1}^{T} (\phi(L;\boldsymbol{\varphi}) x_t(\delta))^2 - \left(\sum_{t=1}^{T} \phi(L;\boldsymbol{\varphi}) x_t(\delta) h_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi})\right)^2\right). \tag{S.12}$$

S.2.4.1 Proof of (S.11) for $i=4$

Given (97), we first apply the bound

$$\frac{1}{T}\sum_{t=1}^{T} (\phi(L;\boldsymbol{\varphi}) x_t(\delta))^2 \geq \frac{1}{T}\sum_{t=1}^{T} (\phi(L;\boldsymbol{\varphi}) u_t(\delta-\delta_0))^2$$
$$- \frac{2|\mu_0|}{T}\left|\sum_{t=1}^{T} \phi(L;\boldsymbol{\varphi}) u_t(\delta-\delta_0) c_{t-1}(\gamma_0-\delta,\boldsymbol{\varphi})\right| \tag{S.13}$$

and note that $\delta_0-\delta \leq 1/2-\eta$ when $\delta \in \mathcal{I}_4 \cup \mathcal{I}_5$, so $u_t(\delta-\delta_0)$ is asymptotically stationary. In view of (S.12) and (S.13), the proof of (S.11) for $i=4$ then follows by Hualde and Robinson (2011) (see the proof of their (2.7) for $i=4$) by showing

$$\sup_{\mathcal{W}_4} \frac{1}{T}\left|\sum_{t=1}^{T} \phi(L;\boldsymbol{\varphi}) u_t(\delta-\delta_0) c_{t-1}(\gamma_0-\delta,\boldsymbol{\varphi})\right| = o_p(1), \tag{S.14}$$

$$\sup_{\mathcal{W}_4} \frac{1}{T}\left(\sum_{t=1}^{T} \phi(L;\boldsymbol{\varphi}) x_t(\delta) h_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi})\right)^2 = o_p(1). \tag{S.15}$$

First, noting that $\delta \in \mathcal{I}_4 \cup \mathcal{I}_5$ implies $\delta_0 - \delta \leq 1/2 - \eta$ and $\gamma_0 - \delta \leq 1/2 + \gamma_0 - \delta_0 - \eta$, (S.95) of Lemma S.2 implies that the left-hand side of (S.14) is $O_p(T^{\gamma_0 - \delta_0 - 1/2 - 2\eta} + T^{-\varsigma - \eta} + T^{-1} \log^2 T) = o_p(1)$.

Next, using (97), (S.15) follows by showing

$$\sup_{\mathcal{W}_4} T^{-1/2} \sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) = o_p(1), \tag{S.16}$$

$$\sup_{\mathcal{W}_4} T^{-1/2} \sum_{t=1}^{T} c_{t-1}(\gamma_0 - \delta, \boldsymbol{\varphi}) h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) = o(1). \tag{S.17}$$

Here, (134) of Lemma 16 shows that the left-hand side of (S.16) is $O_p(T^{\theta - 1/2} + T^{-\eta}) = o_p(1)$ by choosing $\theta < 1/2$, while (S.92) of Lemma S.1 shows that the left-hand side of (S.17) $O(T^{\theta - 1/2} + T^{\gamma_0 - 1/2 - \delta_0 - \eta}) = o(1)$, to conclude the proof of (S.11) for $i = 4$.

S.2.4.2 Proof of (S.11) for $i = 3$

Noting (47), (S.12), and (S.13), the proof follows on showing

$$\sup_{\mathcal{W}_3} \frac{1}{T} \left| \sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) c_{t-1}(\gamma_0 - \delta, \boldsymbol{\varphi}) \right| = O_p(1), \tag{S.18}$$

$$\sup_{\mathcal{W}_3} \frac{1}{T} \left(\sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) x_t(\delta) h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) \right)^2 = O_p(1). \tag{S.19}$$

Both (S.18) and (S.19) follow straightforwardly by identical steps as those given in the proofs of (S.14) and (S.15) just replacing η by 0.

S.2.4.3 Proof of (S.11) for $i = 2$

Clearly,

$$\begin{aligned} \Pr\left(\inf_{\mathcal{W}_2} R_T(\boldsymbol{\vartheta}) \leq \sigma_0^2 + \epsilon \right) &\leq \Pr\left(\inf_{\mathcal{W}_2} \frac{T^{2(\delta_0 - \delta)}}{T} \inf_{\mathcal{W}_2} \frac{T}{T^{2(\delta_0 - \delta)}} R_T(\boldsymbol{\vartheta}) \leq \sigma_0^2 + \epsilon \right) \\ &= \Pr\left(\inf_{\mathcal{W}_2} \frac{T}{T^{2(\delta_0 - \delta)}} R_T(\boldsymbol{\vartheta}) \leq \sigma_0^2 + \epsilon \right). \end{aligned} \tag{S.20}$$

Thus, in view of (55), (S.12), and (S.13), (S.11) for $i = 2$ follows on showing

$$\sup_{\mathcal{W}_2} \frac{1}{T^{2(\delta_0 - \delta)}} \left| \sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) c_{t-1}(\gamma_0 - \delta, \boldsymbol{\varphi}) \right| = O_p(1), \tag{S.21}$$

$$\sup_{\mathcal{W}_2} \frac{1}{T^{2(\delta_0 - \delta)}} \left(\sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) x_t(\delta) h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) \right)^2 = O_p(1). \tag{S.22}$$

The proofs of (S.21) and (S.22) are almost identical to those of (S.14) and (S.15), taking into account the different normalization, which implies using (S.93) instead of (S.92) in Lemma S.1, (S.96) instead of (S.95) in Lemma S.2, and (135) instead of (134) in Lemma 16.

S.2.4.4 Proof of (S.11) for $i = 1$

Following identical steps to those given in (S.20),

$$\Pr\left(\inf_{\mathcal{W}_1} R_T(\boldsymbol{\vartheta}) \leq \sigma_0^2 + \epsilon\right) \leq \Pr\left(\inf_{\mathcal{W}_1} \frac{T}{T^{2(\delta_0-\delta)}} R_T(\boldsymbol{\vartheta}) \leq \frac{\sigma_0^2 + \epsilon}{T^{2\eta}}\right).$$

Letting $\alpha > 0$ be arbitrarily small (in particular $\alpha < (\varsigma - 1/2)/3$) and defining the sets $\Phi_1 = \{\boldsymbol{\vartheta} \in \Xi : \gamma - \delta \leq 1/2 - \alpha\}$, $\Phi_2 = \{\boldsymbol{\vartheta} \in \Xi : 1/2 - \alpha \leq \gamma - \delta \leq 1/2 + \alpha\}$, and $\Phi_3 = \{\boldsymbol{\vartheta} \in \Xi : \gamma - \delta \geq 1/2 + \alpha\}$, the required result follows on showing

$$\Pr\left(\inf_{\mathcal{W}_1 \cap \Phi_j} \frac{T}{T^{2(\delta_0-\delta)}} R_T(\boldsymbol{\vartheta}) > \epsilon\right) \to 1 \text{ as } T \to \infty \tag{S.23}$$

for $j = 1, 2, 3$. In the proof of their (2.7) for $i = 1$, Hualde and Robinson (2011) showed that

$$\Pr\left(\inf_{\|\boldsymbol{\tau}-\boldsymbol{\tau}_0\| \geq \varepsilon, \boldsymbol{\tau} \in \mathcal{T}_1} \frac{1}{T^{2(\delta_0-\delta)}} \sum_{t=1}^{T} \left(\phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0)\right)^2 > \epsilon\right) \to 1 \text{ as } T \to \infty,$$

so in view of (97), (S.12), and (S.13), (S.23) for $j = 1, 2$ holds if we show that

$$\sup_{\mathcal{W}_1} \frac{1}{T^{2(\delta_0-\delta)}} \sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) c_{t-1}(\gamma_0 - \delta, \boldsymbol{\varphi}) = o_p(1), \tag{S.24}$$

$$\sup_{\mathcal{W}_1} \frac{1}{T^{\delta_0-\delta}} \sum_{t=1}^{T} c_{t-1}(\gamma_0 - \delta, \boldsymbol{\varphi}) h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi}) = o(1), \tag{S.25}$$

$$\sup_{\mathcal{W}_1 \cap \Phi_j} \frac{1}{T^{2(\delta_0-\delta)}} \left(\sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) h_{t-1,T}(\gamma - \delta, \boldsymbol{\varphi})\right)^2 = o_p(1). \tag{S.26}$$

First, noting that $\delta_0 - \delta \geq 1/2 + \eta$ and $\gamma_0 - \delta \geq 1/2 + \eta + \gamma_0 - \delta_0$, by (S.96) of Lemma S.2 and (S.93) of Lemma S.1 with $\theta < 1/2 + \eta$, the left-hand sides of (S.24) and (S.25) are $O_p(T^{\max\{\gamma_0-\delta_0-1/2, -\varsigma-\eta\}} + T^{-1-2\eta}\log^2 T) = o_p(1)$ and $O(T^{\gamma_0-\delta_0-1/2} + T^{-1/2-\eta+\theta}) = o(1)$, respectively. Next, for $j = 1$, by (125) of Lemma 13, the left-hand side of (S.26) is

$$\sup_{\mathcal{W}_1 \cap \Phi_1} T^{2(\delta-\delta_0)} \left(\sum_{t=1}^{T} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) c_{t-1}(\gamma - \delta, \boldsymbol{\varphi})\right)^2,$$

which is easily shown to be $O_p(T^{-2\alpha}) = o_p(1)$ by (133) of Lemma 15 and (121) of Lemma 12. For $j = 2$, we use (17) to bound the left-hand side of (S.26) by

$$\frac{\sup_{\mathcal{W}_1 \cap \Phi_2} T \left(\sum_{t=1}^{T} T^{-(\delta_0-\delta)} \phi(L; \boldsymbol{\varphi}) u_t(\delta - \delta_0) T^{-(\gamma-\delta)} c_{t-1}(\gamma - \delta, \boldsymbol{\varphi})\right)^2}{\inf_{\mathcal{W}_1 \cap \Phi_2} T^{-2(\gamma-\delta)+1} \sum_{j=1}^{T} c_{j-1}^2(\gamma - \delta, \boldsymbol{\varphi})},$$

where the denominator can be made arbitrarily large by (126) of Lemma 13 and the numerator is easily seen to be $O_p(1)$ by direct application of (133) of Lemma 15 and (122) of Lemma 12.

We finally show (S.23) for $j = 3$. Using (29), applying (S.24) and (S.25) together with Lemmas 14, 15, we have

$$T^{1-2(\delta_0-\delta)} R_T(\boldsymbol{\vartheta}) \geq T^{-2(\delta_0-\delta)} \sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}) + q_{1,T}(\boldsymbol{\vartheta}),$$

where $\sup_{\mathcal{W}_1\cap\Phi_3} |q_{1,T}(\boldsymbol{\vartheta})| = o_p(1)$. Thus, (S.23) for $j = 3$ holds on showing

$$\Pr\left(\inf_{\mathcal{W}_1\cap\Phi_3} \frac{1}{T^{2(\delta_0-\delta)}} \sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}) > \epsilon\right) \to 1 \text{ as } T \to \infty. \tag{S.27}$$

First we show that

$$T^{-2(\delta_0-\delta)} \sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}) \geq \phi^2(1;\boldsymbol{\varphi})\,\omega^2(1;\boldsymbol{\varphi}_0)\, Z_T^2(\delta,\gamma) - |q_{2,T}(\boldsymbol{\vartheta})|, \tag{S.28}$$

where $\sup_{\mathcal{W}_1\cap\Phi_3} |q_{2,T}(\boldsymbol{\tau},\gamma_2)| = o_p(1)$. To prove (S.28) we first justify that

$$\begin{aligned}\sum_{t=1}^{T} s_t^2(\boldsymbol{\vartheta}) &= \sum_{t=1}^{T} \left(\phi(L;\boldsymbol{\varphi})\, u_t(\delta-\delta_0)\right)^2 - \left(\sum_{t=1}^{T} \phi(L;\boldsymbol{\varphi})\, u_t(\delta-\delta_0)\, h_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi})\right)^2 \\ &\geq \frac{\left(\sum_{t=1}^{T-2}\left(2\phi(L;\boldsymbol{\varphi})\, u_t(\delta-\delta_0-1) - \phi(L;\boldsymbol{\varphi})\, u_{T-1}(\delta-\delta_0-1)\right) c_t(\gamma-\delta,\boldsymbol{\varphi}) + q_{3,T}(\boldsymbol{\vartheta})\right)^2}{T^2 \sum_{t=1}^{T} c_{t-1}^2(\gamma-\delta,\boldsymbol{\varphi})},\end{aligned} \tag{S.29}$$

where $\sup_{\mathcal{W}_1\cap\Phi_3} T^{2\delta-\delta_0-\gamma-1/2} |q_{3,T}(\boldsymbol{\vartheta})| = o_p(1)$. To see this, first note that the equality in (S.29) follows because $\sum_{t=1}^{T} h_{t-1,T}^2(d,\boldsymbol{\varphi}) = 1$ and then note that left-hand side of the inequality in (S.29) is

$$\begin{aligned}&\sum_{t=1}^{T} \left(\phi(L;\boldsymbol{\varphi})\, u_t(\delta-\delta_0)\right)^2 \sum_{s=1}^{T} h_{s-1,T}^2(\gamma-\delta,\boldsymbol{\varphi}) - \left(\sum_{t=1}^{T} \phi(L;\boldsymbol{\varphi})\, u_t(\delta-\delta_0)\, h_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi})\right)^2 \\ &= \sum_{t=1}^{T-1}\sum_{k=t+1}^{T} \left(\phi(L;\boldsymbol{\varphi})\, u_t(\delta-\delta_0)\, h_{k-1,T}(\gamma-\delta,\boldsymbol{\varphi}) - \phi(L;\boldsymbol{\varphi})\, u_k(\delta-\delta_0)\, h_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi})\right)^2 \\ &\geq \frac{\left(\sum_{t=1}^{T-1}\sum_{k=t+1}^{T}\left(\phi(L;\boldsymbol{\varphi})\, u_t(\delta-\delta_0)\, h_{k-1,T}(\gamma-\delta,\boldsymbol{\varphi}) - \phi(L;\boldsymbol{\varphi})\, u_k(\delta-\delta_0)\, h_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi})\right)\right)^2}{T(T-1)/2}\end{aligned} \tag{S.30}$$

by the Cauchy-Schwarz inequality. By summation by parts, the contribution of the first term in the parenthesis on the right-hand side of (S.30) is

$$\begin{aligned}\sum_{t=1}^{T-1}\sum_{k=t+1}^{T} h_{k-1,T}(\gamma-\delta,\boldsymbol{\varphi})\,\phi(L;\boldsymbol{\varphi})\, u_t(\delta-\delta_0) &= h_{T-1,T}(\gamma-\delta,\boldsymbol{\varphi})\,\phi(L;\boldsymbol{\varphi})\, u_{T-1}(\delta-\delta_0-1) \\ &\quad + \sum_{t=1}^{T-2} \phi(L;\boldsymbol{\varphi})\, u_t(\delta-\delta_0-1)\, h_{t,T}(\gamma-\delta,\boldsymbol{\varphi}).\end{aligned} \tag{S.31}$$

The contribution of the second term in the parenthesis on the right-hand side of (S.30) is

$$\begin{aligned}\sum_{t=1}^{T-1}\sum_{k=t+1}^{T}&\phi\left(L;\boldsymbol{\varphi}\right)u_k\left(\delta-\delta_0\right)h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\\&=\sum_{t=2}^{T}\phi\left(L;\boldsymbol{\varphi}\right)u_t\left(\delta-\delta_0\right)\sum_{k=1}^{t-1}h_{k-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\\&=\sum_{t=2}^{T}\phi\left(L;\boldsymbol{\varphi}\right)u_t\left(\delta-\delta_0+1\right)\sum_{k=1}^{t-1}h_{k-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\\&\quad+\sum_{t=2}^{T}\phi\left(L;\boldsymbol{\varphi}\right)u_{t-1}\left(\delta-\delta_0\right)\sum_{k=1}^{t-1}h_{k-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right),\end{aligned}\tag{S.32}$$

and by summation by parts the second term on the right-hand side of (S.32) is

$$\begin{aligned}&\phi\left(L;\boldsymbol{\varphi}\right)u_{T-1}\left(\delta-\delta_0-1\right)\sum_{k=1}^{T-1}h_{k-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\\&\quad-\sum_{t=2}^{T-1}\phi\left(L;\boldsymbol{\varphi}\right)u_{t-1}\left(\delta-\delta_0-1\right)h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right).\end{aligned}\tag{S.33}$$

Thus, in view of (S.30)–(S.33), (17), and noting that $h_{0,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)=1$ and $T(T-1)/2\le T^2$, (S.29) holds with

$$\begin{aligned}q_{3,T}\left(\boldsymbol{\vartheta}\right)=&c_{T-1}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\phi\left(L;\boldsymbol{\varphi}\right)u_{T-1}\left(\delta-\delta_0-1\right)\\&-\sum_{t=2}^{T}\phi\left(L;\boldsymbol{\varphi}\right)u_t\left(\delta-\delta_0+1\right)\sum_{k=1}^{t-1}c_{k-1}\left(\gamma-\delta,\boldsymbol{\varphi}\right)-\phi\left(L;\boldsymbol{\varphi}\right)u_{T-1}\left(\delta-\delta_0-1\right),\end{aligned}$$

and consequently $\sup_{\mathcal{W}_1\cap\Phi_3}T^{2\delta-\delta_0-\gamma-1/2}\left|q_{3,T}\left(\boldsymbol{\vartheta}\right)\right|=O_p(T^{-1}+T^{-\alpha-1/2})=o_p(1)$ by application of (133) of Lemma 15 and (122) of Lemma 12. Next, by (67), (68), and (69),

$$\begin{aligned}&\sum_{t=1}^{T-2}\left(2\phi\left(L;\boldsymbol{\varphi}\right)u_t\left(\delta-\delta_0-1\right)-\phi\left(L;\boldsymbol{\varphi}\right)u_{T-1}\left(\delta-\delta_0-1\right)\right)c_t\left(\gamma-\delta,\boldsymbol{\varphi}\right)\\&=\phi^2\left(1;\boldsymbol{\varphi}\right)\omega\left(1;\boldsymbol{\varphi}_0\right)\sum_{t=1}^{T-2}\left(2\varepsilon_t\left(\delta-\delta_0-1\right)-\varepsilon_{T-1}\left(\delta-\delta_0-1\right)\right)\pi_t\left(\gamma-\delta\right)+q_{4,T}\left(\boldsymbol{\vartheta}\right),\end{aligned}\tag{S.34}$$

where

$$\begin{aligned}q_{4,T}\left(\boldsymbol{\vartheta}\right)=&\sum_{t=1}^{T-2}\left(2\phi\left(L;\boldsymbol{\varphi}\right)u_t\left(\delta-\delta_0-1\right)-\phi\left(L;\boldsymbol{\varphi}\right)u_{T-1}\left(\delta-\delta_0-1\right)\right)\\&\quad\times\left(c_{2,t+1}\left(\gamma-\delta,\boldsymbol{\varphi}\right)+c_{3,t+1}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right)\\&+\phi\left(1;\boldsymbol{\varphi}\right)\sum_{t=1}^{T-2}\left(2p_t\left(\boldsymbol{\vartheta}\right)-p_{T-1}\left(\boldsymbol{\vartheta}\right)\right)\pi_t\left(\gamma-\delta\right)\end{aligned}$$

and satisfies $\sup_{\mathcal{W}_1\cap\Phi_3} T^{2\delta-\delta_0-\gamma-1/2}\left|q_{4,T}\left(\boldsymbol{\vartheta}\right)\right| = o_p\left(1\right)$ by the same arguments as for (71).

By (S.29), (S.34), and (124) in Lemma 13, we then have the bound (S.28), where $\sup_{\mathcal{W}_1\cap\Phi_3}\left|q_{2,T}\left(\boldsymbol{\vartheta}\right)\right| = o_p\left(1\right)$ and

$$Z_T\left(\delta,\gamma\right) = \frac{1}{T^{\delta_0-\delta+1}}\frac{\sum_{t=1}^{T-2}\left(2\varepsilon_t\left(\delta-\delta_0-1\right)-\varepsilon_{T-1}\left(\delta-\delta_0-1\right)\right)\pi_t\left(\gamma-\delta\right)}{\left(\sum_{t=1}^{T}\pi_{t-1}^2(\gamma-\delta)\right)^{1/2}}.$$

Then considering $Z_T\left(\delta,\gamma\right)$ as a continuous process indexed by $\left(\delta,\gamma\right)$,

$$Z_T\left(\delta,\gamma\right) \Rightarrow Z\left(\delta,\gamma\right) \tag{S.35}$$

follows by exactly the same arguments as those in the proof of (73), where

$$Z\left(\delta,\gamma\right) = \left(2\left(\gamma-\delta\right)-1\right)^{1/2}\int_0^1 r^{\gamma-\delta-1}\left(2W\left(r;1+\delta_0-\delta\right)-W\left(1;1+\delta_0-\delta\right)\right)dr.$$

Then, noting Assumption A1(iv), (8) and the definition of $\mathcal{L}$ given in the proof of (21) for $i=1$, (S.27) follows on showing that

$$\Pr\left(\inf_{\mathcal{L}} Z_T^2\left(\delta,\gamma\right) > \epsilon\right) \to 1 \text{ as } T\to\infty. \tag{S.36}$$

By (S.35) and the continuous mapping theorem, $\inf_{\mathcal{L}} Z_T^2\left(\delta,\gamma\right) \to_d \inf_{\mathcal{L}} Z^2\left(\delta,\gamma\right)$ as $T\to\infty$, so that $\Pr\left(\inf_{\mathcal{L}} Z_T^2\left(\delta,\gamma\right) > \epsilon\right) \to \Pr\left(\inf_{\mathcal{L}} Z^2\left(\delta,\gamma\right) > \epsilon\right)$ as $T\to\infty$, and (S.36) follows because ϵ is arbitrarily small and $Z^2\left(\delta,\gamma\right)$ is the square of a Gaussian random variable. This completes the proof of (S.27) and therefore that of (S.11) for $i=1$.

S.2.5 Proof of Lemma 5

First we show the first equality in (115). By definition of $s_t\left(\boldsymbol{\vartheta}\right)$,

$$s_t\left(\boldsymbol{\vartheta}_0\right)-\varepsilon_t = -\sum_{j=t}^{\infty}\phi_j(\boldsymbol{\varphi}_0)u_{t-j} - h_{t-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right)\sum_{j=1}^{T}s_{1j}\left(\boldsymbol{\tau}_0\right)h_{j-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right), \tag{S.37}$$

where $s_{1j}\left(\boldsymbol{\tau}_0\right) = \varepsilon_j - \sum_{k=j}^{\infty}\phi_k(\boldsymbol{\varphi}_0)u_{j-k}$, so the result holds if

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}\sum_{j=t}^{\infty}\phi_j(\boldsymbol{\varphi}_0)u_{t-j}\frac{\partial s_{1t}\left(\boldsymbol{\tau}_0\right)}{\partial\boldsymbol{\tau}} = o_p\left(1\right), \tag{S.38}$$

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}h_{t-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right)\frac{\partial s_{1t}\left(\boldsymbol{\tau}_0\right)}{\partial\boldsymbol{\tau}}\sum_{j=1}^{T}s_{1j}\left(\boldsymbol{\tau}_0\right)h_{j-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right) = o_p\left(1\right). \tag{S.39}$$

First, for $t\geq 2$,

$$\frac{\partial s_{1t}\left(\boldsymbol{\tau}_0\right)}{\partial\delta} = -\sum_{j=1}^{t-1}\sum_{k=0}^{j-1}\phi_k(\boldsymbol{\varphi}_0)(j-k)^{-1}u_{t-j}, \tag{S.40}$$

so, noting (7) and applying Lemma 9,

$$E\left|\frac{\partial s_{1t}\left(\boldsymbol{\tau}_0\right)}{\partial\delta}\right| \leq K\sum_{j=1}^{t-1}\sum_{k=1}^{j-1}k^{-1-\varsigma}(j-k)^{-1} \leq K\sum_{j=1}^{t-1}j^{-1}\log j \leq K\log^2 t.$$

Similarly, by (12),

$$E\left\|\frac{\partial s_{1t}\left(\boldsymbol{\tau}_0\right)}{\partial \boldsymbol{\varphi}}\right\| = E\left\|\sum_{j=1}^{t-1}\frac{\partial \phi_j(\boldsymbol{\varphi}_0)}{\partial \boldsymbol{\varphi}}u_{t-j}\right\| = O\left(1\right), \tag{S.41}$$

so that

$$E\left\|\frac{\partial s_{1t}\left(\boldsymbol{\tau}_0\right)}{\partial \boldsymbol{\tau}}\right\| = O\left(\log^2 t\right).$$

Thus, noting (107),

$$E\left\|\frac{1}{T^{1/2}}\sum_{t=1}^{T}\sum_{j=t}^{\infty}\phi_j(\boldsymbol{\varphi}_0)u_{t-j}\frac{\partial s_{1t}\left(\boldsymbol{\tau}_0\right)}{\partial \boldsymbol{\tau}}\right\| \leq \frac{K}{T^{1/2}}\sum_{t=1}^{T}t^{-1/2-\varsigma}\log^2 t \leq KT^{-1/2} = o\left(1\right)$$

because $\varsigma > 1/2$, which proves (S.38). Next, by (127) of Lemma 14 and (107), it is straightforward to show that

$$\begin{aligned}
&\sum_{j=1}^{T}s_{1j}\left(\boldsymbol{\tau}_0\right)h_{j-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right) \\
&\qquad = \sum_{j=1}^{T}(\varepsilon_j - \sum_{k=j}^{\infty}\phi_k(\boldsymbol{\varphi}_0)u_{j-k})h_{j-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right) = O_p(T^{\theta}),
\end{aligned} \tag{S.42}$$

so (S.39) holds on showing that

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}h_{t-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right)\frac{\partial s_{1t}\left(\boldsymbol{\tau}_0\right)}{\partial \boldsymbol{\tau}} = O_p\left(T^{\theta-1/2}\log T\right) \tag{S.43}$$

and setting $\theta < 1/4$. Noting (S.40), by simple calculations,

$$\frac{\partial s_{1t}\left(\boldsymbol{\tau}_0\right)}{\partial \delta} = -\sum_{j=1}^{t-1}\frac{1}{j}\varepsilon_{t-j} + \sum_{j=1}^{t-1}\frac{1}{j}\sum_{k=t-j}^{\infty}\phi_k(\boldsymbol{\varphi}_0)u_{t-j-k}. \tag{S.44}$$

The contribution of the first term on the right-hand side of (S.44) to the left-hand side of (S.43) is, by (127) of Lemma 14,

$$\begin{aligned}
-\frac{1}{T^{1/2}}\sum_{t=2}^{T}\sum_{j=1}^{t-1}j^{-1}\varepsilon_{t-j}h_{t-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right) &= -\frac{1}{T^{1/2}}\sum_{t=1}^{T-1}\varepsilon_t\sum_{j=1}^{T-t}j^{-1}h_{t+j-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right) \\
&= O_p\left(\frac{T^{\theta}}{T^{1/2}}\left(\sum_{t=1}^{T-1}\left(\sum_{j=1}^{T-t}\frac{1}{j}\left(t+j\right)^{-1/2-\theta}\right)^2\right)^{1/2}\right) \\
&= O_p(T^{\theta-1/2}\log T)
\end{aligned} \tag{S.45}$$

because

$$\sum_{j=1}^{T-t}j^{-1}\left(t+j\right)^{-1/2-\theta} \leq t^{-1/2-\theta}\sum_{j=1}^{T}j^{-1} \leq Kt^{-1/2-\theta}\log T.$$

Similarly, the contribution of the second term on the right-hand side of (S.44) to the left-hand side of (S.43) is

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}h_{t-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right)\sum_{j=1}^{t-1}j^{-1}\sum_{k=t-j}^{\infty}\phi_k(\boldsymbol{\varphi}_0)u_{t-j-k},$$

which can be easily shown to be $O_p\left(T^{\theta-1/2}\log T\right)$ by (127) of Lemma 14, (107), and Lemma 9. Next, the derivative with respect to $\boldsymbol{\varphi}$ on the left-hand side of (S.43) is

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T-1}u_t\sum_{j=1}^{T-t}\frac{\partial\phi_j(\boldsymbol{\varphi}_0)}{\partial\boldsymbol{\varphi}}h_{t+j-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right),$$

which, by very similar arguments to (S.45), can easily be shown to be $O_p\left(T^{\theta-1/2}\right)$ by (12) and (127) of Lemma 14, to conclude the proof of (S.43) and hence of the first equality in (115).

Next, because $\sum_{t=1}^{T}c_{t-1}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right)\partial d_t\left(\boldsymbol{\vartheta}_0\right)/\partial\gamma=0$, the proof of the second equality in (115) follows by showing that

$$\frac{1}{T^{\gamma_0-\delta_0-1/2}}\sum_{t=1}^{T}\sum_{j=t}^{\infty}\phi_j(\boldsymbol{\varphi}_0)u_{t-j}\frac{\partial d_t\left(\boldsymbol{\vartheta}_0\right)}{\partial\gamma}=o_p\left(1\right),$$

which, noting the proof of (82) and (107), follows easily by previous arguments.

The proofs of the two equalities in (116) are almost identical, but the second is simpler, so we show only the first. By (S.37), the first equality in (116) holds if

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}s_{1t}\left(\boldsymbol{\tau}_0\right)\frac{\partial s_{2t}\left(\boldsymbol{\vartheta}_0\right)}{\partial\boldsymbol{\tau}}=o_p\left(1\right), \tag{S.46}$$

$$\frac{1}{T^{1/2}}\sum_{t=1}^{T}h_{t-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right)\frac{\partial s_{2t}\left(\boldsymbol{\vartheta}_0\right)}{\partial\boldsymbol{\tau}}\sum_{j=1}^{T}s_{1j}\left(\boldsymbol{\tau}_0\right)h_{j-1,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right)=o_p\left(1\right). \tag{S.47}$$

As defined before, $s_{2t}(\boldsymbol{\vartheta})=h_{t-1,T}(\gamma-\delta,\boldsymbol{\varphi})\sum_{j=1}^{T}s_{1j}(\boldsymbol{\tau})h_{j-1,T}(\gamma-\delta,\boldsymbol{\varphi})$ so that

$$\begin{aligned}\frac{\partial s_{2t}\left(\boldsymbol{\vartheta}\right)}{\partial\boldsymbol{\tau}}&=\frac{\partial h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)}{\partial\boldsymbol{\tau}}\sum_{j=1}^{T}s_{1j}\left(\boldsymbol{\tau}\right)h_{j-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\\&\quad+h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\sum_{j=1}^{T}\frac{\partial s_{1j}\left(\boldsymbol{\tau}\right)}{\partial\boldsymbol{\tau}}h_{j-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\\&\quad+h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\sum_{j=1}^{T}s_{1j}\left(\boldsymbol{\tau}\right)\frac{\partial h_{j-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)}{\partial\boldsymbol{\tau}}.\end{aligned} \tag{S.48}$$

First, given that $\gamma_0-\delta_0>1/2$, setting $\theta<\gamma_0-\delta_0-1/2$, by a simple modification of the proof of (128) of Lemma 14,

$$\left\|\frac{\partial h_{t,T}\left(\gamma_0-\delta_0,\boldsymbol{\varphi}_0\right)}{\partial\boldsymbol{\tau}}\right\|=O\left(t^{-1/2}\left(\frac{T}{t}\right)^{\theta}\log T\right). \tag{S.49}$$

Then noting (S.42), (S.43), and by application of Lemma 14, it follows that

$$\frac{\partial s_{2t}\left(\boldsymbol{\vartheta}_0\right)}{\partial \boldsymbol{\tau}} = O_p(t^{-1/2-\theta}T^{2\theta}\log T). \tag{S.50}$$

By (S.50) and (107), it follows that the left-hand side of (S.46) is $O_p(T^{2\theta-1/2}\log T) = o_p(1)$ by setting $\theta < 1/4$. Similarly, by (127) of Lemma 14, (S.42), and (S.50), the left-hand side of (S.47) is $O_p(T^{4\theta-1/2}\log T) = o_p(1)$ by setting $\theta < 1/8$. This concludes the proof of the first equality in (116).

Finally, the proofs for the results in part (b) are heavily based on the arguments employed in the proofs of (115) and (116), and are therefore omitted.

S.2.6 Proof of Lemma 6

As in the proof of Theorem 1(i), noting (15), (16), (18), (75), the result holds on establishing that

$$\Pr\left(\inf_{\boldsymbol{\vartheta}\in\overline{M}_{\varepsilon}^{*}} S_T\left(\boldsymbol{\vartheta}\right) \le 0\right) \to 0 \text{ as } T \to \infty, \tag{S.51}$$

$$\Pr\left(\widehat{\boldsymbol{\vartheta}} \in \overline{N}_{\varepsilon}^{*} \cap M_{\varepsilon}^{*}, \inf_{\overline{N}_{\varepsilon}^{*}\cap M_{\varepsilon}^{*}} R_T(\widehat{\boldsymbol{\tau}},\gamma) - R_T(\widehat{\boldsymbol{\tau}},\gamma_0) \le 0\right) \to 0 \text{ as } T \to \infty, \tag{S.52}$$

where

$$M_{\varepsilon}^{*} = \left\{\boldsymbol{\vartheta} \in \Xi : \|\boldsymbol{\tau}-\boldsymbol{\tau}_0\| < \varepsilon T^{-\varkappa}\right\}, \quad \overline{M}_{\varepsilon}^{*} = \{\boldsymbol{\vartheta} \in \Xi : \varepsilon T^{-\varkappa} \le \|\boldsymbol{\tau}-\boldsymbol{\tau}_0\| < \varepsilon\},$$
$$N_{\varepsilon}^{*} = \left\{\boldsymbol{\vartheta} \in \Xi : |\gamma-\gamma_0| < \varepsilon T^{-\varkappa}\right\}, \quad \overline{N}_{\varepsilon}^{*} = \{\boldsymbol{\vartheta} \in \Xi : \varepsilon T^{-\varkappa} \le |\gamma-\gamma_0| < \varepsilon\}.$$

We first prove (S.51), which, defining $\mathcal{J}_i = \{\boldsymbol{\vartheta} \in \overline{M}_{\varepsilon}^{*} : \delta \in \mathcal{I}_i\}$, $i = 4, 5$, holds if

$$\Pr\left(\inf_{\mathcal{J}_i} S_T\left(\boldsymbol{\vartheta}\right) \le 0\right) \to 0 \text{ as } T \to \infty \tag{S.53}$$

for $i = 4, 5$. Note here that $\boldsymbol{\vartheta} \in \overline{M}_{\varepsilon}^{*}$ implies $\|\boldsymbol{\tau}-\boldsymbol{\tau}_0\| < \varepsilon$, so necessarily $\delta \in \mathcal{I}_4 \cup \mathcal{I}_5$ and there is no need to consider the intervals $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$. Clearly, (S.53) for $i = 5$ would hold if

$$\Pr\left(\inf_{\mathcal{J}_5} T^{2\varkappa} S_T\left(\boldsymbol{\vartheta}\right) \le 0\right) \to 0 \text{ as } T \to \infty. \tag{S.54}$$

Proceeding as in the proof of (19)–(21) for $i = 5$, (S.54) holds if

$$\inf_{\mathcal{J}_5} T^{2\varkappa} U\left(\boldsymbol{\tau}\right) > \epsilon, \tag{S.55}$$

$$\frac{1}{T^{1-2\varkappa}} \sum_{t=1}^{T} \left[\left(\phi(L;\boldsymbol{\varphi}_0)\left\{u_t \mathbb{I}(t>0)\right\}\right)^2 - \sigma_0^2\right] = o_p\left(1\right), \tag{S.56}$$

$$\sup_{\mathcal{J}_5} \frac{1}{T^{1-2\varkappa}} \sum_{t=1}^{T} \left[\left(\phi(L;\boldsymbol{\varphi}) u_t\left(\delta-\delta_0\right)\right)^2 - E\left(\left(\phi(L;\boldsymbol{\varphi})\Delta^{\delta-\delta_0} u_t\right)^2\right)\right] = o_p\left(1\right), \tag{S.57}$$

$$\sup_{\mathcal{J}_5} \frac{1}{T^{1-2\varkappa}} \left|\sum_{t=1}^{T} d_t\left(\boldsymbol{\vartheta}\right) s_t\left(\boldsymbol{\vartheta}\right)\right| = o_p\left(1\right), \tag{S.58}$$

$$\sup_{\mathcal{J}_5} \frac{1}{T^{1-2\varkappa}} \left(\sum_{t=1}^{T} \phi(L;\boldsymbol{\varphi}) u_t\left(\delta-\delta_0\right) h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right)^2 = o_p\left(1\right). \tag{S.59}$$

First, we justify (S.55). Clearly

$$U(\boldsymbol{\tau}) = \sigma_0^2 \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left|\phi\left(e^{i\lambda};\boldsymbol{\varphi}\right)\right|^2}{\left|\phi\left(e^{i\lambda};\boldsymbol{\varphi}_0\right)\right|^2} \left|1-e^{i\lambda}\right|^{2(\delta-\delta_0)} d\lambda - 1 \right),$$

and we show that $U(\boldsymbol{\tau})$ is a strictly convex function at $\boldsymbol{\tau}_0$ with a strict local minimum at $\boldsymbol{\tau}=\boldsymbol{\tau}_0$. Noting that

$$\int_{-\pi}^{\pi} \frac{e^{iq\lambda}}{\phi\left(e^{i\lambda};\boldsymbol{\varphi}_0\right)} d\lambda = 0 \text{ for any } q = \pm 1, \pm 2, ..., \tag{S.60}$$

and $\int_{-\pi}^{\pi} \log(2-2\cos\lambda)\, d\lambda = 0$, it is straightforward to show that $\partial U(\boldsymbol{\tau}_0)/\partial\boldsymbol{\tau} = 0$. Similarly, using again (S.60),

$$\begin{aligned}\frac{\partial^2 U(\boldsymbol{\tau}_0)}{\partial\boldsymbol{\tau}\partial\boldsymbol{\tau}'} &= \begin{pmatrix} \int_{-\pi}^{\pi} \log^2(2-2\cos\lambda)\, d\lambda & 2\int_{-\pi}^{\pi} \frac{\partial\phi(e^{i\lambda};\boldsymbol{\varphi}_0)/\partial\boldsymbol{\varphi}'}{\phi(e^{i\lambda};\boldsymbol{\varphi}_0)} \log(2-2\cos\lambda)\, d\lambda \\ 2\int_{-\pi}^{\pi} \frac{\partial\phi(e^{i\lambda};\boldsymbol{\varphi}_0)/\partial\boldsymbol{\varphi}}{\phi(e^{i\lambda};\boldsymbol{\varphi}_0)} \log(2-2\cos\lambda)\, d\lambda & 2\int_{-\pi}^{\pi} \frac{\partial\phi(e^{i\lambda};\boldsymbol{\varphi}_0)/\partial\boldsymbol{\varphi}\,\partial\phi(e^{i\lambda};\boldsymbol{\varphi}_0)/\partial\boldsymbol{\varphi}'}{\left|\phi(e^{i\lambda};\boldsymbol{\varphi}_0)\right|^2} d\lambda \end{pmatrix} \\ &= \begin{pmatrix} 2\pi^3/3 & -4\pi\sum_{j=1}^{\infty} \boldsymbol{b}_j'(\boldsymbol{\varphi}_0)/j \\ -4\pi\sum_{j=1}^{\infty} \boldsymbol{b}_j(\boldsymbol{\varphi}_0)/j & 4\pi\sum_{j=1}^{\infty} \boldsymbol{b}_j(\boldsymbol{\varphi}_0)\boldsymbol{b}_j'(\boldsymbol{\varphi}_0) \end{pmatrix},\end{aligned}$$

which by A4(iii) is positive definite, to complete the proof of strict convexity of $U(\boldsymbol{\tau})$ at $\boldsymbol{\tau}_0$. Thus, by continuity there exists a point $\boldsymbol{\tau}^*$ such that $\|\boldsymbol{\tau}_0 - \boldsymbol{\tau}^*\| = \varepsilon T^{-\varkappa}$ and $\inf_{\mathcal{J}_5} U(\boldsymbol{\tau}) = U(\boldsymbol{\tau}^*)$. Then, noting that $U(\boldsymbol{\tau}_0) = 0$ and $\partial U(\boldsymbol{\tau}_0)/\partial\boldsymbol{\tau} = 0$, by Taylor's expansion,

$$U(\boldsymbol{\tau}^*) \geq \frac{1}{2}(\boldsymbol{\tau}^* - \boldsymbol{\tau}_0)' \frac{\partial^2 U(\boldsymbol{\tau}_0)}{\partial\boldsymbol{\tau}\partial\boldsymbol{\tau}'}(\boldsymbol{\tau}^* - \boldsymbol{\tau}_0) - |w_T|, \tag{S.61}$$

where it can be shown that $w_T = O(T^{-3\varkappa})$. Here, the main issue is to justify that the third derivative of $U(\boldsymbol{\tau})$ evaluated at an arbitrarily small neighborhood of $\boldsymbol{\tau}_0$ is bounded, but this follows straightforwardly from A4(ii). Additionally,

$$(\boldsymbol{\tau}^* - \boldsymbol{\tau}_0)' \frac{\partial^2 U(\boldsymbol{\tau}_0)}{\partial\boldsymbol{\tau}\partial\boldsymbol{\tau}'}(\boldsymbol{\tau}^* - \boldsymbol{\tau}_0) \geq \underline{\lambda} \|\boldsymbol{\tau}^* - \boldsymbol{\tau}_0\|^2,$$

where $\underline{\lambda}$ denotes the minimum eigenvalue of the matrix $\partial^2 U(\boldsymbol{\tau}_0)/\partial\boldsymbol{\tau}\partial\boldsymbol{\tau}'$, so by (S.61) and noting that $\underline{\lambda}$ is strictly positive, for a sufficiently small $\epsilon > 0$,

$$U(\boldsymbol{\tau}^*) > \frac{\epsilon}{\varepsilon^2} \|\boldsymbol{\tau}^* - \boldsymbol{\tau}_0\|^2,$$

which justifies (S.55). The proofs of (S.56)–(S.59) are omitted as, for small enough $\varkappa$, they follow by almost identical arguments to those of (24)–(27).

Next, the proof of (S.53) for $i = 4$ is omitted because it is basically identical to those of (19)–(21) for $i = 4$. The only difference is that now $\varepsilon T^{-\varkappa} \leq \|\boldsymbol{\tau} - \boldsymbol{\tau}_0\| \leq \varepsilon$ instead of $\|\boldsymbol{\tau} - \boldsymbol{\tau}_0\| \geq \varepsilon$, but this does not make any difference. This completes the justification of (S.51).

Finally, we prove (S.52). For the same reason as in the proof of (75), we need to prove that

$$\inf_{\boldsymbol{\vartheta}\in\overline{N}_{\varepsilon}^{*}\cap M_{\varepsilon}^{*}} \frac{T^{2\varkappa}}{T^{2(\gamma_0-\delta)-1}} \sum_{t=1}^{T} d_t^2\left(\boldsymbol{\vartheta}\right) > \epsilon, \tag{S.62}$$

$$\sup_{\boldsymbol{\vartheta}\in\overline{N}_{\varepsilon}^{*}\cap M_{\varepsilon}^{*}} \frac{T^{2\varkappa}}{T^{2(\gamma_0-\delta)-1}} \left|\sum_{t=1}^{T} d_t\left(\boldsymbol{\vartheta}\right) s_{1t}\left(\boldsymbol{\tau}\right)\right| = o_p(1), \tag{S.63}$$

$$\sup_{\boldsymbol{\vartheta}\in\overline{N}_{\varepsilon}^{*}\cap M_{\varepsilon}^{*}} \frac{T^{2\varkappa}}{T^{2(\gamma_0-\delta)-1}} \left(\sum_{j=1}^{T} s_{1j}\left(\boldsymbol{\tau}\right) h_{j-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right)^2 = o_p(1). \tag{S.64}$$

As in (34), the proof of (S.62) follows by Lemma 3, whereas the proofs of (S.63) and (S.64) hold as in (77) and (78) for $\varkappa > 0$ sufficiently small.

S.2.7 Proof of Lemma 7

First we show (117). Clearly

$$\begin{aligned} c_t(\widehat{\delta},\widehat{\boldsymbol{\varphi}}) - c_t(\delta_0,\boldsymbol{\varphi}_0) &= \sum_{j=0}^{t}\left(\phi_j\left(\widehat{\boldsymbol{\varphi}}\right)-\phi_j\left(\boldsymbol{\varphi}_0\right)\right)\pi_{t-j}\left(\delta_0\right) + \sum_{j=0}^{t}(\pi_{t-j}(\widehat{\delta})-\pi_{t-j}\left(\delta_0\right))\phi_j\left(\boldsymbol{\varphi}_0\right) \\ &+ \sum_{j=0}^{t}\left(\phi_j\left(\widehat{\boldsymbol{\varphi}}\right)-\phi_j\left(\boldsymbol{\varphi}_0\right)\right)(\pi_{t-j}(\widehat{\delta})-\pi_{t-j}\left(\delta_0\right)). \end{aligned} \tag{S.65}$$

Fix $\epsilon < 1/2$. Then

$$\phi_j\left(\widehat{\boldsymbol{\varphi}}\right)-\phi_j\left(\boldsymbol{\varphi}_0\right) = \left(\phi_j\left(\widehat{\boldsymbol{\varphi}}\right)-\phi_j\left(\boldsymbol{\varphi}_0\right)\right)\left(\mathbb{I}\left(\|\widehat{\boldsymbol{\varphi}}-\boldsymbol{\varphi}_0\|<\epsilon\right)+\mathbb{I}\left(\|\widehat{\boldsymbol{\varphi}}-\boldsymbol{\varphi}_0\|\geq\epsilon\right)\right), \tag{S.66}$$

so by the mean value theorem the left-hand side of (S.66) is bounded by

$$\sup_{\|\boldsymbol{\varphi}-\boldsymbol{\varphi}_0\|<\epsilon}\left\|\frac{\partial\phi_j\left(\boldsymbol{\varphi}\right)}{\partial\boldsymbol{\varphi}}\right\|\|\widehat{\boldsymbol{\varphi}}-\boldsymbol{\varphi}_0\| + K\sup_{\boldsymbol{\varphi}\in\Psi}\left|\phi_j\left(\boldsymbol{\varphi}\right)\right|\frac{\|\widehat{\boldsymbol{\varphi}}-\boldsymbol{\varphi}_0\|^N}{\epsilon^N}, \tag{S.67}$$

for any arbitrarily large fixed number N. Then by (7) and the $T^{-\varkappa}$-consistency of $\widehat{\boldsymbol{\tau}}$, the second term in (S.67) is of smaller order, whereas by (12), the first one is $O_p(T^{-\varkappa}j^{-1-\varsigma})$. This implies that the first term on the right-hand side of (S.65) is $O_p(T^{-\varkappa}t^{\max\{\delta_0-1,-1-\varsigma\}}\log t)$ by Lemmas 9 and 11. Next, by straightforward application of Lemma 11 and a second-order Taylor expansion, $\pi_{t-j}(\widehat{\delta})-\pi_{t-j}\left(\delta_0\right) = O_p((t-j)^{\delta_0-1}\left(\log t\right)T^{-\varkappa})$, so by (7) and Lemma 9 the second term on the right-hand side of (S.65) is $O_p(T^{-\varkappa}t^{\max\{\delta_0-1,-1-\varsigma\}}\log^2 t)$. Finally, combining the arguments for the first two terms, the third term on the right-hand side of (S.65) is of smaller order, to conclude the proof of (117). The proof of (118) is omitted because it is almost identical to that of (117) with the only difference that the coefficients $\pi_{t-j}^{(1)}(\cdot)$ instead of $\pi_{t-j}(\cdot)$ lead to an extra $\log t$ factor, see Lemma 11. Finally, we show (119). Clearly the left-hand side of (119) is

$$\sum_{j=0}^{t-1}\phi_j\left(\boldsymbol{\varphi}_0\right)u_{t-j} + \sum_{j=0}^{t-1}\left(\phi_j\left(\widehat{\boldsymbol{\varphi}}\right)-\phi_j\left(\boldsymbol{\varphi}_0\right)\right)u_{t-j} + \sum_{j=0}^{t-1}\phi_j\left(\widehat{\boldsymbol{\varphi}}\right)(u_{t-j}(\widehat{\delta}-\delta_0)-u_{t-j}). \tag{S.68}$$

Using the mean value theorem as in (S.66) and (S.67) and summation by parts, it can be shown that the second term in (S.68) is $O_p\left(T^{-\varkappa}\right)$. Similarly, by Lemma C.5 of Robinson and Hualde (2003) and (7), the third term in (S.68) is also $O_p\left(T^{-\varkappa}\right)$, to conclude the proof of (119).

S.2.8 Proof of Lemma 8

First, for any $\epsilon > 0$, clearly

$$\begin{aligned}\Pr\left(\left\|T^{1/2}\partial S_T(\widehat{\boldsymbol{\vartheta}})/\partial\boldsymbol{\tau}\right\|\geq\epsilon\right) &= \Pr\left(\left\|T^{1/2}\partial S_T(\widehat{\boldsymbol{\vartheta}})/\partial\boldsymbol{\tau}\right\|\geq\epsilon, \|\widehat{\boldsymbol{\tau}}-\boldsymbol{\tau}_0\|<\varepsilon\right)\\ &\quad + \Pr\left(\left\|T^{1/2}\partial S_T(\widehat{\boldsymbol{\vartheta}})/\partial\boldsymbol{\tau}\right\|\geq\epsilon, \|\widehat{\boldsymbol{\tau}}-\boldsymbol{\tau}_0\|\geq\varepsilon\right)\\ &\leq \Pr\left(\left\|T^{1/2}\partial S_T(\widehat{\boldsymbol{\vartheta}})/\partial\boldsymbol{\tau}\right\|\geq\epsilon, \|\widehat{\boldsymbol{\tau}}-\boldsymbol{\tau}_0\|<\varepsilon\right)\\ &\quad + \Pr\left(\|\widehat{\boldsymbol{\tau}}-\tau_0\|\geq\varepsilon\right),\end{aligned}$$

so, in view of Theorem 1(ii) and (109), the result holds on showing

$$\sup_{\boldsymbol{\vartheta}\in M_\varepsilon}\frac{1}{T^{1/2}}\sum_{t=1}^{T}\left(d_t(\boldsymbol{\vartheta})-s_{2t}(\boldsymbol{\vartheta})\right)\left(\frac{\partial d_t(\boldsymbol{\vartheta})}{\partial\boldsymbol{\tau}}-\frac{\partial s_{2t}(\boldsymbol{\vartheta})}{\partial\boldsymbol{\tau}}\right)=o_p(1), \tag{S.69}$$

$$\sup_{\boldsymbol{\vartheta}\in M_\varepsilon}\frac{1}{T^{1/2}}\sum_{t=1}^{T}s_{1t}(\boldsymbol{\tau})\left(\frac{\partial d_t(\boldsymbol{\vartheta})}{\partial\boldsymbol{\tau}}-\frac{\partial s_{2t}(\boldsymbol{\vartheta})}{\partial\boldsymbol{\tau}}\right)=o_p(1), \tag{S.70}$$

$$\sup_{\boldsymbol{\vartheta}\in M_\varepsilon}\frac{1}{T^{1/2}}\sum_{t=1}^{T}\frac{\partial s_{1t}(\boldsymbol{\tau})}{\partial\boldsymbol{\tau}}\left(d_t(\boldsymbol{\vartheta})-s_{2t}(\boldsymbol{\vartheta})\right)=o_p(1). \tag{S.71}$$

The proof of (S.69) follows upon showing that, for any $\theta > 0$ and ε such that $0 < \varepsilon < \theta$,

$$\sup_{\boldsymbol{\vartheta}\in M_\varepsilon}|d_t(\boldsymbol{\vartheta})| = O(t^{\max\{\gamma_0-\delta_0+\varepsilon-1,-1-\varsigma\}}+T^{2\theta}t^{-1/2-\theta}), \tag{S.72}$$

$$\sup_{\boldsymbol{\vartheta}\in M_\varepsilon}|s_{2t}(\boldsymbol{\vartheta})| = O_p(T^{2\theta}t^{-1/2-\theta}), \tag{S.73}$$

$$\sup_{\boldsymbol{\vartheta}\in M_\varepsilon}\left\|\frac{\partial d_t(\boldsymbol{\vartheta})}{\partial\boldsymbol{\tau}}\right\| = O(t^{\max\{\gamma_0-\delta_0+\varepsilon-1,-1-\varsigma\}}\log t+T^{4\theta}t^{-1/2-\theta}), \tag{S.74}$$

$$\sup_{\boldsymbol{\vartheta}\in M_\varepsilon}\left\|\frac{\partial s_{2t}(\boldsymbol{\vartheta})}{\partial\boldsymbol{\tau}}\right\| = O_p(T^{4\theta}t^{-1/2-\theta}), \tag{S.75}$$

and then letting θ be sufficiently small. We only show (S.74) and (S.75) because the proofs for (S.72) and (S.73) are very similar but simpler. First, by (121) of Lemma 12,

$$\sup_{\boldsymbol{\vartheta}\in M_\varepsilon}|c_t(\gamma_0-\delta,\boldsymbol{\varphi})| = O(t^{\max\{\gamma_0-\delta_0+\varepsilon-1,-1-\varsigma\}}), \tag{S.76}$$

and by a simple modification of that result,

$$\sup_{\boldsymbol{\vartheta}\in M_\varepsilon}\left\|\frac{\partial c_t(\gamma_0-\delta,\boldsymbol{\varphi})}{\partial\boldsymbol{\tau}}\right\| = O(t^{\max\{\gamma_0-\delta_0+\varepsilon-1,-1-\varsigma\}}\log t). \tag{S.77}$$

Then (S.74) follows by direct application of (127) of Lemma 14, (S.76) and (S.77), noting that the bound in (127) also applies if the derivative is taken with respect to $\boldsymbol{\tau}$.

To prove (S.75) we apply (S.48), where the $\sup_{\boldsymbol{\vartheta}\in M_\varepsilon}$ of the absolute values of the first and third terms on the right-hand side are $O_p(T^{4\theta}t^{-1/2-\theta})$ by direct application of (127) of Lemma 14 and (134) of Lemma 16, noting that $\delta_0-\delta\leq\varepsilon$ and that these bounds also apply if the derivatives are taken with respect to $\boldsymbol{\tau}$. For the second term on the right-hand side of (S.48), noting that $\sum_{l=1}^{t} s_{1l}(\boldsymbol{\tau}) = \sum_{j=0}^{t-1} c_j\left(\delta_0 - \delta, \boldsymbol{\varphi}\right) \sum_{l=1}^{t-j} u_l$, it is straightforward to show that, by (121) of Lemma 12,

$$\sup_{\boldsymbol{\vartheta} \in M_\varepsilon} \left\| \sum_{j=1}^{t} \frac{\partial s_{1j}(\boldsymbol{\tau})}{\partial \boldsymbol{\tau}} \right\| = O_p(t^{1/2+\varepsilon} \log t). \tag{S.78}$$

Therefore, using summation by parts as in the proof of Lemma 16 and by (127) of Lemma 14, the $\sup_{\boldsymbol{\vartheta} \in M_\varepsilon}$ of the absolute value of the second term on the right-hand side of (S.48) is $O_p(T^{2\theta} t^{-1/2-\theta})$, to justify (S.75) and hence (S.69).

Finally, (S.70) and (S.71) can be established by using summation by parts followed by direct application of the results in (S.72), (S.74), (S.78), and Lemma 15, noting also that by previous arguments it can be easily shown that

$$\sup_{\boldsymbol{\vartheta} \in M_\varepsilon} |d_{t+1}(\boldsymbol{\vartheta}) - d_t(\boldsymbol{\vartheta})| = O(t^{\max\{\gamma_0 - \delta_0 + \varepsilon - 2, -1-\varsigma\}} + T^{2\theta} t^{-3/2-\theta}),$$

$$\sup_{\boldsymbol{\vartheta} \in M_\varepsilon} \left\| \frac{\partial d_{t+1}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\tau}} - \frac{\partial d_t(\boldsymbol{\vartheta})}{\partial \boldsymbol{\tau}} \right\| = O(t^{\max\{\gamma_0 - \delta_0 + \varepsilon - 2, -1-\varsigma\}} \log t + T^{4\theta} t^{-3/2-\theta}),$$

$$\sup_{\boldsymbol{\vartheta} \in M_\varepsilon} |s_{2t+1}(\boldsymbol{\vartheta}) - s_{2t}(\boldsymbol{\vartheta})| = O_p(T^{2\theta} t^{-3/2-\theta}),$$

$$\sup_{\boldsymbol{\vartheta} \in M_\varepsilon} \left\| \frac{\partial s_{2t+1}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\tau}} - \frac{\partial s_{2t}(\boldsymbol{\vartheta})}{\partial \boldsymbol{\tau}} \right\| = O_p(T^{4\theta} t^{-3/2-\theta}).$$

S.3 Proofs of technical lemmas

S.3.1 Proof of Lemma 9

The proof of Lemma 9 is given in Lemma B.4 of Johansen and Nielsen (2010).

S.3.2 Proof of Lemma 10

The proof of the first result is straightforward by approximating the sum by an integral. Next, by the mean value theorem, it is simple to show that

$$\frac{1}{T^d} \sum_{t=1}^{T} \log\left(\frac{t+a}{T}\right) t^{d-1} = \frac{1}{T^d} \sum_{t=1}^{T} \log\left(\frac{t}{T}\right) t^{d-1} + o\left(1\right).$$

Approximating the sum by an integral we find

$$\frac{1}{T} \sum_{t=1}^{T} \log\left(\frac{t}{T}\right) \left(\frac{t}{T}\right)^{d-1} \sim \int_0^1 \log\left(x\right) x^{d-1} dx = B\left(d,1\right)\left(\psi\left(d\right) - \psi\left(d+1\right)\right),$$

see p. 535 of Gradshteyn and Ryzhik (2000), where $B\left(x,y\right) = \Gamma\left(x\right)\Gamma\left(y\right)/\Gamma\left(x+y\right)$ is the Beta function and $\psi(\cdot)$ is the digamma function. Thus, the second result follows by the recurrence formulae for the gamma and digamma functions, see pp. 256 and 258 of Abramowitz and Stegun (1970). Similarly, $T^{-d} \sum_{t=1}^{T} \log^2((t+a)/T) t^{d-1}$ can be approximated by

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^{T} \log^2\left(\frac{t}{T}\right) \left(\frac{t}{T}\right)^{d-1} &\sim \int_0^1 \log^2\left(x\right) x^{d-1} dx \\ &= B\left(d,1\right)\left(\left(\psi\left(d\right) - \psi\left(d+1\right)\right)^2 + \psi'\left(d\right) - \psi'\left(d+1\right)\right), \end{aligned}$$

see p. 538 of Gradshteyn and Ryzhik (2000), where $\psi'(\cdot)$ is the trigamma function. Then the third result follows by the recurrence formulae for the gamma, digamma, and trigamma functions, see pp. 256, 258, and 260 of Abramowitz and Stegun (1970).

S.3.3 Proof of Lemma 11

The proof of Lemma 11 is given in Lemma B.3 of Johansen and Nielsen (2010) and Lemma A.5 of Johansen and Nielsen (2012).

S.3.4 Proof of Lemma 12

The proof of Lemma 12 is almost identical to that of Lemma 1 of Hualde and Robinson (2011) and is therefore omitted.

S.3.5 Proof of Lemma 13

First we show (123). By (67),

$$\sum_{t=1}^{T} c_{t-1}^{2}(d,\boldsymbol{\varphi}) \geq \phi^{2}(1;\boldsymbol{\varphi}) \sum_{t=1}^{T} \pi_{t-1}^{2}(d) - 2\phi(1;\boldsymbol{\varphi}) \sum_{t=1}^{T} \pi_{t-1}^{2}(d) \sum_{k=t}^{\infty} \phi_{k}(\boldsymbol{\varphi})$$
$$- 2\phi(1;\boldsymbol{\varphi}) \sum_{t=1}^{T} \pi_{t-1}(d) \sum_{k=0}^{t-2} \pi_{k+1}(d-1) \sum_{l=0}^{k} \phi_{t-1-l}(\boldsymbol{\varphi})$$
$$+ 2 \sum_{t=1}^{T} \pi_{t-1}(d) \sum_{j=t}^{\infty} \phi_{j}(\boldsymbol{\varphi}) \sum_{k=0}^{t-2} \pi_{k+1}(d-1) \sum_{l=0}^{k} \phi_{t-1-l}(\boldsymbol{\varphi}). \tag{S.79}$$

Noting (7), the fourth term on the right-hand side of (S.79) is of smaller order than the third term. Then the proof of (123) follows on showing

$$\sup_{d\geq 1/2+\eta,\boldsymbol{\varphi}\in\Psi} \frac{1}{T^{2d-1}} \left| \sum_{t=1}^{T} \pi_{t-1}^{2}(d) \sum_{k=t}^{\infty} \phi_{k}(\boldsymbol{\varphi}) \right| = o(1), \tag{S.80}$$

$$\sup_{d\geq 1/2+\eta,\boldsymbol{\varphi}\in\Psi} \frac{1}{T^{2d-1}} \left| \sum_{t=1}^{T} \pi_{t-1}(d) \sum_{k=0}^{t-2} \pi_{k+1}(d-1) \sum_{l=0}^{k} \phi_{t-1-l}(\boldsymbol{\varphi}) \right| = o(1). \tag{S.81}$$

First, by (7) and Lemma 11, the left-hand side of (S.80) is bounded by

$$K \sup_{d\geq 1/2+\eta} T \sum_{t=1}^{T} \left(\frac{t}{T}\right)^{2d} t^{-2-\varsigma} \leq KT \sum_{t=1}^{T} \left(\frac{t}{T}\right)^{1+2\eta} t^{-2-\varsigma}$$
$$\leq K \frac{1}{T^{2\eta}} \sum_{t=1}^{T} t^{-1+2\eta-\varsigma} = o(1),$$

so (S.80) holds. Similarly, the left-hand side of (S.81) is bounded by

$$K \sup_{d\geq 1/2+\eta} T \sum_{t=1}^{T} \left(\frac{t}{T}\right)^{d} t^{-1} \sum_{k=1}^{t-1} \left(\frac{k}{T}\right)^{d} k^{-2} (t-k)^{-\varsigma}$$
$$\leq K \frac{1}{T^{2\eta}} \sum_{t=1}^{T} t^{-1/2+\eta} \sum_{k=1}^{t-1} k^{-3/2+\eta} (t-k)^{-\varsigma}$$
$$\leq K \frac{1}{T^{2\eta}} \sum_{t=1}^{T} t^{-1/2+\eta-\varsigma} (1+\log t) = o(1),$$

for η small, noting that $\varsigma > 1/2$, where the second inequality is due to Lemma 9, to conclude the proof of (123). Next, in view of (67), the proof of (124) follows by (S.80), (S.81) and

$$\sup_{d\geq 1/2+\eta, \boldsymbol{\varphi}\in\Psi} \frac{1}{T^{2d-1}} \sum_{t=1}^{T} \pi_{t-1}^{2}(d) \left(\sum_{k=t}^{\infty} \phi_k(\boldsymbol{\varphi})\right)^2 = o(1),$$
$$\sup_{d\geq 1/2+\eta, \boldsymbol{\varphi}\in\Psi} \frac{1}{T^{2d-1}} \sum_{t=1}^{T} \left(\sum_{k=0}^{t-2} \pi_{k+1}(d-1) \sum_{l=0}^{k} \phi_{t-1-l}(\boldsymbol{\varphi})\right)^2 = o(1),$$

which follow by straightforward arguments using (107) and Lemma 9. The proof of (125) is immediate because

$$\inf_{d_1\leq d\leq d_2, \boldsymbol{\varphi}\in\Psi} \sum\nolimits_{t=1}^{T} c_{t-1}^{2}(d,\boldsymbol{\varphi}) \geq \inf_{d_1\leq d\leq d_2, \boldsymbol{\varphi}\in\Psi} c_{0}^{2}(d,\boldsymbol{\varphi}) = 1.$$

For the proof of (126), it clearly holds that

$$\inf_{d\geq 1/2-\alpha, \boldsymbol{\varphi}\in\Psi} T^{1-2d} \sum\nolimits_{t=1}^{T} c_{t-1}^{2}(d,\boldsymbol{\varphi}) \geq \inf_{d\geq 1/2-\alpha, \boldsymbol{\varphi}\in\Psi} T^{1-2d} \sum\nolimits_{t=[T^{1/2}]}^{T} c_{t-1}^{2}(d,\boldsymbol{\varphi}). \tag{S.82}$$

Then, as in (67), the right-hand side of (S.82) is bounded from below by

$$\epsilon \inf_{d\geq 1/2-\alpha} \frac{1}{T^{2d-1}} \sum\nolimits_{t=[T^{1/2}]}^{T} \pi_{t-1}^{2}(d) \tag{S.83}$$
$$- \sup_{d\geq 1/2-\alpha, \boldsymbol{\varphi}\in\Psi} \frac{K}{T^{2d-1}} \left| \sum_{t=[T^{1/2}]}^{T} \pi_{t-1}^{2}(d) \sum_{l=t}^{\infty} \phi_l(\boldsymbol{\varphi}) \right| \tag{S.84}$$
$$- \sup_{d\geq 1/2-\alpha, \boldsymbol{\varphi}\in\Psi} \frac{K}{T^{2d-1}} \left| \sum_{t=[T^{1/2}]}^{T} \pi_{t-1}(d) \sum_{k=0}^{t-2} \pi_{k+1}(d-1) \sum_{l=0}^{k} \phi_{t-1-l}(\boldsymbol{\varphi}) \right|. \tag{S.85}$$

First, noting that $\sup_{d\geq 1/2-\alpha} (t/T)^{2d} = (t/T)^{1-2\alpha}$, by (7) and Lemma 11, it can be readily shown that (S.84) is $O\left(T^{\alpha-\varsigma/2}\right) = o(1)$ because $\alpha < (\varsigma - 1/2)/3 < \varsigma/2$. Similarly, by identical arguments, (S.85) is bounded by

$$KT^{2\alpha} \sum_{t=[T^{1/2}]}^{T} t^{-1/2-\alpha} \sum_{k=1}^{t-1} k^{-3/2-\alpha} (t-k)^{-\varsigma} \leq KT^{2\alpha} \log T \sum_{t=[T^{1/2}]}^{T} t^{-1/2-\alpha} t^{-\varsigma}$$
$$\leq KT^{3\alpha/2+1/4-\varsigma/2} \log T$$

by Lemma 9. Then (S.85) is $o(1)$ because $\alpha < (\varsigma - 1/2)/3$. Finally, by (120) in Lemma 11, (S.83) is bounded from below by

$$\epsilon \inf_{d \geq 1/2-\alpha} \frac{1}{T^{2d-1}} \sum_{t=[T^{1/2}]}^{T} t^{2d-2} \geq \epsilon \frac{1}{T} \sum_{t=[T^{1/2}]}^{T} \left(\frac{t}{T}\right)^{2\alpha-1} \geq \epsilon \int_{[T^{1/2}]/T}^{1} x^{2\alpha-1} dx$$
$$= \epsilon \frac{1-([T^{1/2}]/T)^{2\alpha}}{2\alpha} = \frac{\epsilon}{2\alpha} - O\left(T^{-\alpha}\right).$$

In view of (S.82), (S.84), and (S.85), this proves (126).

S.3.6 Proof of Lemma 14

The left-hand side of (127) is bounded by

$$\sup_{d_1 \leq d \leq 1/2-\theta, \boldsymbol{\varphi} \in \Psi} \left| \frac{\partial^m}{\partial d^m} h_{t-1,T}(d, \boldsymbol{\varphi}) \right| + \sup_{1/2-\theta \leq d \leq d_2, \boldsymbol{\varphi} \in \Psi} \left| \frac{\partial^m}{\partial d^m} h_{t-1,T}(d, \boldsymbol{\varphi}) \right|. \tag{S.86}$$

Suppose first that $m = 0$. Using the definition (17) and applying (121) of Lemma 12 and (125) of Lemma 13, the first term of (S.86) is bounded by

$$\frac{\sup_{d_1 \leq d \leq 1/2-\theta, \boldsymbol{\varphi} \in \Psi} |c_{t-1}(d, \boldsymbol{\varphi})|}{\inf_{d_1 \leq d \leq 1/2-\theta, \boldsymbol{\varphi} \in \Psi} \left(\sum_{j=1}^{T} c_{j-1}^2 (d, \boldsymbol{\varphi})\right)^{1/2}} \leq \sup_{d_1 \leq d \leq 1/2-\theta, \boldsymbol{\varphi} \in \Psi} |c_{t-1}(d, \boldsymbol{\varphi})|$$
$$= O(t^{-1/2-\theta}),$$

so the bound in (127) applies to the first term of (S.86) (although it is not tight). Next, the second term of (S.86) is bounded by

$$\frac{\sup_{1/2-\theta \leq d \leq d_2, \boldsymbol{\varphi} \in \Psi} T^{-d} |c_{t-1}(d, \boldsymbol{\varphi})|}{\inf_{1/2-\theta \leq d \leq d_2, \boldsymbol{\varphi} \in \Psi} \left(\sum_{j=1}^{T} T^{-2d} c_{j-1}^2 (d, \boldsymbol{\varphi})\right)^{1/2}}.$$

By (122) of Lemma 12 the numerator is $O(t^{-1}(t/T)^{1/2-\theta})$ and by Lemma 13 the denominator is bounded from below by $\epsilon\theta^{-1/2}T^{-1/2}$. Thus (127) for $m = 0$ follows.

Next, for the derivative we find

$$\frac{\partial h_{t-1,T}(d, \boldsymbol{\varphi})}{\partial d} = \frac{c_{t-1}^{(1)}(d, \boldsymbol{\varphi})}{\left(\sum_{j=1}^{T} c_{j-1}^2(d, \boldsymbol{\varphi})\right)^{1/2}} - \frac{h_{t-1,T}(d, \boldsymbol{\varphi}) \sum_{j=1}^{T} c_{j-1}(d, \boldsymbol{\varphi}) c_{j-1}^{(1)}(d, \boldsymbol{\varphi})}{\sum_{j=1}^{T} c_{j-1}^2(d, \boldsymbol{\varphi})}. \tag{S.87}$$

First we show (127). Proceeding as in the proof for $m = 0$, taking into account the extra log-term arising from (121) in Lemma 12, the first term of (S.87) is $O(t^{-1/2}(T/t)^{\theta} \log T)$, so the bound in (127) applies. Next, using again (121) in Lemma 12 and also (127) for $m = 0$, the second term of (S.87) is $O(t^{-1/2}(T/t)^{\theta} T^{2\theta} \sum_{j=1}^{T} j^{-1-2\theta} \log j)$, so the bound in (127) applies for $m = 1$.

Next we show (128). Clearly

$$\frac{\partial h_{t-1,T}(d,\boldsymbol{\varphi})}{\partial d}=\frac{\partial T^{-d}c_{t-1}(d,\boldsymbol{\varphi})/\partial d}{\left(\sum_{j=1}^{T}T^{-2d}c_{j-1}^{2}(d,\boldsymbol{\varphi})\right)^{1/2}}-\frac{T^{-d}c_{t-1}(d,\boldsymbol{\varphi})\sum_{j=1}^{T}T^{-d}c_{j-1}(d,\boldsymbol{\varphi})\,\partial T^{-d}c_{j-1}(d,\boldsymbol{\varphi})/\partial d}{\left(\sum_{j=1}^{T}T^{-2d}c_{j-1}^{2}(d,\boldsymbol{\varphi})\right)^{3/2}}.$$

First,

$$\sup_{d\geq 1/2+\theta,\boldsymbol{\varphi}\in\Psi}\left|\frac{\partial T^{-d}c_{t-1}(d,\boldsymbol{\varphi})/\partial d}{\left(\sum_{j=1}^{T}T^{-2d}c_{j-1}^{2}(d,\boldsymbol{\varphi})\right)^{1/2}}\right|\leq\frac{\sup_{d\geq 1/2+\theta,\boldsymbol{\varphi}\in\Psi}\left|\partial T^{-d}c_{t-1}(d,\boldsymbol{\varphi})/\partial d\right|}{\left(\inf_{d\geq 1/2+\theta,\boldsymbol{\varphi}\in\Psi}\sum_{j=1}^{T}T^{-2d}c_{j-1}^{2}(d,\boldsymbol{\varphi})\right)^{1/2}}$$

$$=O\left(t^{-1/2}\left(\frac{t}{T}\right)^{\theta}(1+|\log(t/T)|)\right),\qquad \text{(S.88)}$$

by (122) of Lemma 12 and (126) of Lemma 13. Similarly, like in (S.88),

$$\sup_{d\geq 1/2+\theta,\boldsymbol{\varphi}\in\Psi}\left|\frac{T^{-d}c_{t-1}(d,\boldsymbol{\varphi})}{\left(\sum_{j=1}^{T}T^{-2d}c_{j-1}^{2}(d,\boldsymbol{\varphi})\right)^{1/2}}\right|=O\left(t^{-1/2}\left(\frac{t}{T}\right)^{\theta}\right),\qquad \text{(S.89)}$$

so by (S.88) and (S.89) it is straightforward to show that

$$\sup_{d\geq 1/2+\theta,\boldsymbol{\varphi}\in\Psi}\left|\frac{T^{-d}c_{t-1}(d,\boldsymbol{\varphi})\sum_{j=1}^{T}T^{-d}c_{j-1}(d,\boldsymbol{\varphi})\,\partial T^{-d}c_{j-1}(d,\boldsymbol{\varphi})/\partial d}{\left(\sum_{j=1}^{T}T^{-2d}c_{j-1}^{2}(d,\boldsymbol{\varphi})\right)^{3/2}}\right|=O\left(t^{-1/2}\left(\frac{t}{T}\right)^{\theta}\right),$$

to conclude the proof of (128).

The proofs of (129)–(131) are omitted because they follow by identical arguments, noting that

$$h_{t,T}(d,\boldsymbol{\varphi})-h_{t-1,T}(d,\boldsymbol{\varphi})=\frac{c_{t}(d-1,\boldsymbol{\varphi})}{\left(\sum_{j=1}^{T}c_{j-1}^{2}(d,\boldsymbol{\varphi})\right)^{1/2}}.$$

S.3.7 Proof of Lemma 15

First we show (132). Write $\phi(L;\boldsymbol{\varphi})\,u_t(-d)=\sum_{j=0}^{t-1}c_j(d,\boldsymbol{\varphi})\,u_{t-j}$ and apply summation by parts,

$$\sum_{j=0}^{t-1}c_j(d,\boldsymbol{\varphi})\,u_{t-j}=c_{t-1}(d,\boldsymbol{\varphi})\sum_{j=0}^{t-1}u_{t-j}-\sum_{j=0}^{t-2}\left(c_{j+1}(d,\boldsymbol{\varphi})-c_j(d,\boldsymbol{\varphi})\right)\sum_{l=0}^{j}u_{t-l}.\qquad \text{(S.90)}$$

Noting that $c_{j+1}(d,\boldsymbol{\varphi})-c_j(d,\boldsymbol{\varphi})=c_{j+1}(d-1,\boldsymbol{\varphi})$, the right-hand side of (S.90) is bounded by

$$|c_{t-1}(d,\boldsymbol{\varphi})|\left|\sum_{j=0}^{t-1}u_{t-j}\right|+\sum_{j=0}^{t-2}|c_{j+1}(d-1,\boldsymbol{\varphi})|\left|\sum_{l=0}^{j}u_{t-l}\right|.\qquad \text{(S.91)}$$

Under our conditions, $E\left|\sum_{l=1}^{t} u_l\right| = O\left(t^{1/2}\right)$, so, in view of (121) of Lemma 12, the expectation of the left-hand side of (132) is bounded by

$$Kt^{\max\{g-1/2,-1/2-\varsigma\}} + K\sum_{j=1}^{t} j^{\max\{g-3/2,-1/2-\varsigma\}} \leq K(t^{g-1/2} + \log t\mathbb{I}\left(g = 1/2\right) + \mathbb{I}\left(g < 1/2\right))$$

to conclude the proof of (132). The proof of (133) is omitted because it is almost identical to that for (132).

S.3.8 Proof of Lemma 16

By summation by parts as in (S.97), we find

$$\begin{aligned}\left|\sum_{t=1}^{T} \phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right| &\leq \left|h_{T-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right|\left|\phi(L;\boldsymbol{\varphi})u_T\left(\delta-\delta_0-1\right)\right| \\ &+\sum_{t=1}^{T-1}\left|h_{t,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)-h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right|\left|\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0-1\right)\right|.\end{aligned}$$

First, application of (127), (129) of Lemma 14 together with (132), (133) of Lemma 15 implies (134) and (135). Next, (136) and (137) follow from (128), (130) of Lemma 14 and (132), (133) of Lemma 15.

S.3.9 Proof of Lemma 17

Letting $d_t\left(\boldsymbol{\tau},\gamma\right) = d_t\left(\boldsymbol{\vartheta}\right)$, noting (29) and that $d_t\left(\boldsymbol{\tau},\gamma_0\right) = 0$, by the mean value theorem,

$$\left|\sum_{t=1}^{T} d_t\left(\boldsymbol{\vartheta}\right)s_t\left(\boldsymbol{\vartheta}\right)\right| \leq \left|\gamma-\gamma_0\right|\left|\frac{\partial}{\partial\gamma}\sum_{t=1}^{T} d_t\left(\boldsymbol{\tau},\overline{\gamma}\right)\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)\right|,$$

where $\left|\overline{\gamma}-\gamma_0\right| \leq \left|\gamma-\gamma_0\right|$. Then we find the bound

$$\begin{aligned}&\left|\frac{\partial}{\partial\gamma}\sum_{t=1}^{T} d_t\left(\boldsymbol{\tau},\gamma\right)\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)\right| \\ &\leq \left|\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)\frac{\partial h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)}{\partial\gamma}\sum_{j=1}^{T} c_{j-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)h_{j-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right| \\ &\quad+\left|\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\sum_{j=1}^{T} c_{j-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)\frac{\partial h_{j-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)}{\partial\gamma}\right|.\end{aligned}$$

The results (139)–(141) now all follow by direct application of (S.92), (S.93) of Lemma S.1 with $\theta < g - 1/2$ and (134), (135) of Lemma 16. Results (142) and (143) are derived straightforwardly from (S.94) of Lemma S.1 and (136), (137) of Lemma 16.

S.4 Additional technical lemmas

Lemma S.1 *Let θ be an arbitrary number such that $0 < \theta < \varsigma - 1/2$. Then, under Assumptions A1 and A3, for $m = 0, 1$ and uniformly in $\boldsymbol{\vartheta} \in \Xi$,*

$$\sup_{\gamma_0-\delta\leq g}\left|\sum_{t=1}^{T} c_{t-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)\frac{\partial^m}{\partial\gamma^m}h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right| = O(T^{\max\{\theta,g-1/2\}+2\theta m}), \tag{S.92}$$

$$\sup_{\gamma_0-\delta\geq g}\frac{1}{T^{\gamma_0-\delta}}\left|\sum_{t=1}^{T} c_{t-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)\frac{\partial^m}{\partial\gamma^m}h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right| = O(T^{\max\{\theta,g-1/2\}-g+2\theta m}), \tag{S.93}$$

and for $g > 1/2$ and $\theta < g - 1/2$, uniformly in $\boldsymbol{\vartheta} \in \Xi$,

$$\sup_{\gamma_0-\delta\geq g,\gamma-\delta\geq 1/2+\theta}\frac{1}{T^{\gamma_0-\delta}}\left|\sum_{t=1}^{T} c_{t-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)\frac{\partial}{\partial\gamma}h_{t-1,T}\left(\gamma-\delta,\boldsymbol{\varphi}\right)\right| = O(T^{-1/2}). \tag{S.94}$$

Proof. The results follow by direct application of (127), (128) of Lemma 14 and (121), (122) of Lemma 12. ■

Lemma S.2 *Under Assumptions A1–A3, uniformly in $\boldsymbol{\vartheta} \in \Xi$,*

$$\begin{aligned}&\sup_{\delta_0-\delta\leq g_1,\gamma_0-\delta\leq g_2}\frac{1}{T}\left|\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)c_{t-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)\right| \\ &= O_p(T^{\max\{g_2-1,-\varsigma\}+g_1-1/2}+T^{-1}\log^2 T+T^{\max\{g_2-1,-1-\varsigma\}-1}(\log T)\mathbb{I}\left(g_1\leq -1/2\right)),\end{aligned} \tag{S.95}$$

$$\begin{aligned}&\sup_{\delta_0-\delta\geq g_1,\gamma_0-\delta\geq g_2}\frac{1}{T^{\gamma_0+\delta_0-2\delta}}\left|\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)c_{t-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)\right| \\ &= O_p(T^{\max\{g_2-1,-\varsigma\}-g_2+1/2}+T^{-g_1-g_2}\log^2 T+T^{\max\{g_2-1,-\varsigma\}-g_1-g_2}(\log T)\mathbb{I}\left(g_1\leq -1/2\right)).\end{aligned} \tag{S.96}$$

Proof. By summation by parts and (65) we find

$$\begin{aligned}\left|\sum_{t=1}^{T}\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0\right)c_{t-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)\right| &\leq \left|c_{T-1}\left(\gamma_0-\delta,\boldsymbol{\varphi}\right)\right|\left|\phi(L;\boldsymbol{\varphi})u_T\left(\delta-\delta_0-1\right)\right| \\ &+\left|\sum_{t=1}^{T-1}c_t\left(\gamma_0-\delta-1,\boldsymbol{\varphi}\right)\phi(L;\boldsymbol{\varphi})u_t\left(\delta-\delta_0-1\right)\right|.\end{aligned} \tag{S.97}$$

The result (S.95) then follows by application of (132) of Lemma 15 and (121) of Lemma 12, while the result (S.96) follows by application of (133) of Lemma 15 and (122) of Lemma 12. ■

References

1. Abramowitz, M. and Stegun, I.A. (1970). *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, D.C.

2. Gradshteyn, I.S. and Ryzhik, I.M. (2000). *Tables of Integrals, Series and Products*, Academic Press, New York.
3. Hosoya, Y. (2005). Fractional invariance principle. *Journal of Time Series Analysis* **26**, 463–486.
4. Hualde, J. and Nielsen, M.Ø. (2017). Truncated sum of squares estimation of fractional time series models with deterministic trends. Working paper.
5. Hualde, J. and Robinson, P.M. (2011). Gaussian pseudo-maximum likelihood estimation of fractional time series models. *Annals of Statistics* **39**, 3152–3181.
6. Johansen, S. and Nielsen, M.Ø. (2010). Likelihood inference for a nonstationary fractional autoregressive model. *Journal of Econometrics* **158**, 51–66.
7. Johansen, S. and Nielsen, M.Ø. (2012). Likelihood inference for a fractionally cointegrated vector autoregressive model. *Econometrica* **80**, 2667–2732.
8. Robinson, P.M. and Hualde, J. (2003). Cointegration in fractional systems with unknown integration orders. *Econometrica* **71**, 1727–1766.
9. Robinson, P.M. and Iacone, F. (2005). Cointegration in fractional systems with deterministic trends. *Journal of Econometrics* **129**, 263–298.