Hualde, Javier; Nielsen, Morten Ørregaard

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

Queen's Economics Department Working Paper, No. 1376

Provided in Cooperation with:
Queen's University, Department of Economics (QED)

Suggested Citation: Hualde, Javier; Nielsen, Morten Ørregaard (2017): Truncated sum of squares estimation of fractional time series models with deterministic trends, Queen's Economics Department Working Paper, No. 1376, Queen's University, Department of Economics, Kingston (Ontario)

This Version is available at:
http://hdl.handle.net/10419/149102

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Truncated sum of squares estimation of fractional time series models with deterministic trends

Javier Hualde
Universidad Publica de Navarra

Morten Ørregaard Nielsen
Queen’s University and CREATES

Department of Economics
Queen’s University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

1-2017
TRUNCATED SUM OF SQUARES ESTIMATION OF FRACTIONAL TIME SERIES MODELS WITH DETERMINISTIC TRENDS

BY JAVIER HUALDE†

Universidad Pública de Navarra
E-mail: javier.hualde@unavarra.es

AND

BY MORTEN ØRREGAARD NIELSEN

Queen’s University and CREATES
E-mail: mon@econ.queensu.ca

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 non-invertibility. 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.

∗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.

†Corresponding author.

MSC 2010 subject classifications: Primary 62M10; secondary 62F12

Keywords and phrases: Asymptotic normality, consistency, deterministic trend, fractional process, generalized polynomial trend, noninvertibility, nonstationarity, truncated sum of squares estimation
1. Introduction. Empirical evidence suggests that observed time series in most fields of research are formed by both stochastic and deterministic components. Traditionally, the most common method of 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” $\delta_0$ equal to zero, and unit root nonstationary process, where $\delta_0 = 1$. A zero-mean FARIMA($p_1, \delta_0, p_2$) process $z_t$ is given by

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

(1)

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

(2)

where $L$ is the lag operator and $\Delta^\zeta_+$ is given by $\Delta^\zeta_+ x_t = \Delta^\zeta x_t I (t \geq 1) = \sum_{i=0}^{t-1} \pi_i (-\zeta) x_{t-i}$ with

$$\pi_i (v) = \frac{\Gamma (v + i)}{\Gamma (v) \Gamma (1 + i)} \frac{v(v + 1) \ldots (v + i - 1)}{i!}, \ i \geq 0,$$

(3)

and $\pi_i (v) = 0, i < 0$, denoting the coefficients in the usual binomial expansion of $(1 - z)^{-v}$ and $I(\cdot)$ 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 $\varepsilon_t$ 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

$$u_t = \omega(L; \varphi_0)\varepsilon_t, \ t = 0, \pm 1, \pm 2, \ldots,$$

(4)

where $\varphi_0$ is an unknown $p \times 1$ vector and $\omega(s; \varphi) = \sum_{j=0}^\infty \omega_j(\varphi)s^j$ with, for all $\varphi$, $\omega_0(\varphi) = 1$ and $|\omega(s; \varphi)| \neq 0, |s| \leq 1$. Like $\alpha$ and $\beta$ in (2), $\omega$ in (4) characterizes parametric short memory autocorrelation.

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) has been usually complemented with deterministic components, although many theoretical developments have exclusively assumed a purely random process. For a review and discussion of these, see [10] or [13].
A simple extension of the fractional model is to allow a non-zero mean as in, for example,
\begin{equation}
\Delta_+^{\delta_0} x_t = \mu_0 + u_t \text{ or } x_t = \Delta_+^{\delta_0} \mu_0 + \Delta_+^{\delta_0} u_t,
\end{equation}
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
\begin{equation}
x_t = \mu_0 \pi_{t-1} (\gamma_0) + z_t,
\end{equation}
where \( \mu_0 \) and \( \gamma_0 \) 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 \( \pi_{t-1}(1) = I(t \geq 1) \) and \( \pi_{t-1}(2) = tI(t \geq 1) \), see (3)), but, given that \( \gamma_0 \) 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, therefore mimicking the fractional behavior of the stochastic component. This can be made more precise using the terminology of [22], where different definitions of trends appear. Because \( \text{Var} (x_t) \) 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 \( t^{\gamma_0 - 1} \), according to [22], 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 \( \delta_0 \) and \( \gamma_0 \), respectively, and hence letting \( \gamma_0 \) be real-valued appears as natural as letting \( \delta_0 \) be real-valued, it just affects another aspect of the distribution of \( x_t \).

We note that if \( \gamma_0 \) 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. [10]. Having thus estimated \( \delta_0 \), and denoting this estimator \( \hat{\delta} \), one could estimate \( \mu_0 \) by regression, i.e. by \( \hat{\mu}(\hat{\delta}) \), where \( \hat{\mu}(\delta) = \)}
\[ \sum_{t=1}^{T} \pi_{t-1} (\gamma_0 - \delta) \Delta_t x_t / \sum_{t=1}^{T} \pi_{t-1}^2 (\gamma_0 - \delta), \] and the problem reduces to a question of whether \( \hat{\mu}(\hat{\delta}) \) has identical limiting distribution to \( \hat{\mu}(\delta_0) \). This question will be covered by our more general theory, which assumes \( \gamma_0 \) unknown. The inclusion of deterministic trends of known order and the idea of differencing-and-adding-back is also considered by [3, 20, 21] in combination with tapering of the periodogram in frequency domain methods. Similarly, [16] considers \( M \)-estimation of a model like (6) (although involving more deterministic terms) with known \( \gamma_0 \) 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 \( \gamma_0 \) 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, [23] 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 \( \mu \) and \( \gamma \) have different normalizations. The analysis of [23] was generalized by [14] 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, [17] considers (6), 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 [11], 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, [18, 19] 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. Importantly, we let the memory (\( \delta_0 \)) and exponent (\( \gamma_0 \)) parameters lie in arbitrarily large, but finite, intervals, so \( x_t \) can display many different behaviours. As in related work, e.g. [10, 13], 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, $\mu_0$ and $\gamma_0$, cannot be consistently estimated, but, remarkably, those related to the stochastic part of the model, i.e. $\delta_0$ and $\varphi_0$, are still consistently estimated and their asymptotic normal distribution is unaffected by the presence of the remaining unestimable parameters. 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. Some concluding remarks are presented in Section 4. The proofs of the main theorems are given in Section 5, which applies some auxiliary lemmas provided in Section 6. To conserve space, the proofs of these lemmas, along with additional proof details for the main theorems, some additional technical lemmas as well as a Monte Carlo simulation study, are included in the supplementary material [9].

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 $\mu_0, \gamma_0, \delta_0$, and $\varphi_0$ are unknown parameters to be estimated.

We first impose an assumption on the short memory component, $\omega$, where $\varphi_0$ is assumed to lie in $\Psi$, which is a compact and convex subset of $\mathbb{R}^p$.

A1. (i) for all $\varphi \in \Psi \backslash \{\varphi_0\}, \left| \omega(s; \varphi) \right| \neq \left| \omega(s; \varphi_0) \right|$ on a set $S \subset \{s : |s| = 1\}$ of positive Lebesgue measure;

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

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

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

Assumption A1 is identical to A1 in [10]. 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 $j$, $\sup_{\varphi \in \Psi} |\omega_j(\varphi)| = O \left( j^{-1-\varsigma} \right)$ as $j \to \infty$. Also, writing $\omega^{-1}(s; \varphi) = \phi(s; \varphi) = \sum_{j=0}^{\infty} \phi_j(\varphi)s^j$, it holds that $\phi_0(\varphi) = 1$ for all
\( \varphi \), and (ii), (iii), and (iv) imply that
\[
(7) \quad \sup_{\varphi \in \Psi} |\phi_j(\varphi)| = O\left(j^{-1-\varsigma}\right) \text{ as } j \to \infty,
\]
whereas (ii) also implies that
\[
(8) \quad \inf_{|s|=1, \varphi \in \Psi} |\phi(s; \varphi)| > 0.
\]

A1 is easily satisfied in the stationary and invertible ARMA case. Another model covered by A1 is the exponential spectrum model of [1], which leads to a relatively simple covariance matrix formula in the context of fractional time series models, see [15]. More generally, A1 is also similar to other conditions employed in asymptotic theory for the estimate \( \hat{\tau} = (\hat{\delta}, \hat{\varphi}') \) below, see [10, 13], as well as Whittle estimators that restrict to stationarity, e.g. [4, 5, 6]. Assumption A1 can be readily verified because \( \omega \) is a known parametric function. In fact \( \omega \) satisfying A1 are invariably employed by practitioners.

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

Assumption A2 is identical to A2 in [10]. It does not impose independence or identity of distribution of \( \varepsilon_t \), but rules out conditional heteroskedasticity. It is standard in the time series asymptotics literature since [8].

A3. The parameter space for \( \vartheta = (\delta, \varphi', \gamma)' \) is given by \( \Xi = [\nabla_1, \nabla_2] \times \Psi \times [\square_1, \square_2] \) with \( \nabla_1 < \nabla_2 \) and \( \square_1 < \square_2 \), where \( \Psi \) is compact and convex and \( \vartheta_0 = (\delta_0, \varphi_0', \gamma_0)' \in \Xi \). For \( \mu \) the parameter space is \( \mathbb{R} \), and we assume that \( \mu_0 \neq 0 \).

We assume that \( \mu_0 \neq 0 \) since otherwise \( \gamma_0 \) is not identified. Note that the model where \( \gamma_0 \) is known, e.g., the model with a constant term or a linear trend, is a special case of our model with unknown \( \gamma_0 \). Hence, the asymptotic results below can easily be specialized to this situation. In general, though, \([\nabla_1, \nabla_2]\) 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 \( \tau = (\delta, \varphi')' \) with true value \( \tau_0 = (\delta_0, \varphi_0')' \), and denote the estimator (to be defined below) by \( \hat{\tau} = (\hat{\delta}, \hat{\varphi}')' \). We also use the notation \( \vartheta = (\tau', \gamma)' \), \( \vartheta_0 = (\tau_0', \gamma_0)' \), and \( \hat{\vartheta} = (\hat{\tau}', \hat{\gamma})' \). The Gaussian log-likelihood, conditional on \( x_t = 0 \) for \( t \leq 0 \), is, apart from a
constant, given by

\[ L_T(\vartheta, \mu, \sigma^2) = -\frac{T}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^{T} \left( \phi(L; \varphi) \Delta^\delta_t (x_t - \mu \tau_{t-1} (\gamma)) \right)^2 \]

(9)

\[ = -\frac{T}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^{T} (\phi(L; \varphi) x_t(\delta) - \mu c_{t-1}(\gamma - \delta, \varphi))^2, \]

where \( x_t(\delta) = \Delta^\delta_t x_t \) and we have defined

(10) \[ c_{t-1}(d, \varphi) = \phi(L; \varphi) \tau_{t-1}(d) = \sum_{j=0}^{t-1} \phi_j(\varphi) \tau_{t-j-1}(d). \]

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

(11) \[ \hat{\mu}(\vartheta) = \frac{\sum_{t=1}^{T} \phi(L; \varphi) x_t(\delta) c_{t-1}(\gamma - \delta, \varphi)}{\sum_{t=1}^{T} c_{t-1}^2(\gamma - \delta, \varphi)}, \]

and we then propose the estimator

\[ \hat{\vartheta} = \arg \min_{\vartheta \in \Xi} R_T(\vartheta), \quad R_T(\vartheta) = \frac{1}{T} \sum_{t=1}^{T} (\phi(L; \varphi) x_t(\delta) - \hat{\mu}(\vartheta) c_{t-1}(\gamma - \delta, \varphi))^2, \]

along with \( \hat{\mu} = \hat{\mu}(\hat{\vartheta}) \). This estimator, often termed nonlinear least squares or conditional sum-of-squares, although we prefer the term truncated sum-of-squares as suggested by [10], goes back to, at least, [2] for estimation of nonfractional ARMA models (where \( \delta_0 \) is a known integer). In the context of fractional time series, the estimator was first analyzed by [12] in stationary FARIMA models with \( 0 < \delta_0 < 1/2 \) and later by [10, 13] for \( \delta_0 \) 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 \( \hat{\vartheta} \to_p \vartheta_0 \) as \( T \to \infty \).

(ii) If \( \gamma_0 - 1/2 < \delta_0 \) then \( \hat{\tau} \to_p \tau_0 \) as \( T \to \infty \).

We note that the result in part (i) of Theorem 1 includes consistency of the estimator of the parameter vector \( \vartheta_0 \). Under (i) \( \mu_0 \) can be also consistently estimated, but we do not report this result here: \( \hat{\mu} \) 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 $\tau_0$. In fact, $\gamma_0$ and $\mu_0$ 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 one cannot consistently estimate trends of order $\gamma_0 - 1 < 1/2$ or the example $\delta_0 = 0$ (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 $\hat{\gamma}$ and $\hat{\mu}$ are consistent as long as $\delta_0 < 1/2$, i.e. $z_t$ is (asymptotically) stationary. Thus, the remarkable feature about part (ii) of Theorem 1 is that, even though $\gamma_0$ and $\mu_0$ cannot be consistently estimated, the remaining parameters $\tau_0$ can still be consistently estimated.

As in related work, e.g. [10, 13], the proof of Theorem 1 is challenging due to non-uniform convergence of the objective function over a large admissible parameter space for $\delta$. 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 5.1.1.

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

**A4.** (i) $\vartheta_0 \in \text{int}(\Xi)$;

(ii) for all $\lambda$, $\omega(e^{i\lambda}; \varphi)$ is thrice continuously differentiable in $\varphi$ on a closed neighbourhood $\mathcal{N}_\epsilon(\varphi_0)$ of radius $\epsilon \in (0, 1/2)$ about $\varphi_0$;

(iii) the matrix

$$A = \begin{pmatrix} \pi^2/6 & -\sum_{j=1}^{\infty} b_j (\varphi_0) / j \\ -\sum_{j=1}^{\infty} b_j (\varphi_0) / j & \sum_{j=1}^{\infty} b_j (\varphi_0) b'_j (\varphi_0) \end{pmatrix}$$

is nonsingular, where $b_j (\varphi_0) = \sum_{k=0}^{j-1} \omega_k (\varphi_0) \partial \phi_{j-k} / \partial \varphi$.}

This assumption is almost identical to A3 in [10], with the only difference that our A4(ii) is slightly stronger than their A3(ii) in imposing thrice instead of twice continuously differentiable $\omega(e^{i\lambda}; \varphi)$. The reason for strengthening the assumption in [10] is that now third derivatives of $\phi_j(\varphi)$ are involved in the proof of convergence of the Hessian matrix below. As in [10], by compactness of $\mathcal{N}_\epsilon(\varphi_0)$ and continuity of $\partial \phi_j(\varphi) / \partial \varphi_i$, $\partial^2 \phi_j(\varphi) / \partial \varphi_i \partial \varphi_l$, and $\partial^3 \phi_j(\varphi) / \partial \varphi_i \partial \varphi_l \partial \varphi_k$ for all $j$, with $i, l, k = 1, \ldots, p$, where $\varphi_i$ denotes
the $i$-th element of $\varphi$, A1(ii), A1(iv) and A4(ii) imply that, as $j \to \infty$, 
\[
\sup_{\varphi \in \mathcal{N}_i(\varphi_0)} \left| \frac{\partial \hat{\phi}_j(\varphi)}{\partial \varphi_i} \right| = O(j^{-1-\epsilon}), \quad \sup_{\varphi \in \mathcal{N}_i(\varphi_0)} \left| \frac{\partial^2 \hat{\phi}_j(\varphi)}{\partial \varphi_i \partial \varphi_l} \right| = O(j^{-1-\epsilon}), 
\]

\[
(12) \quad \sup_{\varphi \in \mathcal{N}_i(\varphi_0)} \left| \frac{\partial^3 \hat{\phi}_j(\varphi)}{\partial \varphi_i \partial \varphi_l \partial \varphi_k} \right| = O(j^{-1-\epsilon}). 
\]

Again A4 is satisfied in the ARMA case.

Define the scaling matrices 
\[
P_T = \begin{pmatrix} M_T & 0 \\ 0 & T^{1-(\gamma_0-\delta_0) \log T} \end{pmatrix}, \quad M_T = \begin{pmatrix} I_{p+1} & 0 \\ 0 & T^{1-(\gamma_0-\delta_0)} \end{pmatrix}, 
\]

and also 
\[
V = \begin{pmatrix} \sigma_0^2 A & 0 \\ 0 & \frac{\mu_0^2 d^2(1; \varphi_0)}{\Gamma^2(\gamma_0-\delta_0)(2(\gamma_0-\delta_0)-1)^3} \end{pmatrix}. 
\]

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

(i) If $\gamma_0 - 1/2 > \delta_0$, then, as $T \to \infty$, 
\[T^{1/2}P_T^{-1} \left( \hat{\vartheta} - \vartheta_0 \right) \to_d \begin{pmatrix} I_{p+2} & 0 \\ 0 & -\mu_0 \end{pmatrix} N, \]

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

(ii) If $\gamma_0 - 1/2 < \delta_0$, then, as $T \to \infty$, 
\[T^{1/2} (\hat{\tau} - \tau_0) \to_d N \left( 0, A^{-1} \right). \]

A remarkable feature of the results in Theorem 2 is that the asymptotic distribution of $\hat{\tau}$ is unaffected by the presence of the deterministic component in (6), and $\hat{\tau}$ has the same asymptotic distribution as in, e.g., Theorem 2.2 of [10]. Moreover, as with the consistency result in Theorem 1, the asymptotic distribution result for $\hat{\tau}$ 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 $\gamma_0$ and $\mu_0$ cannot be consistently estimated, the asymptotic distribution of $\hat{\vartheta}$ is unaffected.

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

The joint asymptotic distribution of $\hat{\vartheta}$ and $\hat{\mu}$ given in (13) is singular, which makes testing of joint hypotheses on $\vartheta_0$ and $\mu_0$ impossible. However,
separate inference can be conducted on $\vartheta_0$ and $\mu_0$. For example, it is straightforward given (13) to construct confidence intervals and test hypotheses, e.g. that $\gamma_0 = 1$ (constant).

4. 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, $\gamma_0$ and $\delta_0$, 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 (antipersistence, weak dependence, long memory). Our results depend crucially on whether the deterministic 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 ($\hat{\tau}$) are identical to those achieved in the simpler, purely stochastic, setting of [10]. When the deterministic signal is weak, i.e., $\gamma_0 - \delta_0 < 1/2$, $\gamma_0$ and $\mu_0$ cannot be consistently estimated, but, nicely, $\hat{\tau}$ 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 various deterministic terms characterized by different power law parameters. Considering this further generality 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, so the proofs present in a clear way the essence of the problem of the competition between deterministic and stochastic terms. Secondly, a semiparametric approach which focuses on estimating $\gamma_0$ and $\delta_0$ without making parametric assumptions about the structure of $z_t$ seems possible. Thirdly, 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.

5. Proofs of theorems.

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

5.1.1. Overall design of the proof. Throughout, $\epsilon$ will denote a generic arbitrarily small positive constant, and $K$ a generic arbitrarily large positive.
constant. Fix $\varepsilon > 0$ and let $M_\varepsilon = \{ \vartheta \in \Xi : \| \tau - \tau_0 \| < \varepsilon \}$, $\overline{M}_\varepsilon = \{ \vartheta \in \Xi : \| \tau - \tau_0 \| \geq \varepsilon \}$. Then $\text{Pr}(\| \hat{\vartheta} - \vartheta_0 \| \geq \varepsilon) \to 0$ as $T \to \infty$, is implied by
\begin{align}
\text{Pr}(\hat{\vartheta} \in \overline{M}_\varepsilon) \to 0 & \quad \text{as } T \to \infty, \\
\text{Pr}(\hat{\vartheta} \in \overline{N}_\varepsilon \cap M_\varepsilon) \to 0 & \quad \text{as } T \to \infty.
\end{align}
Strictly, $\varepsilon$ should be $\varepsilon/\sqrt{2}$ in (15) and (16), but since $\varepsilon$ is arbitrary this is irrelevant and we continue without the $\sqrt{2}$ factor.

We decompose the objective function as $R_T(\vartheta) = \frac{1}{T} \sum_{t=1}^{T} (d_t(\vartheta) + s_t(\vartheta))^2$ with
\begin{align}
d_t(\vartheta) &= \mu_0 \left( c_{t-1}(\gamma_0 - \delta, \varphi) - h_{t-1,T}(\gamma - \delta, \varphi) \sum_{j=1}^{T} c_{j-1}(\gamma_0 - \delta, \varphi) h_{j-1,T}(\gamma - \delta, \varphi) \right), \\
s_t(\vartheta) &= \phi(L; \varphi) u_t(\delta - \delta_0) - h_{t-1,T}(\gamma - \delta, \varphi) \sum_{j=1}^{T} \phi(L; \varphi) u_j(\delta - \delta_0) h_{j-1,T}(\gamma - \delta, \varphi),
\end{align}
defining also the coefficient
\begin{equation}
h_{t,T}(d, \varphi) = \frac{c_t(d, \varphi)}{\left( \sum_{j=1}^{T} c_{j-1}^2(d, \varphi) \right)^{1/2}},
\end{equation}
which clearly satisfies $\sum_{t=1}^{T} h_{t-1,T}(d, \varphi) = 1$.

The strategy of proof relies on recognizing the competition between the stochastic term $s_t(\vartheta)$ and deterministic term $d_t(\vartheta)$ in $R_T(\vartheta)$, taking into account that when considering (15), just $\tau$ is for sure “far” from $\tau_0$, whereas when dealing with (16), just $\gamma$ is “far” from $\gamma_0$. As will be seen, an important feature of the problem is that when $\gamma = \gamma_0$ we have $d_t(\vartheta) = 0$, which complicates the treatment of (15). In any case, as in [10], we need to carefully consider the cases where $R_T(\vartheta)$ shows distinct behaviours, noting that either the deterministic or the stochastic term might dominate, and below we partition the parameter space accordingly.

5.1.2. Proof of (15). To prove (15) we use
\begin{equation}
\text{Pr}(\hat{\vartheta} \in \overline{M}_\varepsilon) = \text{Pr} \left( \inf_{\vartheta \in \overline{M}_\varepsilon} R_T(\vartheta) \leq \inf_{\vartheta \in M_\varepsilon} R_T(\vartheta) \right) \leq \text{Pr} \left( \inf_{\vartheta \in \overline{M}_\varepsilon} S_T(\vartheta) \leq 0 \right),
\end{equation}
where $S_T(\vartheta) = R_T(\vartheta) - R_T(\vartheta_0)$. Fix an arbitrarily small $\eta > 0$ such that $\eta < (\gamma_0 - \delta_0 - 1/2)/2$ and suppose that $\nabla_1 < \delta_0 - 1/2 - \eta$ and $\nabla_2 > \gamma_0 - 1 -$
\( \eta \). 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 \( I_1 = \{ \delta : \gamma_1 \leq \delta \leq \delta_0 - 1/2 - \eta \} \), \( I_2 = \{ \delta : \delta_0 - 1/2 - \eta \leq \delta \leq \delta_0 - 1/2 \} \), \( I_3 = \{ \delta : \delta_0 - 1/2 \leq \delta \leq \delta_0 - 1/2 + \eta \} \), \( I_4 = \{ \delta : \delta_0 - 1/2 + \eta \leq \delta \leq \gamma_0 - 1 - \eta \} \), and \( I_5 = \{ \delta : \gamma_0 - 1 - \eta \leq \delta \leq \gamma_2 \} \), noting that the upper bound for \( \eta \) guarantees that \( I_4 \) is non-empty. Correspondingly define \( T_i = I_i \times \Psi \) and, fixing \( \xi > 0 \) and \( \varrho > 0 \), such that \( \varrho < \eta/2 \), also define \( \mathcal{H}_i = \{ \vartheta \in M_\varepsilon : \tau \in T_i, |\gamma - \gamma_0| < \xi T^{-\kappa_i} \} \), \( \mathcal{H}_i = \{ \vartheta \in M_\varepsilon : \tau \in T_i, \xi T^{-\kappa_i} \leq |\gamma - \gamma_0| \leq \varrho \} \) and \( \overline{\mathcal{H}}_i = \{ \vartheta \in M_\varepsilon : \tau \in T_i, |\gamma - \gamma_0| \geq \varrho \} \), \( i = 1, \ldots, 5 \), where \( \kappa_i > 0 \) will be defined subsequently, noting that \( \overline{\mathcal{H}}_i \) is non-empty for any \( \xi, \varrho \), for \( T \) large enough. Then, by \( (18) \), \( (15) \) is justified by showing

\[
\begin{align*}
(19) & \quad \text{Pr} \left( \inf_{\mathcal{H}_i} S_T(\vartheta) \leq 0 \right) \to 0 \text{ as } T \to \infty \text{ for } i = 1, \ldots, 5, \\
(20) & \quad \text{Pr} \left( \inf_{\mathcal{H}_i} S_T(\vartheta) \leq 0 \right) \to 0 \text{ as } T \to \infty \text{ for } i = 1, \ldots, 5, \\
(21) & \quad \text{Pr} \left( \inf_{\mathcal{H}_i} S_T(\vartheta) \leq 0 \right) \to 0 \text{ as } T \to \infty \text{ for } i = 1, \ldots, 5.
\end{align*}
\]

We note that \( \mathcal{H}_i \), \( \mathcal{H}_i \), and \( \overline{\mathcal{H}}_i \) are designed exactly such that in \( \mathcal{H}_i \) the stochastic term dominates \( S_T(\vartheta) \) while in \( \mathcal{H}_i \cup \overline{\mathcal{H}}_i \) it is the deterministic term that dominates. As will be seen, the analysis on \( \overline{\mathcal{H}}_i \) is much simpler because \( \gamma \) is “far” from \( \gamma_0 \), whereas a much more delicate treatment in necessary for \( \mathcal{H}_i \). 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 \( \mathcal{H}_5 \cup \mathcal{H}_5 \cup \overline{\mathcal{H}}_5 \), where \( \delta_0 - \delta \leq 1 + \delta_0 - \gamma_0 + \eta < 1/2 \), so \( u_t(\delta - \delta_0) \) is asymptotically stationary. Let

\[
(22) S_T(\vartheta) = U(\tau) - r_T(\vartheta),
\]

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

\[
\begin{align*}
r_T(\vartheta) &= \frac{1}{T} \sum_{t=1}^{T} \left( (\phi(L; \varphi_0) \{ u_t \mathbb{I}(t > 0) \})^2 - \sigma_0^2 \right) \\
&\quad - \frac{1}{T} \sum_{t=1}^{T} \left( (\phi(L; \varphi) u_t(\delta - \delta_0))^2 - E \left( (\phi(L; \varphi) \Delta^{\delta - \delta_0} u_t)^2 \right) \right) \\
&\quad - \frac{1}{T} \left( \sum_{t=1}^{T} \phi(L; \varphi_0) (u_t \mathbb{I}(t > 0)) h_{t-1,T}(\gamma_0 - \delta_0, \varphi_0) \right)^2.
\end{align*}
\]
\[
+ \frac{1}{T} \left( \sum_{t=1}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) h_{t-1,T} (\gamma - \delta, \varphi) \right)^2 \\
- \frac{2}{T} \sum_{t=1}^{T} d_t (\vartheta) s_t (\vartheta) - \frac{1}{T} \sum_{t=1}^{T} d_t^2 (\vartheta).
\]

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

\[
(23) \quad \inf_{\|\tau - \tau_0\| \geq \varepsilon, \tau \in T_5} U (\tau) > \varepsilon,
\]

\[
(24) \quad \frac{1}{T} \sum_{t=1}^{T} \left( (\phi(L; \varphi_0) \{u_t \mathbb{I} (t > 0)\})^2 - \sigma_0^2 \right) = o_p (1),
\]

\[
(25) \quad \sup_{\|\tau - \tau_0\| \geq \varepsilon, \tau \in T_5} \frac{1}{T} \sum_{t=1}^{T} \left( (\phi(L; \varphi) u_t (\delta - \delta_0))^2 - E \left( (\phi(L; \varphi) \Delta^{\delta - \delta_0} u_t)^2 \right) \right) = o_p (1),
\]

\[
(26) \quad \sup_{\mathcal{H}_3 \cup \mathcal{H}_5 \cup \mathcal{H}_5} \frac{1}{T} \sum_{t=1}^{T} d_t (\vartheta) s_t (\vartheta) = o_p (1),
\]

\[
(27) \quad \sup_{\mathcal{H}_3 \cup \mathcal{H}_5 \cup \mathcal{H}_5} \frac{1}{T} \left( \sum_{t=1}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) h_{t-1,T} (\gamma - \delta, \varphi) \right)^2 = o_p (1).
\]

First, (23), (24), and (25) follow by identical arguments to those in the proofs of (2.8) and (2.9) in [10]. Next, by (111) of Lemma 6 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 \( O_p (T^\max\{\vartheta, 1 + \delta_0 - \gamma_0 + \eta\} + 2\vartheta) \), and by (106) of Lemma 5, the left-hand side of (27) is \( O_p (T^2 \max\{\vartheta, 1 + \delta_0 - \gamma_0 + \eta\} - 1) \). Both are \( o_p (1) \) for \( \vartheta \) and \( \eta \) 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 7 and that \( d_t (\vartheta_0) = 0 \), holds if, for \( i = 1, \ldots, 4 \),

\[
\Pr \left( \inf_{\mathcal{P}_i} \frac{1}{T} \sum_{t=1}^{T} (d_t (\vartheta) + s_t (\vartheta))^2 \leq \sigma_0^2 + \varepsilon \right) \to 0 \text{ as } T \to \infty.
\]

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

\[
\Pr \left( \inf_{\mathcal{P}_i} \frac{T^{2(\gamma_0 - \delta) - 1}}{T} \inf_{\mathcal{P}_i} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} (d_t (\vartheta) + s_t (\vartheta))^2 \leq \sigma_0^2 + \varepsilon \right)
\]
where \( v \) if \( \kappa \)

Thus, (19) for \( i = 1, \ldots, 4 \) follows for \( \theta \) small enough by (113) of Lemma 6, noting also that when \( \delta \in \cup_{i=1}^4 I_i \), \( \delta_0 - \delta \geq \delta_0 - \gamma_0 + 1 + \eta \), and by Lemma 8.

Proof of (20) and (21) for \( i = 4 \). Fix \( \zeta \) such that \( 0 < \zeta < \eta \) and let \( \kappa_4 = \gamma_0 - \delta - 1 - \zeta \), noting that \( \kappa_4 \geq \eta - \zeta > 0 \) when \( \delta \in I_4 \). Then, because \( d_t(\theta_0) = 0 \), (20) holds if

\[
(28) \quad \Pr \left( \inf_{\mathcal{H}_4} \frac{T^{2\kappa_4}}{T^{2(\gamma_0 - \delta)}} \left( \sum_{t=1}^T d_t^2(\varphi) - 2 \sum_{t=1}^T d_t(\varphi) s_t(\varphi) \right) \leq 0 \right) \to 0
\]

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 \( \sum_{t=1}^T d_t(\varphi) c_{t-1}(\gamma - \delta, \varphi) = 0 \), it holds that

\[
(29) \quad \sum_{t=1}^T d_t(\varphi) s_t(\varphi) = \sum_{t=1}^T d_t(\varphi) \phi(L; \varphi) u_t(\delta - \delta_0),
\]

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

\[
(30) \quad \Pr \left( \inf_{\mathcal{H}_4} \frac{T^{2\kappa_4}}{T^{2(\gamma_0 - \delta)}} \left( \sum_{t=1}^T d_t^2(\varphi) (1 - 2v_T(\varphi)) - \sum_{t=1}^T s_t^2(\theta_0) \right) \leq 0 \right),
\]

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

\[
(31) \quad \sup_{\mathcal{H}_4} \frac{T^{2\kappa_4}}{T^{2(\gamma_0 - \delta)}} \sum_{t=1}^T s_t^2(\theta_0) = o_p(1),
\]

(32) \Pr \left( \inf_{\mathcal{H}_4} \frac{T^{2\kappa_4}}{T^{2(\gamma_0 - \delta)}} \sum_{t=1}^T d_t^2(\varphi) (1 - 2v_T(\varphi)) \leq \epsilon \right) \to 0 \quad \text{as } T \to \infty.
First, given that \( T^{2^{\alpha_4} - 2(\gamma_0 - \delta) + 1} = T^{-1 - 2\zeta}, \) (31) follows immediately by Lemma 7. Next, fixing \( c \) such that \( 0 < c < 1/2, \) the probability in (32) equals

\[
\Pr \left( \inf_{\hat{\varphi}_T} \frac{T^{2^{\alpha_4}}}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_T^2 (\vartheta) (1 - 2v_T (\vartheta)) \leq \epsilon, \sup_{\hat{\varphi}_T} v_T (\vartheta) \leq c \right) \\
+ \Pr \left( \inf_{\hat{\varphi}_T} \frac{T^{2^{\alpha_4}}}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_T^2 (\vartheta) (1 - 2v_T (\vartheta)) \leq \epsilon, \sup_{\hat{\varphi}_T} v_T (\vartheta) > c \right)
\]

(33)

\[
\leq \Pr \left( \inf_{\hat{\varphi}_T} \frac{T^{2^{\alpha_4}}}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_T^2 (\vartheta) (1 - 2c) \leq \epsilon \right) + \Pr \left( \sup_{\hat{\varphi}_T} v_T (\vartheta) > c \right),
\]

so (32) holds on showing

\[
\lim_{T \to \infty} \inf_{\hat{\varphi}_T} \frac{T^{2^{\alpha_4}}}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_T^2 (\vartheta) > \epsilon,
\]

(34)

\[
\sup_{\hat{\varphi}_T} v_T (\vartheta) = o_p (1).
\]

By the Cauchy-Schwarz inequality, \( \sum_{t=1}^{T} d_T^2 (\vartheta) \geq T^{-1} \overline{d}_T^2 (\vartheta) \), where \( \overline{d}_T (\vartheta) = \sum_{t=1}^{T} d_t (\vartheta) \), so that (34) holds by (117) of Lemma 9. To show (35), note that

\[
\sup_{\hat{\varphi}_T} v_T (\vartheta) \leq \left( \frac{\sup_{\hat{\varphi}_T} T^{-1 - 2\zeta} \sum_{t=1}^{T} (\phi(L; \varphi) u_t (\delta - \delta_0))^2}{\inf_{\hat{\varphi}_T} T^{2^{\alpha_4} - (2(\gamma_0 - \delta) - 1)} \sum_{t=1}^{T} d_T^2 (\vartheta)} \right)^{1/2}
\]

(36)

using \( \alpha_4 = \gamma_0 - \delta - 1 - \zeta \), where \( \sup_{\hat{\varphi}_T} T^{-1 - 2\zeta} \sum_{t=1}^{T} (\phi(L; \varphi) u_t (\delta - \delta_0))^2 = o_p (1) \) by Lemma 2 because \( \delta_0 - \delta \leq 1/2 - \eta \). 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 \( \sum_{t=1}^{T} d_T^2 (\vartheta) \) as in (34) because \( d_t (\vartheta) = 0 \) when \( \gamma = \gamma_0 \). However, we can instead take advantage of \( |\gamma - \gamma_0| \leq \zeta T^{-\alpha_4} \) in \( \hat{\varphi}_T \) and apply the mean value theorem. First note that \( \delta \in \mathcal{I}_4 \) implies that \( \delta_0 - \delta \leq 1/2 - \eta \) and \( \gamma_0 - \delta \geq 1 + \eta \), so that \( u_t (\delta - \delta_0) \) 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 $\delta \in I_4$, and showing also that
\[
\sup_{H_4} \frac{1}{T} \left( \sum_{t=1}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) h_{t-1,T} (\gamma - \delta, \varphi) \right)^2 = o_p(1),
\]
\[
(37)
\]
\[
(38)
\]
From (106) of Lemma 5, the left-hand side of (37) is $O_p(T^{-2\eta}) = o_p(1)$ by choosing $\theta < 1/2 - \eta$. Next, because $|\gamma - \gamma_0| < \xi T^{-\kappa_4}$ in $H_4$, by (110) and (112) of Lemma 6 the left-hand side of (38) is $O_p(T^{-\eta + 2\theta}) = o_p(1)$ for $\theta$ small enough because $\zeta < \eta$.

**Proof of (20) and (21) for $i = 3$.** Fix $\kappa_3 = \gamma_0 - \delta - 1$, so noting that $\delta \in I_3$, $\kappa_3 \geq \gamma_0 - \delta_0 - 1/2 - \eta > 0$. Then, by the Cauchy-Schwarz inequality,
\[
\Pr \left( \inf_{H_3} s_T (\vartheta) \leq 0 \right) \leq \Pr \left( \inf_{H_3} \frac{1}{T^2} (d_T (\vartheta) + \pi_T (\vartheta))^2 - 1 \sum_{t=1}^{T} s^2_t (\vartheta_0) \leq 0 \right),
\]
where $d_T (\vartheta) = \sum_{t=1}^{T} d_t (\vartheta)$ and
\[
\pi_T (\vartheta) = \sum_{t=1}^{T} s_t (\vartheta) = \phi(L; \varphi) u_T (\delta - \delta_0 - 1)
\]
\[
- \sum_{t=1}^{T} h_{t-1, T} (\gamma - \delta, \varphi) \sum_{j=1}^{T} \phi(L; \varphi) u_j (\delta - \delta_0) h_{j-1, T} (\gamma - \delta, \varphi).
\]
\[
(40)
\]
The right-hand side of (39) is thus bounded by
\[
\Pr \left( \inf_{H_3} \frac{1}{T^2} d_T^2 (\vartheta) (1 - 2 |\pi_T (\vartheta)|) - 1 \sum_{t=1}^{T} s^2_t (\vartheta_0) \leq 0 \right),
\]
where $\pi_T (\vartheta) = \pi_T (\vartheta) / d_T (\vartheta)$. Applying Lemma 7, (20) for $i = 3$ would then hold if
\[
\Pr \left( \inf_{H_3} \frac{1}{T^2} d_T^2 (\vartheta) (1 - 2 |\pi_T (\vartheta)|) \leq K \right) \to 0 \text{ as } T \to \infty,
\]
for an arbitrarily large $K$. As in (33), fixing $c$ such that $0 < c < 1/2$, the probability in (42) is bounded by
\[
\Pr \left( \inf_{H_3} \frac{1}{T^2} d_T^2 (\vartheta) (1 - 2c) \leq K \right) + \Pr \left( \sup_{H_3} |\pi_T (\vartheta)| > c \right),
\]
\[
(43)
\]
so, as in (36), (42) holds if

\begin{equation}
\sup_{F_3} \frac{1}{T} |\mathcal{S}_T(\vartheta)| = O_p(1),
\end{equation}

\begin{equation}
\lim_{T \to \infty} \inf_{F_3} \frac{1}{T^2} \mathcal{S}_T(\vartheta)^2 > K.
\end{equation}

For \( \delta \in I_3 \) it holds that \( \delta_0 - \delta \leq 1/2 \), so in view of (40) the proof of (44) is immediate using (98) in Lemma 1 together with Lemmas 2 and 5 with \( \theta < 1/2 \). Finally, given that \( T^{\gamma_0 - \delta} = T^{-1} \), the proof of (45) follows by Lemma 9, to conclude the proof of (20) for \( i = 3 \).

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

\begin{equation}
\Pr \left( \inf_{\tau \in T_3} \frac{1}{T} \left( \sum_{t=1}^T s_t^2(\vartheta) - 2 \sum_{t=1}^T d_t(\vartheta) s_t(\vartheta) - \sum_{t=1}^T s_t^2(\vartheta_0) \right) \leq 0 \right) \to 0 \text{ as } T \to \infty,
\end{equation}

where

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

In the proof of their (2.7) for \( i = 3 \), [10] showed that

\begin{equation}
\Pr \left( \inf_{\tau \in T_3} \frac{1}{T} \sum_{t=1}^T (\phi(L; \varphi) u_t(\delta - \delta_0))^2 > K \right) \to 1 \text{ as } T \to \infty
\end{equation}

for any arbitrarily large fixed constant \( K \) (for small enough \( \eta \)). Thus, noting (46), (21) for \( i = 3 \) holds by (47) and Lemma 7 on showing

\begin{equation}
\sup_{H_3} \frac{1}{T} \left( \sum_{t=1}^T \phi(L; \varphi) u_t(\delta - \delta_0) h_{t-1,T}(\gamma - \delta, \varphi) \right)^2 = O_p(1),
\end{equation}

\begin{equation}
\sup_{H_3} \frac{1}{T} \left| \sum_{t=1}^T d_t(\vartheta) s_t(\vartheta) \right| = O_p(\xi),
\end{equation}

because, even if \( \xi \) had to be set large enough in the proof of (45) (see the proof of Lemma 9), this can be dominated by the constant \( K \) fixed in (47), which can be chosen arbitrarily large by setting \( \eta \) small enough. First, by (106) of Lemma 5, the left-hand side of (48) holds by choosing \( \theta < 1/2 \) because \( \delta_0 - \delta \leq 1/2 \) when \( \delta \in I_3 \). Next, noting that \( \sup_{H_3} |\gamma - \gamma_0| \leq \xi T^{-\gamma_3} \) and that \( \delta \in I_3 \) implies \( \gamma_0 - \delta \geq \gamma_0 - \delta_0 + 1/2 - \eta > 1 \) and \( \delta_0 - \delta \leq 1/2 \), it follows by (110) and (114) of Lemma 6 that the left-hand side of (49) is \( \xi O_p(1) = O_p(\xi) \) by choosing \( \theta < 1/2 \).
Proof of (20) and (21) for $i = 2$. Fix $\gamma_2 = \gamma_0 - \delta_0 - 1/2 > 0$. 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

$$\Pr \left( \inf_{H \in \mathcal{H}_2} \frac{1}{T^{2(\delta_0 - \delta) + 1}} \overline{d}_T^2(\vartheta) (1 - 2|\varpi_T(\vartheta)|) - \sup_{H \in \mathcal{H}_2} \frac{1}{T^{2(\delta_0 - \delta)}} \sum_{t=1}^{T} s_t^2(\vartheta_0) \leq 0 \right).$$

Then, given Lemma 7,

$$\sup_{H \in \mathcal{H}_2} \frac{1}{T^{2(\delta_0 - \delta) + 1}} \sum_{t=1}^{T} s_t^2(\vartheta_0) = \sup_{H \in \mathcal{H}_2} \frac{T}{T^{2(\delta_0 - \delta)}} \frac{1}{T} \sum_{t=1}^{T} s_t^2(\vartheta_0) = O_p(1),$$

because when $\delta \in I_2$, $\delta_0 - \delta \geq 1/2$. Thus (20) for $i = 2$ would hold if

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

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

$$\sup_{H \in \mathcal{H}_2} \frac{1}{T^{2(\delta_0 - \delta) + 1/2}} |\varpi_T(\vartheta)| = O_p(1),$$

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

The proof of (52) is almost identical to that of (44), again applying Lemmas 1, 2, and 5. Also, given that $T^{\gamma_2 - (\gamma_0 - \delta)} = T^{-(\delta_0 - \delta) - 1/2}$, (53) follows by Lemma 9, to conclude the proof of (20) for $i = 2$.

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

$$\Pr \left( \inf_{H \in \mathcal{H}_2} \frac{1}{T^{2(\delta_0 - \delta)}} \left( \sum_{t=1}^{T} s_t^2(\vartheta) - 2 \left| \sum_{t=1}^{T} d_t(\vartheta) s_t(\vartheta) \right| - \sum_{t=1}^{T} s_t^2(\vartheta_0) \right) \leq 0 \right) \to 0$$

as $T \to \infty$. In the proof of their (2.7) for $i = 2$, [10] showed that

$$\Pr \left( \inf_{\|\tau - \tau_0\| \geq \varepsilon, \tau \in \mathcal{I}_2} \frac{1}{T^{2(\delta_0 - \delta)}} \sum_{t=1}^{T} (\phi(L; \varphi) u_t(\delta - \delta_0))^2 \geq K \right) \to 1$$

as $T \to \infty$ for any arbitrarily large fixed constant $K$ (for small enough $\eta$). Thus, in view of (46), (50), (55), and Lemma 5 with $\theta < 1/2$, it follows that (54) holds if

$$\frac{1}{T^{2(\delta_0 - \delta)}} \left| \sum_{t=1}^{T} d_t(\vartheta) s_t(\vartheta) \right| = O_p(\xi).$$
Again, noting that \( \sup_{H_2} |\gamma - \gamma_0| \leq \xi T^{-\kappa_2} \) and that \( \delta \in I_2 \) implies \( \gamma_0 - \delta \geq \gamma_0 - \delta_0 + 1/2 > 1 \) and \( \delta_0 - \delta \geq 1/2 \), (56) follows from (110) and (115) of Lemma 6 setting \( \theta < 1/2 \).

**Proof of (20) and (21) for \( i = 1 \).** Fix \( \kappa_1 = \gamma_0 - \delta_0 - 1/2 > 0 \). As in the treatment of (28), (20) for \( i = 1 \) holds if

\[
(57) \quad \sup_{H_1} \frac{T^{2\kappa_1}}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} s_t^2 (\vartheta_0) = o_p(1),
\]

(58) \[ \Pr \left( \inf_{H_1} \frac{T^{2\kappa_1}}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_t^2 (\vartheta) (1 - 2v_T (\vartheta)) \leq \epsilon \right) \to 0 \text{ as } T \to \infty, \]

for an arbitrarily small \( \epsilon \). First, (57) holds by Lemma 7, noting that \( 2\kappa_1 - 2 (\gamma_0 - \delta) + 1 = 2 (\delta - \delta_0) \) and \( \sup_{H_1} 2 (\delta - \delta_0) = -1 - 2\eta < -1 \). Next, as in the proof of (51), see also (32) and (36), (58) follows if

\[
(59) \quad \sup_{H_1} \frac{1}{T^{2(\delta_0 - \delta)}} \sum_{t=1}^{T} \phi (L; \varphi) u_t (\delta - \delta_0)^2 = O_p (1),
\]

(60) \[ \lim_{T \to \infty} \inf_{H_1} \frac{1}{T^{2(\delta_0 - \delta)}} \sum_{t=1}^{T} d_t^2 (\vartheta) > K. \]

First, (59) follows immediately from (100) of Lemma 2, noting that \( \delta_0 - \delta \geq 1/2 + \eta \). Next, by the Cauchy-Schwarz inequality, noting that \( T^{\kappa_1 - (\gamma_0 - \delta)} = T^{-(\delta_0 - \delta) - 1/2} \), (60) follows by Lemma 9.

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

\[
(61) \quad \Pr \left( \inf_{H_1} \frac{1}{T^{2(\delta_0 - \delta) + 1}} \left( \sum_{t=1}^{T} (d_t (\vartheta) + s_t (\vartheta))^2 - \sum_{t=1}^{T} s_t^2 (\vartheta_0) \right) \leq 0 \right) \to 0
\]

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

\[
\Pr \left( \inf_{H_1} \frac{1}{T^{2(\delta_0 - \delta) + 1}} \left( \left( \bar{d}_T (\vartheta) + \bar{s}_T (\vartheta) \right)^2 - \sup_{H_1} \frac{1}{T^{2(\delta_0 - \delta)}} \sum_{t=1}^{T} s_t^2 (\vartheta_0) \leq 0 \right) \right).
\]

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

(62) \[ \Pr \left( \inf_{H_1} \frac{1}{T^{2(\delta_0 - \delta) + 1}} \left( T^{\kappa_1} (\gamma - \gamma_0) \frac{1}{T^{\kappa_1}} \frac{\partial d_T (\tau, \gamma)}{\partial \gamma} + \bar{s}_T (\vartheta) \right)^2 > \epsilon \right) \to 1 \]
as $T \to \infty$, where $|\tau - \gamma_0| \leq |\gamma - \gamma_0|$. Note that in $\mathcal{H}_1$, $\gamma_0 - \delta \geq 1 + \eta$, so there exists $\alpha > 0$ such that for $T$ sufficiently large $\gamma - \delta \geq 1 + \alpha$. Let $g = T^{x_1} (\gamma - \gamma_0)$ and $\varpi_T (\tau, \gamma) = \varpi_T (\varnothing)$. Define the set $\mathcal{G}_1 = \{(g, \tau, \gamma_1, \gamma_2 : |g| < \xi, \tau \in \mathcal{T}_1, \gamma_i - \delta \geq 1 + \alpha, \gamma_i \in [\square_1, \square_2], i = 1, 2\}$. Then, noting that two different values of $\gamma$ appear within the probability in (62), (21) for $i = 1$ holds if

$$
\Pr \left( \inf_{\mathcal{G}_1} \left( \frac{1}{T^{x_1 + \delta_0 - \delta + 1/2}} \frac{\partial \varpi_T (\tau, \gamma_1)}{\partial \gamma} + \frac{1}{T^{\delta_0 - \delta + 1/2}} \varpi_T (\tau, \gamma_2) \right)^2 > \epsilon \right) \to 1
$$

as $T \to \infty$. First, by (116) in Lemma 9, noting that $x_1 + \delta_0 - \delta + 1/2 = \gamma_0 - \delta$,

$$
\frac{1}{T^{\gamma_0 - \delta}} \frac{\partial \varpi_T (\tau, \gamma_1)}{\partial \gamma} = \mu_0 \phi (1; \varphi) D (\delta, \gamma_1) + m_{1,T} (\tau, \gamma_1),
$$

where

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

and $\sup_{\mathcal{G}_1} |m_{1,T} (\tau, \gamma_1)| = o_p (1)$. Next, it is shown that

$$
\frac{1}{T^{\delta_0 - \delta + 1/2}} \varpi_T (\tau, \gamma_2) = \phi (1; \varphi) V_T (\delta, \gamma_2) + m_{2,T} (\tau, \gamma_2),
$$

where $\sup_{\mathcal{G}_1} |m_{2,T} (\tau, \gamma_2)| = o_p (1)$ and, considering $V_T (\delta, \gamma)$ as a continuous process indexed by $(\delta, \gamma)$,

$$
V_T (\delta, \gamma) \Rightarrow V (\delta, \gamma).
$$

Here, $\Rightarrow$ means weak convergence in the space of continuous functions on $\mathcal{L} = \{\delta, \gamma : \delta \in \mathcal{I}_1, \gamma - \delta \geq 1 + \alpha, \gamma \in [\square_1, \square_2]\}$ endowed with the uniform topology and $V (\delta, \gamma)$ is a fractional Brownian motion functional. The proofs of (64) and (65) are quite similar to derivations in [10, 13, 18], so these are given in Section S.3 of [9], along with the precise definitions of $V_T (\delta, \gamma)$ and $V (\delta, \gamma)$.

Then, noting (8) and defining $\mathcal{L}_1 = \{g, \delta, \gamma_1, \gamma_2 : |g| \leq \xi, \delta \in \mathcal{I}_1, \gamma_i - \delta \geq 1 + \alpha, \gamma_i \in [\square_1, \square_2], i = 1, 2\}$, (63) holds if

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

It follows from (65) and the continuous mapping theorem that, as $T \to \infty$,

$$
\inf_{\mathcal{L}_1} (\mu_0 g D (\delta, \gamma_1) + V_T (\delta, \gamma_2))^2 \rightarrow_d \inf_{\mathcal{L}_1} (\mu_0 g D (\delta, \gamma_1) + V (\delta, \gamma_2))^2,
$$
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 $T \to \infty$,
\[
\Pr \left( \inf_{\ell_1} (\mu_0 g_D (\delta, \gamma_1) + V_T (\delta, \gamma_2))^2 > \epsilon \right) \to \Pr \left( \inf_{\ell_1} (\mu_0 g_D (\delta, \gamma_1) + V (\delta, \gamma_2))^2 > \epsilon \right),
\]
and (66) follows because $\epsilon$ is arbitrarily small, to conclude the proof of (21) for $i = 1$ and therefore that for (15).

5.1.3. Proof of (16). In this case, let $R_T (\tau, \gamma) = R_T (\vartheta), d_t (\tau, \gamma) = d_t (\vartheta)$, and $s_t (\tau, \gamma) = s_t (\vartheta) = s_{1t} (\tau) - s_{2t} (\vartheta)$ with $s_{1t} (\tau) = \phi (L; \varphi) u_t (\delta - \delta_0)$ and $s_{2t} (\tau, \gamma) = s_{2t} (\vartheta) = h_{t-1, T} (\gamma - \delta, \varphi) \sum_{j=1}^{T} s_{1j} (\tau) h_{j-1, T} (\gamma - \delta, \varphi)$, so that, noting $\sum_{t=1}^{T} h_{t-1, T}^2 (\gamma - \delta, \varphi) = 1$,
\[
\sum_{t=1}^{T} s_{2t}^2 (\vartheta) = \sum_{t=1}^{T} s_{1t} (\tau) s_{2t} (\vartheta) = \left( \sum_{j=1}^{T} s_{1j} (\tau) h_{j-1, T} (\gamma - \delta, \varphi) \right)^2.
\]
Noting also (29),
\[
R_T (\vartheta) = \frac{1}{T} \sum_{t=1}^{T} d_t^2 (\vartheta) + \frac{1}{T} \sum_{t=1}^{T} s_{1t}^2 (\tau) - \frac{1}{T} \left( \sum_{j=1}^{T} s_{1j} (\tau) h_{j-1, T} (\gamma - \delta, \varphi) \right)^2 + \frac{2}{T} \sum_{t=1}^{T} d_t (\vartheta) s_{1t} (\vartheta).
\]
Clearly, if $\hat{\vartheta} \in \mathcal{N}_\varepsilon \cap M_\varepsilon$, then $\inf_{\mathcal{N}_\varepsilon \cap M_\varepsilon} R_T (\hat{\vartheta}, \gamma) \leq R_T (\hat{\vartheta}, \gamma_0)$, so that
\[
\Pr (\hat{\vartheta} \in \mathcal{N}_\varepsilon \cap M_\varepsilon) \leq \Pr \left( \hat{\vartheta} \in \mathcal{N}_\varepsilon \cap M_\varepsilon, \inf_{\mathcal{N}_\varepsilon \cap M_\varepsilon} R_T (\hat{\vartheta}, \gamma) - R_T (\hat{\vartheta}, \gamma_0) \leq 0 \right).
\]
Recalling that $d_t (\tau, \gamma_0) = 0, R_T (\hat{\vartheta}, \gamma_0) = T^{-1} \sum_{t=1}^{T} s_{1t} (\hat{\vartheta})$ and this cancels with the corresponding term in $R_T (\hat{\vartheta}, \gamma)$, see (67). Thus, (16) holds if
\[
\lim_{T \to \infty} \inf_{\vartheta \in \mathcal{N}_\varepsilon \cap M_\varepsilon} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_t^2 (\vartheta) > \epsilon,
\]
\[
\sup_{\vartheta \in \mathcal{N}_\varepsilon \cap M_\varepsilon} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \left| \sum_{t=1}^{T} d_t (\vartheta) s_{1t} (\tau) \right| = o_p (1),
\]
\[
\sup_{\vartheta \in \mathcal{N}_\varepsilon \cap M_\varepsilon} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \left( \sum_{j=1}^{T} s_{1j} (\tau) h_{j-1, T} (\gamma - \delta, \varphi) \right)^2 = o_p (1),
\]

\[\text{ESTIMATION OF FRACTIONAL TIME SERIES WITH TRENDS} \quad 21\]
noting the change in the normalization compared with (68) \((T^{2(\gamma_0-\delta_0)}-1)\) instead of \(T\), which is justified because the right-hand side of the inequality inside the probability in (68) is 0, so multiplying the left- and right-hand sides of the inequality by a positive number does not alter the probability.

First, (69) follows from Lemma 8, noting that in \(\mathbb{N}_\varepsilon \cap M_\varepsilon\), \(\gamma_0 - \delta \geq \gamma_0 - \delta_0 - \varepsilon > 1/2\) setting \(\varepsilon\) small enough. Next, letting both \(\varepsilon\) and \(\theta\) be sufficiently small and noting that in \(\mathbb{N}_\varepsilon \cap M_\varepsilon\), \(\delta_0 - \delta \geq -\varepsilon\), by (113) of Lemma 6 the left-hand side of (70) is \(O_p(T^{1/2}+\delta_0-\gamma_0+3\theta+\varepsilon) = o_p(1)\). Finally, by (107) of Lemma 5 the left-hand side of (71) is \(O_p(T^{1/2+\delta_0-\gamma_0+3\theta+\varepsilon}) = o_p(1)\), to conclude the proof of (16) and therefore that of consistency of \(\hat{\vartheta}\).

5.2. Proof of Theorem 1(ii): the \(\gamma_0 - 1/2 < \delta_0\) case. Clearly
\[
\Pr (||\hat{\varphi} - \varphi_0|| \geq \varepsilon) = \Pr \left( \inf_{\varphi \in \mathbb{M}_\varepsilon} R_T (\varphi) \leq \inf_{\varphi \in \mathbb{M}_\varepsilon} R_T (\varphi) \right),
\]
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 \(o(1)\), or, in view of Lemma 7, that
\[
(72) \quad \Pr \left( \inf_{\varphi \in \mathbb{M}_\varepsilon} R_T (\varphi) \leq \sigma_0^2 + \varepsilon \right) \rightarrow 0 \text{ as } T \rightarrow \infty.
\]
The proof of (72), which uses very similar techniques to those employed in the proof of (15), is given in Section S.4 of [9] and completes the proof of consistency of \(\hat{\varphi}\).

5.3. Proof of Theorem 2(i): the \(\gamma_0 - 1/2 > \delta_0\) case. We first show that
\[
(73) \quad T^{1/2} M^{-1}_T (\hat{\varphi} - \varphi_0) \rightarrow_d N (0, \sigma_0^2 V^{-1}).
\]
By the mean value theorem,
\[
(74) \quad \hat{\varphi} - \varphi_0 = - \left( \frac{\partial^2 R_T (\varphi)}{\partial \varphi \partial \varphi'} \right)^{-1} \frac{\partial R_T (\varphi_0)}{\partial \varphi},
\]
where \(\varphi\) represents an intermediate point which is allowed to vary across the different rows of \(\partial^2 R_T (\cdot) / \partial \varphi \partial \varphi'\).

We first analyze the score in (74). It can be easily seen that \(\partial d_t (\varphi_0) / \partial \tau = 0\) and \(\partial s_{1t} (\tau) / \partial \gamma = 0\), so, recalling that \(d_t (\varphi_0) = 0\) and the decomposition (67),
\[
\frac{\partial R_T (\varphi_0)}{\partial \varphi} = \frac{2}{T} \sum_{t=1}^{T} s_t (\varphi_0) \left[ \left( \frac{\partial s_{1t} (\tau_0)}{\partial \tau} \right) - \frac{\partial s_{2t} (\varphi_0)}{\partial \tau} \right].
\]
With appropriate normalization it is fairly straightforward, using previous arguments, to show that

\[
\frac{T^{1/2}}{2} M T \frac{\partial R_T (\vartheta_0)}{\partial \vartheta} = \sum_{t=2}^{T} \varepsilon_t \eta_{t,T} + o_p \left(1\right),
\]

where \( \eta_{t,T} = \left( T^{-1/2} \sum_{j=1}^{\infty} m_j (\varphi_0) \varepsilon_{t-j}, T^\delta_0 + 1/2 - \gamma_0 \mu_0 \phi(1; \varphi_0) g_{t,T} (\gamma_0 - \delta_0) \right)' \) and \( m_j (\varphi_0) = \left( -j^{-1}, b_j' (\varphi_0) \right)' \). The proof of (75) is given in Section S.5 of [9]. Defining \( F_{t,T} = F_t \) for any \( 1 \leq t \leq T \), Assumption A2 implies that \( \{ \varepsilon_t \eta_{t,T}, F_{t,T}, 1 \leq t \leq T, T \geq 1 \} \) is a martingale difference array. For any \( (p+2) \)-dimensional vector \( \xi \), define \( \xi_{t,T} = \varepsilon_t \xi' \eta_{t,T} / \sigma_0 (\xi' V \xi)^{1/2} \) and \( B_T^2 = \sum_{t=2}^{T} E (\xi_{t,T}^2 | F_{t-1,T}) \). Then, by Corollary 3.1 of [7], if

\[
B_T^2 \rightarrow_p 1,
\]

and, for all \( \epsilon > 0 \),

\[
\sum_{t=2}^{T} E \left( \xi_{t,T}^4 | (| \xi_{t,T} | > \epsilon) \right) \rightarrow_p 0,
\]

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

\[
\sum_{t=2}^{T} \varepsilon_t \eta_{t,T} \rightarrow_d N \left(0, \sigma_0^2 V \right)
\]

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

\[
E \left( \xi_{t,T}^2 | F_{t-1,T} \right) = \frac{\xi' \eta_{t,T} \eta_{t,T}' \xi}{\xi' V \xi},
\]

so that (76) holds if \( \sum_{t=2}^{T} \eta_{t,T} \eta_{t,T}' \rightarrow_p V \). However, this follows straightforwardly by the same arguments as in the proof of (2.55) of [10] and Lemma 10 because \( \sum_{t=1}^{T} \sum_{j=1}^{\infty} m_j (\varphi_0) \varepsilon_{t-j} = O_p (t^{1/2}) \), which implies, by summation by parts, that

\[
\frac{1}{T^{\gamma_0 - \delta_0}} \sum_{t=2}^{T} g_{t,T} (\gamma_0 - \delta_0) \sum_{j=1}^{\infty} m_j (\varphi_0) \varepsilon_{t-j} = o_p \left(1\right).
\]

Now (77) holds if, e.g., \( \sum_{t=2}^{T} E (\xi_{t,T}^4 | F_{t-1,T}) \rightarrow_p 0 \), which, given that the fourth moment of \( \varepsilon_t \) is finite, holds if \( \sum_{t=2}^{T} (\xi' \eta_{t,T} \eta_{t,T}' \xi)^2 \rightarrow_p 0 \), and this can be easily justified by previous arguments. This completes the proof of (78).
Next, noting (74), (75), and (78), the proof of (73) is completed by showing

\[ \left( \frac{\partial^2 R_T(\bar{\vartheta})}{\partial \vartheta \partial \vartheta'} - \frac{\partial^2 R_T(\vartheta_0)}{\partial \vartheta \partial \vartheta'} \right) M_T = o_p (1) \]

and

\[ \frac{1}{2} \left( \frac{\partial^2 R_T(\vartheta_0)}{\partial \vartheta \partial \vartheta'} \right) M_T \rightarrow_p V. \]

In Section S.6 of [9] it is shown that, using similar arguments to those in the proof of Theorem 1(i), for some fixed \( \kappa > 0 \),

\[ T^{-\kappa} (\hat{\vartheta} - \vartheta_0) \rightarrow_p 0, \]

and in light of this the proof of (80) is relatively straightforward. It consists of deriving all terms in \( \frac{\partial^2 R_T(\vartheta_0)}{\partial \vartheta \partial \vartheta'} \) and checking that the differences with respect the corresponding ones in \( \frac{\partial^2 R_T(\vartheta_0)}{\partial \vartheta \partial \vartheta'} \) satisfy (80). This requires the use of the mean value theorem and Assumption A4(ii), where typically the derivatives involve additional log factors which are compensated by the factor \( T^{-\kappa} \) that arises because \( \vartheta - \vartheta_0 = O_p (T^{-\kappa}) \).

Now we show (81). In Section S.7 of [9] it is shown, using previous arguments, that

\[ \frac{1}{2} T^{-2} \sum_{t=1}^{T} \left( \frac{\partial s_{11}(\tau_0)}{\partial \gamma} \right)^2 \rightarrow_p A, \]

so (81) holds by showing that, as \( T \rightarrow \infty \),

\[ \frac{1}{T} \sum_{t=1}^{T} \frac{\partial s_{11}(\tau_0)}{\partial \gamma} \left( \frac{\partial d_t(\vartheta_0)}{\partial \gamma} \right)^2 \rightarrow \sum_{t=1}^{T} \frac{\mu_t^2 \phi^2 (1; \varphi_0)}{\Gamma^2 (\gamma_0 - \delta_0) (2 (\gamma_0 - \delta_0) - 1)^3}. \]

Here, (84) follows from (2.53) of [10] and (85) follows by arguments used in the proof of (76).

Next, given (73), the remaining part of (13) is justified as follows. Noting

\[ \phi (L; \varphi) x_t (\delta) = \mu_0 c_{t-1} (\gamma_0 - \delta, \varphi) + \phi (L; \varphi) u_t (\delta - \delta_0), \]

it follows from (11) that

\[ \hat{\mu} = \hat{\mu} (\bar{\vartheta}) = \mu_0 \sum_{t=1}^{T} c_{t-1} (\gamma_0 - \delta, \varphi) k_{t-1,T} (\gamma - \delta, \varphi) + \sum_{t=1}^{T} \phi (L; \varphi) u_t (\delta - \delta_0) k_{t-1,T} (\gamma - \delta, \varphi), \]
where $k_{t-1,T}(\gamma - \delta, \varphi) = c_{t-1} (\gamma - \delta, \varphi) / \sum_{t=1}^{T} c_{t}^{2} (\gamma - \delta, \varphi)$. By straightforward application of the mean value theorem,

$$k_{t-1,T}(\tilde{\gamma} - \tilde{\delta}, \tilde{\varphi}) = k_{t-1,T}(\gamma_{0} - \tilde{\delta}, \tilde{\varphi}) + k_{t-1,T}^{(1)}(\gamma_{0}, \tilde{\delta}, \tilde{\varphi})(\tilde{\gamma} - \gamma_{0}),$$

where $k_{t-1,T}^{(1)}(\cdot, \cdot)$ is the derivative of $k_{t,T}(\cdot, \cdot)$ with respect to the first argument and $|\tilde{\gamma} - \gamma_{0}| \leq |\tilde{\gamma} - \gamma_{0}|$. Thus,

$$\hat{\mu} = \mu_{0} + \mu_{0}(\gamma_{0} - \gamma_{0}) \sum_{t=1}^{T} c_{t-1}(\gamma_{0} - \tilde{\delta}, \tilde{\varphi})k_{t-1,T}^{(1)}(\gamma - \tilde{\delta}, \tilde{\varphi}) + \sum_{t=1}^{T} \phi(L; \tilde{\varphi}) u_{t} (\tilde{\delta} - \delta_{0}) k_{t-1,T}(\tilde{\gamma} - \tilde{\delta}, \tilde{\varphi}),$$

which implies that

$$\frac{T^{\gamma_{0} - \delta_{0} - 1/2}}{\log T} (\hat{\mu} - \mu_{0}) = \mu_{0} T^{\gamma_{0} - \delta_{0} - 1/2}(\tilde{\gamma} - \gamma_{0}) \frac{1}{\log T} \sum_{t=1}^{T} c_{t-1}(\gamma_{0} - \tilde{\delta}, \tilde{\varphi})k_{t-1,T}^{(1)}(\gamma - \tilde{\delta}, \tilde{\varphi})$$

$$+ \frac{T^{\gamma_{0} - \delta_{0} - 1/2}}{\log T} \sum_{t=1}^{T} \phi(L; \tilde{\varphi}) u_{t} (\tilde{\delta} - \delta_{0}) k_{t-1,T}(\tilde{\gamma} - \tilde{\delta}, \tilde{\varphi}).$$

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

$$(87) \quad \frac{1}{\log T} \sum_{t=1}^{T} c_{t-1}(\gamma_{0} - \tilde{\delta}, \tilde{\varphi})k_{t-1,T}^{(1)}(\gamma - \tilde{\delta}, \tilde{\varphi}) \rightarrow_{p} 1,$$

$$(88) \quad \frac{T^{\gamma_{0} - \delta_{0} - 1/2}}{\log T} \sum_{t=1}^{T} \phi(L; \tilde{\varphi}) u_{t} (\tilde{\delta} - \delta_{0}) k_{t-1,T}(\tilde{\gamma} - \tilde{\delta}, \tilde{\varphi}) = o_{p}(1),$$

the proofs of which are based on previous arguments and are given in Section S.8 in [9].

5.4. Proof of Theorem 2(ii): the $\gamma_{0} - 1/2 < \delta_{0}$ case. First, noting (67), the loss function $R_{T}(\vartheta)$ can be decomposed in the sum of two terms, $R_{T}(\vartheta) = Q_{T}(\vartheta) + S_{T}(\vartheta)$, where $Q_{T}(\vartheta) = T^{-1} \sum_{t=1}^{T} s_{2t}^{2}(\vartheta)$ and

$$S_{T}(\vartheta) = \frac{1}{T} \sum_{t=1}^{T} (d_{t}(\vartheta) - s_{2t}(\vartheta))^{2} + \frac{2}{T} \sum_{t=1}^{T} s_{1t}(\vartheta) (d_{t}(\vartheta) - s_{2t}(\vartheta)).$$

Thus, $Q_{T}(\vartheta)$ is the loss function in [10]. Now

$$\frac{\partial R_{T}(\vartheta)}{\partial \vartheta} = 0 = \frac{\partial Q_{T}(\vartheta)}{\partial \vartheta} + \frac{\partial S_{T}(\vartheta)}{\partial \vartheta},$$

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

\[
\frac{\partial Q_T(\hat{\tau})}{\partial \tau} = \frac{\partial Q_T(\tau_0)}{\partial \tau} + \frac{\partial^2 Q_T(\bar{\tau})}{\partial \tau \partial \tau'} (\hat{\tau} - \tau_0),
\]

where \( \bar{\tau} \) represents an intermediate point between \( \hat{\tau} \) and \( \tau_0 \) which is allowed to vary in the different rows of \( \frac{\partial^2 Q_T(\cdot)}{\partial \tau \partial \tau'} \). Inserting (91) in (90) we then find

\[
T^{1/2} (\hat{\tau} - \tau_0) = - \left( \frac{\partial^2 Q_T(\bar{\tau})}{\partial \tau \partial \tau'} \right)^{-1} T^{1/2} \frac{\partial Q_T(\tau_0)}{\partial \tau} - \left( \frac{\partial^2 Q_T(\bar{\tau})}{\partial \tau \partial \tau'} \right)^{-1} T^{1/2} \frac{\partial S_T(\hat{\theta})}{\partial \tau}.
\]

Now, by [10] (see the proof of their Theorem 2.2), the first term on the right-hand side of (92) has a \( N(0, A^{-1}) \) limiting distribution, and \( \frac{\partial^2 Q_T(\bar{\tau})}{\partial \tau \partial \tau'} \) converges in probability to a nonsingular matrix. Thus, in view of (92), Theorem 2(ii) follows if

\[
T^{1/2} \frac{\partial S_T(\hat{\theta})}{\partial \tau} = o_p(1),
\]

which is proven in Section S.9 in [9] based on previous arguments.

6. Auxiliary lemmas.

**Lemma 1.** Let \( \theta \) 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 \( T \geq 1 \),

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

\[
\sup_{d \geq 1/2 + \theta, \phi \in \Psi} \left| \frac{\partial}{\partial d} h_{t-1, T}(d, \phi) \right| = O(t^{-1/2 + \theta} T^{-\theta} (1 + \log (t/T))),
\]

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

\[
\sup_{d \geq 1/2 + \theta, \phi \in \Psi} \left| \frac{\partial}{\partial d} (h_{t, T}(d, \phi) - h_{t-1, T}(d, \phi)) \right| = O(t^{-3/2 + \theta} T^{-\theta} (1 + \log (t/T))),
\]

\[
\sup_{d \in [d_1, d_2], \phi \in \Psi} \left| \sum_{t=1}^T h_{t-1, T}(d, \phi) \right| = O(T^{1/2}).
\]
LEMMA 2. Under Assumptions A1–A3, uniformly in $t = 1, \ldots, T$, $T \geq 1$, and $\varphi \in \Psi$, 

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

(99)

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

(100)

LEMMA 3. Let $\theta$ be an arbitrary number such that $0 < \theta < \varsigma - 1/2$. Then, under Assumptions A1 and A3, for $m = 0, 1$ and uniformly in $\vartheta \in \Xi$,

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

(101)

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

(102)

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

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

(103)

LEMMA 4. Under Assumptions A1–A3, uniformly in $\vartheta \in \Xi$,

$$
\sup_{\delta_0 - \delta \leq \gamma_1, \gamma_0 - \delta \leq \gamma_2} \frac{1}{T} \left| \sum_{t=1}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) c_{t-1} (\gamma_0 - \delta, \varphi) \right|
$$

(104)

$$
= O_p(T^{\max\{g_2-1, -\varsigma\}+g_1-1/2} + T^{-1} \log^2 T + T^{\max\{g_2-1, -\varsigma\}-1}(\log T) \mathbb{I}(g_1 \leq -1/2)),
$$

$$
\sup_{\delta_0 - \delta \geq \gamma_1, \gamma_0 - \delta \geq \gamma_2} \frac{1}{T_{\gamma_0 + \delta_0 - 2\delta}} \left| \sum_{t=1}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) c_{t-1} (\gamma_0 - \delta, \varphi) \right|
$$

(105)

$$
= O_p(T^{\max\{g_2-1, -\varsigma\}-g_2+1/2} + T^{-g_1-g_2} \log^2 T + T^{\max\{g_2-1, -\varsigma\}-g_1-g_2}(\log T) \mathbb{I}(g_1 \leq -1/2)).
$$
Lemma 5. Let \( \theta \) be an arbitrary number such that \( 0 < \theta < \varsigma - 1/2 \). Then, under Assumptions A1–A3, for \( m = 0,1 \), and uniformly in \( \vartheta \in \Xi \),

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

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

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

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

Lemma 6. Under Assumptions A1–A3, for any \( g_2 > 1/2 \) and for any arbitrary \( \theta \) such that \( 0 < \theta < \min\{\varsigma - 1/2, g_2 - 1/2\} \),

\[
\left| \sum_{t=1}^{T} d_t (\vartheta) s_t (\vartheta) \right| \leq |\gamma - \gamma_0| |M_T (\vartheta)|,
\]

where, uniformly in \( \vartheta \in \Xi \),

\[
\sup_{\delta_0 - \delta \leq g_1, \gamma - \delta \leq g_2} |M_T (\vartheta)| = O_p(T_{\text{max} \{g_1, 1/2\}}^{2g+g_2-1/2}),
\]

\[
\sup_{\delta_0 - \delta \leq g_1, \gamma - \delta \geq g_2} T^{\delta - \gamma_0} |M_T (\vartheta)| = O_p(T_{\text{max} \{g_1, 1/2\}}^{2g-1/2}),
\]

\[
\sup_{\delta_0 - \delta \geq g_1, \gamma - \delta \leq g_2} T^{2\delta - \delta_0 - \gamma_0} |M_T (\vartheta)| = O_p(T_{\text{max} \{g_1, 2\}}^{-g_1 - 1/2}),
\]

\[
\sup_{\delta_0 - \delta \geq g_1, \gamma - \delta \geq g_2, \gamma - \delta \geq 1/2+\theta} T^{\delta - \gamma_0} |M_T (\vartheta)| = O_p(T_{\text{max} \{g_1, 1/2\}}^{-1/2}),
\]

\[
\sup_{\delta_0 - \delta \geq g_1, \gamma - \delta \geq g_2, \gamma - \delta \geq 1/2+\theta} T^{2\delta - \delta_0 - \gamma_0} |M_T (\vartheta)| = O_p(T_{\text{max} \{g_1, 1/2\}}^{-1/2}).
\]

Lemma 7. Under Assumptions A1–A3, \( T^{-1} \sum_{t=1}^{T} s_t^2 (\vartheta_0) \rightarrow_p \sigma_0^2 \).
Lemma 8. Under Assumptions A1 and A3, for any \( g > 0 \),
\[
\lim_{T \to \infty} \inf_{\gamma_0 - \delta \geq 1/2 + g, |\gamma - \gamma_0| \geq g} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_t^2(\varphi) > \epsilon.
\]

Lemma 9. Under Assumptions A1 and A3, for \( i = 1,...,4 \),
\[
\frac{1}{T^{\gamma_0 - \delta}} \frac{\partial^2 d_T(\varphi)}{\partial \gamma^2} = \frac{\mu_0 \phi(1; \varphi) 2 (\gamma - \delta)^2 - 2 (\gamma - \delta) + 1 - (\gamma_0 - \delta)}{\Gamma(\gamma_0 - \delta) (\gamma - \delta)^2 (\gamma_0 + \gamma - 2 \delta - 1)^2} + g_T(\varphi),
\]
where \( \sup_{H_i} |g_T(\varphi)| = o(1) \), and for an arbitrarily large \( K \) (setting \( \xi \) large enough),
\[
\lim_{T \to \infty} \inf_{H_i} \frac{T^{2 \kappa_i}}{T^{2(\gamma_0 - \delta)}} d_T^2(\varphi) > K.
\]

Lemma 10. For any \( \delta > 0 \) and any fixed \( a \geq 0 \), as \( T \to \infty \),
\[
\frac{1}{T^{\delta}} \sum_{t=1}^{T} t^{d-1} \to \frac{1}{d}, \quad \frac{1}{T^{d}} \sum_{t=1}^{T} \log \left( \frac{t + a}{T} \right) t^{d-1} \to -\frac{1}{d^2}, \quad \frac{1}{T^{d}} \sum_{t=1}^{T} \log^2 \left( \frac{t + a}{T} \right) t^{d-1} \to \frac{2}{d^3}.
\]

Supplementary Material

Supplement to “Truncated sum of squares estimation of fractional time series models with deterministic trends”

The supplementary material contains a Monte Carlo simulation study, the justification of some particular steps in the proofs of the main theorems, which either use arguments similar to those already given in the proofs in the present paper or can be dealt with separately without affecting the readability of those proofs, along with proofs of Lemmas 1–10 as well as some additional technical lemmas and their proofs.

References


SUPPLEMENT TO
“TRUNCATED SUM OF SQUARES ESTIMATION OF FRACTIONAL TIME SERIES WITH DETERMINISTIC TRENDS”

by

Javier Hualde
Universidad Pública de Navarra

and

Morten Ørregaard Nielsen
Queen’s University and CREATES

Date: January 4, 2017
Without loss of generality we fix \( \gamma = 1 \) and set \( \delta_0 = 1 \) and \( \delta = \gamma_{0} - \delta_0 = -0.9 + 0.2i \) for \( i = 1, \ldots, 13 \). We computed \( \hat{\delta}, \hat{\gamma} \) 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 S.1. Here, the performance of \( \hat{\delta} \) reflects the limiting theory developed in Theorems 1 and 2. The bias of \( \hat{\delta} \) 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 \( \gamma_0 - \delta_0 \geq 0.5 \)), the bias decreases even more.

### S.1. Introduction

This supplements [4] by providing a Monte Carlo simulation study, the justification of some particular steps in the proofs of the theorems in [4], which either use arguments similar to those already given in the proofs in [4], or can be dealt with separately without affecting the readability of those proofs, along with proofs of Lemmas 1–10 as well as some additional technical lemmas and their proofs. Equation references (S.n) for \( n \geq 1 \) refer to equations in this supplement and other equation references are to the main paper, [4].

### S.2. Monte Carlo evidence

We investigate the finite sample performance of our estimators of \( \gamma_0 \) and \( \delta_0 \) by means of a simple Monte Carlo experiment. We generate the observable series \( x_t, t = 1, \ldots, T \), from (6) with \( z_t = \varepsilon_t \) being an independent \( \mathcal{N}(0, 1) \) 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 \( i = 1, \ldots, 13 \). We computed \( \hat{\delta}, \hat{\gamma} \) 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.

#### Table S.1

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

Note: Based on 10,000 Monte Carlo replications.
is higher) results worsen. As expected, the behaviour of $\hat{\gamma}$ is qualitatively different. When $\gamma_0 - \delta_0 \leq 1/2$, the bias of $\hat{\gamma}$ 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 S.2. Results for $\hat{\delta}$ are as expected, but now larger $\gamma_0 - \delta_0$ lead to slightly smaller SD for smaller sample sizes. Regarding $\hat{\gamma}$, 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 $\hat{\gamma}$ decreases with $T$ and is very small for larger $\gamma_0 - \delta_0$.

### S.3. Additional proof details for (64) and (65).

By summation by parts and $c_t (d, \varphi) - c_{t-1} (d, \varphi) = c_t (d - 1, \varphi)$, see (S.92), we find

$$
\sum_{t=1}^{T} c_{t-1} (\gamma - \delta, \varphi) \phi (L; \varphi) u_t (\delta - \delta_0) = c_{T-1} (\gamma - \delta, \varphi) \phi (L; \varphi) u_T (\delta - \delta_0 - 1)
$$

(S.1)

$$
- \sum_{t=1}^{T-1} c_t (\gamma - \delta - 1, \varphi) \phi (L; \varphi) u_t (\delta - \delta_0 - 1).
$$

Also, by summation by parts,

(S.2)  
$$
c_t (d, \varphi) = \phi (1; \varphi) \pi_t (d) + c_{2,t+1} (d, \varphi) + c_{3,t+1} (d, \varphi),
$$

\[ \sum_{t=1}^{T} c_{t-1} (\gamma - \delta, \varphi) \phi (L; \varphi) u_t (\delta - \delta_0) = c_{T-1} (\gamma - \delta, \varphi) \phi (L; \varphi) u_T (\delta - \delta_0 - 1)
\]

(S.1)  

$$
- \sum_{t=1}^{T-1} c_t (\gamma - \delta - 1, \varphi) \phi (L; \varphi) u_t (\delta - \delta_0 - 1).
$$

Also, by summation by parts,

(S.2)  
$$
c_t (d, \varphi) = \phi (1; \varphi) \pi_t (d) + c_{2,t+1} (d, \varphi) + c_{3,t+1} (d, \varphi),
$$
where, see also (S.112) in the proof of Lemma S.4, \( c_{2,t}(d, \varphi) = -\pi_{t-1}(d) \sum_{k=t}^{\infty} \phi_k(\varphi) \) and \( c_{3,t}(d, \varphi) = -\sum_{k=0}^{t-2} \pi_{k+1}(d-1) \sum_{l=0}^{k} \phi_{t-l-1}(\varphi) \). Thus,

\[
\phi(L; \varphi) u_t(\delta - \delta_0 - 1) = \sum_{j=0}^{t-1} c_j (\delta_0 - \delta + 1, \varphi) u_{t-j}
\]

\((S.3)\)

\[
= \phi(1; \varphi) \omega(1; \varphi_0) \varepsilon_t(\delta - \delta_0 - 1) + p_t(\tau),
\]

where

\[
p_t(\tau) = \phi(1; \varphi) (u_t(\delta - \delta_0 - 1) - \omega(1; \varphi_0) \varepsilon_t(\delta - \delta_0 - 1)) \]

\((S.4)\)

\[
+ \sum_{j=0}^{t-1} c_{2,j+1} (\delta_0 - \delta + 1, \varphi) u_{t-j} + \sum_{j=0}^{t-1} c_{3,j+1} (\delta_0 - \delta + 1, \varphi) u_{t-j}.
\]

Substituting \((S.1)-(S.4)\) into \( \bar{\pi}_T(\tau, \gamma_2) \), see (40), we get

\[
\frac{1}{T^{\delta_0-\delta+1/2}} \pi_T(\tau, \gamma_2) = \frac{1}{T^{\delta_0-\delta+1/2}} \phi(L; \varphi) u_T(\delta - \delta_0 - 1) - \frac{c_{T-1}(\gamma_2 - \delta, \varphi) \phi(L; \varphi) u_T(\delta - \delta_0 - 1) \sum_{t=1}^{T} c_{t-1}(\gamma_2 - \delta, \varphi)}{T^{\delta_0-\delta+1/2} \sum_{j=1}^{T} c_j(\gamma_2 - \delta, \varphi)}
\]

\[
+ \frac{\sum_{t=1}^{T} c_{t-1}(\gamma_2 - \delta, \varphi) \sum_{j=1}^{T} c_j(\gamma_2 - \delta - 1, \varphi) \phi(L; \varphi) u_j(\delta - \delta_0 - 1)}{T^{\delta_0-\delta+1/2} \sum_{j=1}^{T} c_j^2(\gamma_2 - \delta, \varphi)}
\]

\[
= \phi(1; \varphi) V_T(\delta, \gamma_2) + m_{2,T}(\tau, \gamma_2),
\]

where

\[
V_T(\delta, \gamma_2) = \frac{\omega(1; \varphi_0)}{T^{\delta_0-\delta+1/2}} \left( 1 - \frac{\pi_{T-1}(\gamma_2 - \delta) \sum_{t=1}^{T} \pi_{t-1}(\gamma_2 - \delta)}{\sum_{j=1}^{T} \pi_j^2(\gamma_2 - \delta)} \right) \varepsilon_T(\delta - \delta_0 - 1)
\]

\[
+ \frac{\omega(1; \varphi_0) \sum_{t=1}^{T} \pi_{t-1}(\gamma_2 - \delta) \sum_{j=1}^{T} \pi_j(\gamma_2 - \delta - 1) \varepsilon_j(\delta - \delta_0 - 1)}{\sum_{j=1}^{T} \pi_j^2(\gamma_2 - \delta)},
\]

and \( m_{2,T}(\tau, \gamma_2) \) collects remaining terms arising from \((S.2)\) and \((S.4)\). Then, by relatively straightforward arguments, it can be shown that

\[(S.5) \quad \sup_{G_1} |m_{2,T}(\tau, \gamma_2)| = o_p(1).\]

The proof of \((S.5)\) involves several results. First, in order to deal with the first term on the right-hand side of \((S.4)\), note that \( u_t = \omega(1; \varphi_0) \varepsilon_t + \tilde{e}_{t-1} - \tilde{e}_t, \)

where \( \tilde{e}_t = \sum_{j=0}^{\infty} \tilde{\omega}_j(\varphi_0) \varepsilon_{t-j}, \) \( \tilde{\omega}_j(\varphi_0) = \sum_{k=j+1}^{\infty} \omega_k(\varphi_0). \) By Assumptions
A1 and A2, $\tilde{\varepsilon}_t$ is well defined in the mean-square sense and $|\tilde{\omega}_j(\varphi_0)| = O(j^{-c})$. We also apply Lemma S.2, which immediately leads to the result that for any $c > 0$, $\sup_{t \geq 1/2+c} \pi_t (d) / T^d = O(t^{-1/2+c}/T^{1/2+c})$, as well as Lemmas 1, 2, and S.1.

We next show (65), where

$$V(\delta, \gamma) = \omega (1; \varphi_0) \frac{1 - (\gamma - \delta)}{\gamma - \delta} W(1; 1 + \delta_0 - \delta)$$

$$+ \omega (1; \varphi_0) \frac{2(\gamma - \delta) - 1}{\gamma - \delta} \int_0^1 r^{\gamma - \delta - 2} W(r; 1 + \delta_0 - \delta) dr,$$

and $W(r; d) = \Gamma(d)^{-1} \int_0^r (1 - s)^{d-1} dB(s)$, $B(s)$ denote fractional (type II) and regular scalar Brownian motions, respectively, both with variance $\sigma_0^2$. Convergence of the finite-dimensional distributions follows by Theorem 1 of [3] (noting that our Assumption A2 implies conditions A(i), A(ii) and A(iii) in [3]) and the continuous mapping theorem (see (A.1) in [9], 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 [6] noting, in particular, that $V_T(\delta, \gamma)$ is continuously differentiable for $\gamma - \delta \geq 1 + \alpha$, which proves (65).

**S.4. Additional proof details for (72).** Recall the intervals $I_i$ and define $\mathcal{W}_i = \{ \vartheta \in \mathcal{M}_\varepsilon : \delta \in I_i \}$ for $i = 1, 2, 3$, and $\mathcal{W}_4 = \{ \vartheta \in \mathcal{M}_\varepsilon : \delta \in \mathcal{I}_4 \cup \mathcal{I}_5 \}$. Then (72) follows on showing

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

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

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

**S.4.1. Proof of (S.6) for $i = 4$.** Given (86), we first apply the bound

$$\frac{1}{T} \sum_{t=1}^T (\phi(L; \varphi) x_t(\delta))^2 \geq \frac{1}{T} \sum_{t=1}^T (\phi(L; \varphi) u_t(\delta - \delta_0))^2$$

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

(S.8)
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.7) and (S.8), the proof of (S.6) for $i = 4$ then follows by [5] (see the proof of their (2.7) for cally stationary. In view of (S.7) and (S.8), the proof of (S.6) for $i = 4$ then follows by [5] (see the proof of their (2.7) for i.e.

(S.9) \[ \sup_{\mathcal{W}_4} \frac{1}{T} \left| \sum_{t=1}^{T} \phi (L; \varphi) u_t (\delta - \delta_0) c_{t-1} (\gamma_0 - \delta, \varphi) \right| = o_p (1), \]

(S.10) \[ \sup_{\mathcal{W}_4} \frac{1}{T} \left( \sum_{t=1}^{T} \phi (L; \varphi) x_t (\delta) h_{t-1,T} (\gamma - \delta, \varphi) \right) \leq o_p (1). \]

First, noting that $\delta \in \mathcal{I}_4 \cup \mathcal{I}_5$ implies $\delta_0 - \delta \leq 1/2 - \eta$ and $\gamma_0 - \delta \leq 1/2 + \gamma_0 - \delta_0 - \eta$, (104) of Lemma 4 implies that the left-hand side of (S.9) is $O_p (T^{\gamma_0 - \delta_0 - 1/2 - 2\eta} + T^{-\infty - \eta} + T^{-1} \log^2 T) = o_p (1)$.

Next, using (86), (S.10) follows by showing

(S.11) \[ \sup_{\mathcal{W}_4} T^{-1/2} \sum_{t=1}^{T} \phi (L; \varphi) u_t (\delta - \delta_0) h_{t-1,T} (\gamma - \delta, \varphi) = o_p (1), \]

(S.12) \[ \sup_{\mathcal{W}_4} T^{-1/2} \sum_{t=1}^{T} c_{t-1} (\gamma_0 - \delta, \varphi) h_{t-1,T} (\gamma - \delta, \varphi) = o (1). \]

Here, (106) of Lemma 5 shows that the left-hand side of (S.11) is $O_p (T^{\theta-1/2+T^{-\eta}}) = o_p (1)$ by choosing $\theta < 1/2$, while (101) of Lemma 3 shows that the left-hand side of (S.12) $O(T^{\theta-1/2+T^{\gamma_0 - \delta_0 - 1/2 - \eta}}) = o (1)$, to conclude the proof of (S.6) for $i = 4$.

S.4.2. Proof of (S.6) for $i = 3$. Noting (47), (S.7), and (S.8), the proof follows on showing

(S.13) \[ \sup_{\mathcal{W}_3} \frac{1}{T} \left| \sum_{t=1}^{T} \phi (L; \varphi) u_t (\delta - \delta_0) c_{t-1} (\gamma_0 - \delta, \varphi) \right| = O_p (1), \]

(S.14) \[ \sup_{\mathcal{W}_3} \frac{1}{T} \left( \sum_{t=1}^{T} \phi (L; \varphi) x_t (\delta) h_{t-1,T} (\gamma - \delta, \varphi) \right)^2 \leq O_p (1). \]

Both (S.13) and (S.14) follow straightforwardly by identical steps as those given in the proofs of (S.9) and (S.10) just replacing $\eta$ by 0.

S.4.3. Proof of (S.6) for $i = 2$. Clearly,

\[ \Pr \left( \inf_{\mathcal{W}_2} R_T (\vartheta) \leq \sigma_0^2 + \epsilon \right) \leq \Pr \left( \inf_{\mathcal{W}_2} \frac{T^{2(\delta_0 - \delta)}}{T} \inf_{\mathcal{W}_2} \frac{T}{T^{2(\delta_0 - \delta)}} R_T (\vartheta) \leq \sigma_0^2 + \epsilon \right) \]
Thus, in view of (55), (S.7), and (S.8), (S.6) for \( i = 2 \) follows on showing

\[
\sup_{W_2} \frac{1}{T^{2(\delta_0 - \delta)}} \left| \sum_{t=1}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) c_{t-1} (\gamma_0 - \delta, \varphi) \right| = O_p(1),
\]

(S.16)

\[
\sup_{W_2} \frac{1}{T^{2(\delta_0 - \delta)}} \left( \sum_{t=1}^{T} \phi(L; \varphi) x_t (\delta) h_{t-1,T} (\gamma - \delta, \varphi) \right)^2 = O_p(1).
\]

(S.17)

The proofs of (S.16) and (S.17) are almost identical to those of (S.9) and (S.10), taking into account the different normalization, which implies using (102) instead of (101) in Lemma 3, (105) instead of (104) in Lemma 4, and (107) instead of (106) in Lemma 5.

S.4.4. Proof of (S.6) for \( i = 1 \). Following identical steps to those given in (S.15),

\[
\Pr \left( \inf_{W_1} R_T (\vartheta) \leq \sigma_0^2 + \epsilon \right) \leq \Pr \left( \inf_{W_1} \frac{T}{T^{2(\delta_0 - \delta)}} R_T (\vartheta) \leq \frac{\sigma_0^2 + \epsilon}{T^{2 \nu}} \right).
\]

(S.18)

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

\[
\Pr \left( \inf_{W_1 \cap \Phi_j} \frac{T}{T^{2(\delta_0 - \delta)}} R_T (\vartheta) > \epsilon \right) \to 1 \text{ as } T \to \infty
\]

for \( j = 1, 2, 3 \). In the proof of their (2.7) for \( i = 1 \), [5] showed that

\[
\Pr \left( \inf_{\|\tau - \tau_0\| \geq \epsilon, \tau \in T_1} \frac{1}{T^{2(\delta_0 - \delta)}} \sum_{t=1}^{T} \left( \phi(L; \varphi) u_t (\delta - \delta_0) \right)^2 > \epsilon \right) \to 1 \text{ as } T \to \infty,
\]

so in view of (86), (S.7), and (S.8), (S.18) for \( j = 1, 2 \) holds if we show that

\[
\sup_{W_1} \frac{1}{T^{2(\delta_0 - \delta)}} \sum_{t=1}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) c_{t-1} (\gamma_0 - \delta, \varphi) = o_p(1),
\]

(S.19)

\[
\sup_{W_1} \frac{1}{T^{\nu_0 - \delta}} \sum_{t=1}^{T} c_{t-1} (\gamma_0 - \delta, \varphi) h_{t-1,T} (\gamma - \delta, \varphi) = o(1),
\]

(S.20)
First, noting that $\delta_0 - \delta \geq 1/2 + \eta$ and $\gamma_0 - \delta \geq 1/2 + \eta + \gamma_0 - \delta_0$, by (105) of Lemma 4 and (102) of Lemma 3 with $\theta < 1/2 + \eta$, the left-hand sides of (S.19) and (S.20) are $O_p(T)$ and $O(T\gamma_0 - \delta_0 - 1/2 + T^{-1/2 - \eta + \delta_0}) = o(1)$, respectively. Next, for $j = 1$, by (S.110) of Lemma S.4, the left-hand side of (S.21) is

$$\sup_{W_1 \cap \Phi_1} \frac{1}{T^{2(\delta_0 - \delta)}} \left( \sum_{t=1}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) h_{t-1,T} (\gamma - \delta, \varphi) \right)^2 = o_p(1).$$

which is easily shown to be $O_p(T - 2\alpha) = o_p(1)$ by (100) of Lemma 2 and (S.106) of Lemma S.3. For $j = 2$, we use (17) to bound the left-hand side of (S.21) by

$$\sup_{W_1 \cap \Phi_2} T^{2(\delta - \delta_0)} \left( \sum_{t=1}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) c_{t-1} (\gamma - \delta, \varphi) \right)^2,$$

where the denominator can be made arbitrarily large by (S.111) of Lemma S.4 and the numerator is easily seen to be $O_p(1)$ by direct application of (100) of Lemma 2 and (S.107) of Lemma S.3.

We finally show (S.18) for $j = 3$. Using (29), applying (S.19) and (S.20) together with Lemmas 1, 2, we have

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

where $\sup_{W_1 \cap \Phi_3} |q_{1,T}(\vartheta)| = o_p(1)$. Thus, (S.18) for $j = 3$ holds on showing

$$\Pr \left( \inf_{W_1 \cap \Phi_3} \frac{1}{T^{2(\delta_0 - \delta)}} \sum_{t=1}^{T} s_t^2(\vartheta) > \epsilon \right) \to 1 \text{ as } T \to \infty.$$
where $\sup_{W \cap \Phi_3} |q_{2,T}(\tau, \gamma)| = o_p(1)$. To prove (S.23) we first justify that

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

(S.24)

$$
\geq \frac{\left( \sum_{t=1}^{T-2} (2\phi(L; \varphi) u_t (\delta - \delta_0) - \phi(L; \varphi) u_{T-1} (\delta - \delta_0 - 1)) c_t (\gamma - \vartheta, \varphi) + q_{3,T}(\vartheta) \right)^2}{T^2 \sum_{t=1}^{T} c_t^2 (\gamma - \vartheta, \varphi)},
$$

where $\sup_{W \cap \Phi_3} T^{2\delta - \delta_0 - \gamma - 1/2} |q_{3,T}(\vartheta)| = o_p(1)$. To see this, first note that the equality in (S.24) follows because $\sum_{t=1}^{T} h_{t-1,T}(d, \varphi) = 1$ and then note that left-hand side of the inequality in (S.24) is

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

(S.25)

$$
= \sum_{t=1}^{T-1} \sum_{k=t+1}^{T} (\phi(L; \varphi) u_t (\delta - \delta_0) h_{k-1,T} (\gamma - \vartheta, \varphi) - \phi(L; \varphi) u_k (\delta - \delta_0) h_{t-1,T} (\gamma - \vartheta, \varphi))^2
$$

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.25) is

$$
\sum_{t=1}^{T-1} \sum_{k=t+1}^{T} h_{k-1,T} (\gamma - \vartheta, \varphi) \phi(L; \varphi) u_t (\delta - \delta_0) = h_{T-1,T} (\gamma - \vartheta, \varphi) \phi(L; \varphi) u_{T-1} (\delta - \delta_0 - 1)
$$

(S.26)

$$
+ \sum_{t=1}^{T-2} \phi(L; \varphi) u_t (\delta - \delta_0 - 1) h_{t,T} (\gamma - \vartheta, \varphi).
$$

The contribution of the second term in the parenthesis on the right-hand side of (S.25) is

$$
\sum_{t=1}^{T-1} \sum_{k=t+1}^{T} \phi(L; \varphi) u_k (\delta - \delta_0) h_{t-1,T} (\gamma - \vartheta, \varphi)
$$

$$
= \sum_{t=2}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) \sum_{k=1}^{t-1} h_{k-1,T} (\gamma - \vartheta, \varphi)
$$
\[
\begin{align*}
&= \sum_{t=2}^{T} \phi(L; \varphi) u_t (\delta - \delta_0 + 1) \sum_{k=1}^{t-1} h_{k-1,T} (\gamma - \delta, \varphi) \\
&\quad + \sum_{t=2}^{T} \phi(L; \varphi) u_{t-1} (\delta - \delta_0) \sum_{k=1}^{t-1} h_{k-1,T} (\gamma - \delta, \varphi), \\
&\text{(S.27)}
\end{align*}
\]

and by summation by parts the second term on the right-hand side of (S.27) is

\[
\phi(L; \varphi) u_{t-1} (\delta - \delta_0 - 1) \sum_{k=1}^{T-1} h_{k-1,T} (\gamma - \delta, \varphi)
\]

\[
- \sum_{t=2}^{T-1} \phi(L; \varphi) u_{t-1} (\delta - \delta_0 - 1) h_{t-1,T} (\gamma - \delta, \varphi).
\]

Thus, in view of (S.25)–(S.28), (17), and noting that \(h_{0,T} (\gamma - \delta, \varphi) = 1\) and \(T(T - 1)/2 \leq T^2\), (S.24) holds with

\[
q_{3,T} (\vartheta) = c_{T-1} (\gamma - \delta, \varphi) \phi(L; \varphi) u_{T-1} (\delta - \delta_0 - 1)
\]

\[
- \sum_{t=2}^{T} \phi(L; \varphi) u_{t} (\delta - \delta_0 + 1) \sum_{k=1}^{t-1} c_{k-1} (\gamma - \delta, \varphi) - \phi(L; \varphi) u_{T-1} (\delta - \delta_0 - 1),
\]

and consequently \(\sup_{W_1 \cap \Phi_3} T^{2\delta - \delta_0 - \gamma - 1/2} |q_{3,T} (\vartheta)| = O_p(T^{-1} + T^{-\alpha - 1/2}) = o_p(1)\) by application of (100) of Lemma 2 and (S.107) of Lemma S.3. Next, by (S.2), (S.3), and (S.4),

\[
\begin{align*}
&\sum_{t=1}^{T-2} \left(2\phi(L; \varphi) u_{t} (\delta - \delta_0 - 1) - \phi(L; \varphi) u_{T-1} (\delta - \delta_0 - 1)\right) c_t (\gamma - \delta, \varphi) \\
&\text{(S.29)}
\end{align*}
\]

\[
= \phi^2 (1; \varphi) \omega (1; \varphi_0) \sum_{t=1}^{T-2} (2\varepsilon_t (\delta - \delta_0 - 1) - \varepsilon_{T-1} (\delta - \delta_0 - 1)) \pi_t (\gamma - \delta) + q_{4,T} (\vartheta),
\]

where

\[
q_{4,T} (\vartheta) = \sum_{t=1}^{T-2} \left(2\phi(L; \varphi) u_{t} (\delta - \delta_0 - 1) - \phi(L; \varphi) u_{T-1} (\delta - \delta_0 - 1)\right)
\]

\[
\times \left(c_{2,t+1} (\gamma - \delta, \varphi) + c_{3,t+1} (\gamma - \delta, \varphi)\right)
\]

\[
+ \phi(1; \varphi) \sum_{t=1}^{T-2} \left(2p_t (\vartheta) - p_{T-1} (\vartheta)\right) \pi_t (\gamma - \delta)
\]
and satisfies \( \sup_{W_1 \cap \Phi_3} T^{2\delta - \delta_0 - \gamma - 1/2} |q_{4,T}(\theta)| = o_p(1) \) by the same arguments as for (S.5).

By (S.24), (S.29), and (S.109) in Lemma S.4, we thus have the bound (S.23), where \( \sup_{W_1 \cap \Phi_3} |q_{2,T}(\theta)| = o_p(1) \) and

\[
Z_T(\delta, \gamma) = \frac{1}{T^{\delta_0 - \delta + 1}} \sum_{t=1}^{T-2} (2\varepsilon_t (\delta - \delta_0 - 1) - \varepsilon_{T-1} (\delta - \delta_0 - 1)) \pi_t (\gamma - \delta) \quad \frac{\pi^2_t}{(\sum_{t=1}^{T} \pi^2_{t-1} (\gamma - \delta))^{1/2}}.
\]

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

(S.30)  
\[
Z_T(\delta, \gamma) \Rightarrow Z(\delta, \gamma)
\]

follows by exactly the same arguments as those in the proof of (65), where

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

Then, noting Assumption A1(iv), (8) and the definition of \( L \) given in the proof of (21) for \( i = 1, \) (S.22) follows on showing that

(S.31)  
\[
\Pr \left( \inf_{\ell} Z^2_T(\delta, \gamma) > \epsilon \right) \rightarrow 1 \quad \text{as} \quad T \rightarrow \infty.
\]

By (S.30) and the continuous mapping theorem, \( \inf_{\ell} Z^2_T(\delta, \gamma) \rightarrow_d \inf_{\ell} Z^2(\delta, \gamma) \) as \( T \rightarrow \infty, \) so that \( \Pr \left( \inf_{\ell} Z^2_T(\delta, \gamma) > \epsilon \right) \rightarrow \Pr \left( \inf_{\ell} Z^2(\delta, \gamma) > \epsilon \right) \) as \( T \rightarrow \infty, \) and (S.31) follows because \( \epsilon \) is arbitrarily small and \( Z^2(\delta, \gamma) \) is the square of a Gaussian random variable. This completes the proof of (S.22) and therefore that of (72).

**S.5. Additional proof details for (75).** First we show that

(S.32)  
\[
\frac{T^{1/2}}{2} M_T \frac{\partial R_1(\theta_0)}{\partial \theta} = \left( \begin{array}{cc} T^{-1/2} I_{p+1} & 0 \\ 0 & T^{1/2 - (\gamma_0 - \delta_0)} \end{array} \right) \sum_{t=1}^{T} \varepsilon_t \left( \frac{\partial s_{1t}(\theta_0)}{\partial \theta} \right) + o_p(1),
\]

for which it is sufficient to prove that the first-order derivatives satisfy

(S.33)  
\[
\frac{1}{T^{1/2}} \sum_{t=1}^{T} (s_t(\theta_0) - \varepsilon_t) \frac{\partial s_{1t}(\theta_0)}{\partial \tau} = o_p(1), \quad \frac{1}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=1}^{T} (s_t(\theta_0) - \varepsilon_t) \frac{\partial d_t(\theta_0)}{\partial \gamma} = o_p(1),
\]

(S.34)  
\[
\frac{1}{T^{1/2}} \sum_{t=1}^{T} s_t(\theta_0) \frac{\partial s_{2t}(\theta_0)}{\partial \tau} = o_p(1), \quad \frac{1}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=1}^{T} s_t(\theta_0) \frac{\partial s_{2t}(\theta_0)}{\partial \gamma} = o_p(1).
\]
First we show the first equality in (S.33). By definition of \( s_t (\vartheta) \),
(S.35)\[ s_t (\vartheta) = - \sum_{j=t}^{\infty} \phi_j (\varphi_0) u_{t-j} - h_{t-1,T} (\gamma_0 - \delta_0, \varphi_0) \sum_{j=1}^{T} s_{1j} (\tau_0) h_{j-1,T} (\gamma_0 - \delta_0, \varphi_0), \]
where \( s_{1j} (\tau_0) = \varepsilon_j - \sum_{k=j}^{\infty} \phi_k (\varphi_0) u_{j-k} \), so the result holds if
(S.36)\[ \frac{1}{T^{1/2}} \sum_{t=1}^{T} \sum_{j=1}^{\infty} \phi_j (\varphi_0) u_{t-j} \frac{\partial s_{1t} (\tau_0)}{\partial \tau} = o_p (1), \]
(S.37)\[ \frac{1}{T^{1/2}} \sum_{t=1}^{T} h_{t-1,T} (\gamma_0 - \delta_0, \varphi_0) \frac{\partial s_{1t} (\tau_0)}{\partial \tau} \sum_{j=1}^{T} s_{1j} (\tau_0) h_{j-1,T} (\gamma_0 - \delta_0, \varphi_0) = o_p (1). \]
First, for \( t \geq 2 \),
(S.38)\[ \frac{\partial s_{1t} (\tau_0)}{\partial \delta} = - \sum_{j=1}^{t-1} \sum_{k=0}^{j-1} \phi_k (\varphi_0) (j-k)^{-1} u_{t-j}, \]
so, noting (7) and applying Lemma S.1,
\[ E \left\| \frac{\partial s_{1t} (\tau_0)}{\partial \delta} \right\| \leq K \sum_{j=1}^{t-1} \sum_{k=1}^{j-1} k^{-1} (j-k)^{-1} \leq K \sum_{j=1}^{t-1} j^{-1} \log j \leq K \log^2 t. \]
Similarly, by (12),
(S.39)\[ E \left\| \frac{\partial s_{1t} (\tau_0)}{\partial \varphi} \right\| = E \left\| \sum_{j=1}^{t-1} \frac{\partial \phi_j (\varphi_0)}{\partial \varphi} u_{t-j} \right\| = O (1), \]
so that
\[ E \left\| \frac{\partial s_{1t} (\tau_0)}{\partial \tau} \right\| = O (\log^2 t). \]
Thus, noting (S.95),
\[ E \left\| \frac{1}{T^{1/2}} \sum_{t=1}^{T} \sum_{j=t}^{\infty} \phi_j (\varphi_0) u_{t-j} \frac{\partial s_{1t} (\tau_0)}{\partial \tau} \right\| \leq \frac{K}{T^{1/2}} \sum_{t=1}^{T} t^{-1/2} \log^2 t \leq KT^{-1/2} = o (1) \]
because $\varsigma > 1/2$, which proves (S.36). Next, by (94) of Lemma 1 and (S.95), it is straightforward to show that

$$\sum_{j=1}^{T} s_{1j} (\tau_0) h_{j-1,T} (\gamma_0 - \delta_0, \varphi_0)$$

(S.40)  

$$= \sum_{j=1}^{T} (\varepsilon_j - \sum_{k=j}^{\infty} \phi_k (\varphi_0) u_{j-k}) h_{j-1,T} (\gamma_0 - \delta_0, \varphi_0) = O_p (T^\theta),$$

so (S.37) holds on showing that

$$\sum_{t=1}^{T} \frac{1}{T^{1/2}} \sum_{j=1}^{T} h_{t-j, T} (\gamma_0 - \delta_0, \varphi_0) \frac{\partial s_{1t} (\tau_0)}{\partial \tau} = O_p \left( T^{\theta - 1/2} \log T \right)$$

(S.41)  

and setting $\theta < 1/4$. Noting (S.38), by simple calculations,

$$\frac{\partial s_{1t} (\tau_0)}{\partial \delta} = - \sum_{j=1}^{t-1} \frac{1}{j} \varepsilon_{t-j} + \sum_{j=1}^{t-1} \frac{1}{j} \sum_{k=t-j}^{\infty} \phi_k (\varphi_0) u_{t-j-k}. $$

(S.42)  

The contribution of the first term on the right-hand side of (S.42) to the left-hand side of (S.41) is, by (94) of Lemma 1,

$$- \frac{1}{T^{1/2}} \sum_{t=2}^{T} \sum_{j=1}^{t-1} \varepsilon_{t-j} h_{t-1,T} (\gamma_0 - \delta_0, \varphi_0) = - \frac{1}{T^{1/2}} \sum_{t=1}^{T-1} \varepsilon_t \sum_{j=1}^{t-1} h_{t+j-1,T} (\gamma_0 - \delta_0, \varphi_0)$$

$$= O_p \left( \frac{T^\theta}{T^{1/2}} \left( \sum_{t=1}^{T-1} \left( \sum_{j=1}^{t-1} \frac{1}{j} (t + j)^{-1/2 - \theta} \right)^2 \right)^{1/2} \right)$$

(S.43)  

$$= O_p (T^{\theta - 1/2} \log T)$$

because

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

Similarly, the contribution of the second term on the right-hand side of (S.42) to the left-hand side of (S.41) is

$$\frac{1}{T^{1/2}} \sum_{t=1}^{T} h_{t-1,T} (\gamma_0 - \delta_0, \varphi_0) \sum_{j=1}^{t-1} \sum_{k=t-j}^{\infty} \phi_k (\varphi_0) u_{t-j-k},$$
which can be easily shown to be \( O_p \left( T^{\theta - 1/2} \log T \right) \) by (94) of Lemma 1, (S.95), and Lemma S.1. Next, the derivative with respect to \( \varphi \) on the left-hand side of (S.41) is

\[
\frac{1}{T^{1/2}} \sum_{t=1}^{T-1} u_t \sum_{j=1}^{T-1} \frac{\partial \phi_j(\varphi_0)}{\partial \varphi} h_{t+j-1,T} (\gamma_0 - \delta_0, \varphi_0),
\]

which, by very similar arguments to (S.43), can easily be shown to be \( O_p \left( T^{\theta - 1/2} \right) \) by (12) and (94) of Lemma 1, to conclude the proof of (S.41) and hence of the first equality in (S.33).

Next, because

\[
\sum_{t=1}^{T} c_{t-1} (\gamma_0 - \delta_0, \varphi_0) \partial d_t (\vartheta_0) / \partial \gamma = 0,
\]

the proof of the second equality in (S.33) follows by showing that

\[
\frac{1}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=1}^{T} \sum_{j=t}^{\infty} \phi_j(\varphi_0) u_{t-j} \frac{\partial d_t (\vartheta_0)}{\partial \gamma} = o_p (1),
\]

which, noting the proof of (S.49) and (S.95), follows easily by previous arguments.

The proofs of the two equalities in (S.34) are almost identical, but the second is simpler, so we show only the first. By (S.35), the first equality in (S.34) holds if

\[
1 \sum_{t=1}^{T} s_{1t} (\tau_0) \frac{\partial s_{2t} (\vartheta_0)}{\partial \tau} = o_p (1), \tag{S.44}
\]

\[
\frac{1}{T^{1/2}} \sum_{t=1}^{T} h_{t-1,T} (\gamma_0 - \delta_0, \varphi_0) \frac{\partial s_{2t} (\vartheta_0)}{\partial \tau} \sum_{j=1}^{T} s_{1j} (\tau_0) h_{j-1,T} (\gamma_0 - \delta_0, \varphi_0) = o_p (1). \tag{S.45}
\]

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

\[
\frac{\partial s_{2t}(\vartheta)}{\partial \tau} = \frac{\partial h_{t-1,T}(\gamma - \delta, \varphi)}{\partial \tau} \sum_{j=1}^{T} s_{1j}(\tau) h_{j-1,T}(\gamma - \delta, \varphi)
\]

\[
+ h_{t-1,T}(\gamma - \delta, \varphi) \sum_{j=1}^{T} \frac{\partial s_{1j}(\tau)}{\partial \tau} h_{j-1,T}(\gamma - \delta, \varphi)
\]

\[
+ h_{t-1,T}(\gamma - \delta, \varphi) \sum_{j=1}^{T} s_{1j}(\tau) \frac{\partial h_{j-1,T}(\gamma - \delta, \varphi)}{\partial \tau}. \tag{S.46}
\]
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 (95) of Lemma 1,

$$\left\| \partial h_{t,T} (\gamma_0 - \delta_0, \varphi_0) \right\| = O \left( t^{-1/2} \left( \frac{T}{t} \right)^{\theta} \log T \right).$$

Then noting (S.40), (S.41), and by application of Lemma 1, it follows that

$$\partial s_{2t} (\varphi_0) = O_p(t^{-1/2 - \theta} T^{2\theta} \log T).$$

By (S.48) and (S.95), it follows that the left-hand side of (S.44) is $O_p(T^{2\theta - 1/2} \log T) = o_p(1)$ by setting $\theta < 1/4$. Similarly, by (94) of Lemma 1, (S.40), and (S.48), the left-hand side of (S.45) 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 (S.34) and hence that of (S.32).

Next, as in (2.54) of [5],

$$\frac{1}{T^{1/2}} \sum_{t=1}^{T} \varepsilon_t \frac{\partial s_{1t} (\tau_0)}{\partial \tau} = \frac{1}{T^{1/2}} \sum_{t=2}^{T} \varepsilon_t \sum_{j=1}^{\infty} m_j (\varphi_0) \varepsilon_{t-j} + o_p(1),$$

where $m_j (\varphi_0) = (-j^{-1}, b'_j (\varphi_0))'$. Also,

$$\frac{\partial d_t (\varphi_0)}{\partial \gamma} = -\mu_0 c^{(1)}_{t-1} (\gamma_0 - \delta_0, \varphi_0)$$

$$+ \mu_0 c_{t-1} (\gamma_0 - \delta_0, \varphi_0) \frac{\sum_{j=1}^{T} c_{j-1} (\gamma_0 - \delta_0, \varphi_0) c^{(1)}_{j-1} (\gamma_0 - \delta_0, \varphi_0)}{\sum_{j=1}^{T} c_j^2 (\gamma_0 - \delta_0, \varphi_0)},$$

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

$$\frac{1}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=1}^{T} \varepsilon_t \frac{\partial d_t (\varphi_0)}{\partial \gamma}$$

$$= \frac{\mu_0}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=2}^{T} \varepsilon_t c_{t-1} (\gamma_0 - \delta_0, \varphi_0) \frac{\sum_{j=1}^{T} c_{j-1} (\gamma_0 - \delta_0, \varphi_0) c^{(1)}_{j-1} (\gamma_0 - \delta_0, \varphi_0)}{\sum_{j=1}^{T} c_j^2 (\gamma_0 - \delta_0, \varphi_0)}$$

$$- \frac{\mu_0}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=2}^{T} \varepsilon_t c^{(1)}_{t-1} (\gamma_0 - \delta_0, \varphi_0).$$
By (10),
\[ c^{(1)}_{d-1}(d, \varphi) = \sum_{j=0}^{t-1} \phi_j(\varphi) \pi^{(1)}_{t-j-1}(d), \]
where \( \pi^{(1)}_j(\cdot) \) is the first derivative of \( \pi_j(\cdot) \) and

\[ \pi^{(1)}_j(d) = (\psi(d + j) - \psi(d)) \pi_j(d), \]

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 S.4, the right-hand side of (S.49) equals

\[ \frac{\mu_0 \phi(1; \varphi_0)}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=2}^{T} \varepsilon_t \pi_{t-1}(\gamma_0 - \delta_0) \sum_{j=1}^{T} \pi_{j-1}(\gamma_0 - \delta_0) \pi^{(1)}_{j-1}(\gamma_0 - \delta_0) \frac{\sum_{j=1}^{T} \pi_{j-1}^2(\gamma_0 - \delta_0)}{\sum_{j=1}^{T} \pi_{j-1}^2(\gamma_0 - \delta_0)} \]

(S.50)

\[ \frac{\mu_0 \phi(1; \varphi_0)}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=2}^{T} \varepsilon_t \pi_{t-1}(\gamma_0 - \delta_0) + o_p(1). \]

(S.51)

Substituting (S.50) (evaluated at \( \gamma_0 - \delta_0 \)) into (S.51), the first two terms of (S.51) become

\[ \frac{\mu_0 \phi(1; \varphi_0)}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=2}^{T} \varepsilon_t \pi_{t-1}(\gamma_0 - \delta_0) \sum_{j=1}^{T} \psi(\gamma_0 - \delta_0 + j - 1) \pi_{j-1}^2(\gamma_0 - \delta_0) \frac{\sum_{j=1}^{T} \pi_{j-1}^2(\gamma_0 - \delta_0)}{\sum_{j=1}^{T} \pi_{j-1}^2(\gamma_0 - \delta_0)} \]

(S.52)

\[ \frac{\mu_0 \phi(1; \varphi_0)}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=2}^{T} \varepsilon_t \psi(\gamma_0 - \delta_0 + t - 1) \pi_{t-1}(\gamma_0 - \delta_0). \]

By the properties of the digamma function it can be shown that (S.52) is

\[ \frac{\mu_0 \phi(1; \varphi_0)}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=2}^{T} \varepsilon_t \pi_{t-1}(\gamma_0 - \delta_0) \sum_{j=1}^{T} \log(\gamma_0 - \delta_0 + j - 1) \pi_{j-1}^2(\gamma_0 - \delta_0) \frac{\sum_{j=1}^{T} \pi_{j-1}^2(\gamma_0 - \delta_0)}{\sum_{j=1}^{T} \pi_{j-1}^2(\gamma_0 - \delta_0)} \]

(S.53)

\[ \frac{\mu_0 \phi(1; \varphi_0)}{T^{\gamma_0 - \delta_0 - 1/2}} \sum_{t=2}^{T} \varepsilon_t \log(\gamma_0 - \delta_0 + t - 1) \pi_{t-1}(\gamma_0 - \delta_0) + o_p(1), \]

where the first two terms in (S.53) equal \( \mu_0 \phi(1; \varphi_0) T^{1/2+\delta_0-\gamma_0} \sum_{t=2}^{T} \varepsilon_t g_{t,T} (\gamma_0 - \delta_0) \)
with

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

Collecting all these terms shows (75).
S.6. Additional proof details for (82). As in the proof of Theorem 1(i), noting (15), (16), (18), (68), the result holds on establishing that

\[
\Pr \left( \inf_{\vartheta \in \hat{M}_T} S_T (\vartheta) \leq 0 \right) \rightarrow 0 \text{ as } T \rightarrow \infty,
\]

(S.54)

\[
\Pr \left( \vartheta \in \hat{N}_T \cap \hat{M}_T^s, \inf_{\hat{N}_T \cap \hat{M}_T} R_T (\vartheta) \leq 0 \right) \rightarrow 0 \text{ as } T \rightarrow \infty,
\]

(S.55)

where

\[
M_T^s = \{ \vartheta \in \Xi : \| \tau - \tau_0 \| < \epsilon \}, \quad \bar{M}_T^s = \{ \vartheta \in \Xi : \epsilon T^{-\kappa} \leq \| \tau - \tau_0 \| < \epsilon \},
\]

\[
N_T^s = \{ \vartheta \in \Xi : |\gamma - \gamma_0| < \epsilon \}, \quad \bar{N}_T^s = \{ \vartheta \in \Xi : \epsilon T^{-\kappa} \leq |\gamma - \gamma_0| < \epsilon \}.
\]

We first prove (S.54), which, defining \( J_i = \{ \vartheta \in \hat{M}_T^s : \delta \in I_i \} \), holds if

\[
\Pr \left( \inf_{J_i} S_T (\vartheta) \leq 0 \right) \rightarrow 0 \text{ as } T \rightarrow \infty
\]

(S.56)

for \( i = 4, 5 \). Note here that \( \vartheta \in \bar{M}_T^s \) implies \( \| \tau - \tau_0 \| < \epsilon \), so necessarily \( \delta \in I_4 \cup I_5 \) and there is no need to consider the intervals \( I_1, I_2, I_3 \). Clearly, (S.56) for \( i = 5 \) would hold if

\[
\Pr \left( \inf_{J_5} T^{2\kappa} S_T (\vartheta) \leq 0 \right) \rightarrow 0 \text{ as } T \rightarrow \infty.
\]

(S.57)

Proceeding as in the proof of (19)–(21) for \( i = 5 \), (S.57) holds if

\[
\inf_{J_5} T^{2\kappa} U (\tau) > \epsilon,
\]

(S.58)

\[
\frac{1}{T^{1-2\kappa}} \sum_{t=1}^{T} \left[ (\phi(L; \varphi_0) \{ u_t(t > 0) \} )^2 - \sigma_0^2 \right] = o_p (1),
\]

(S.59)

\[
\sup_{J_5} \frac{1}{T^{1-2\kappa}} \sum_{t=1}^{T} \left[ (\phi(L; \varphi) u_t (\delta - \delta_0))^2 - E \left( (\phi(L; \varphi) \Delta^{\delta-\delta_0} u_t)^2 \right) \right] = o_p (1),
\]

(S.60)

\[
\sup_{J_5} \frac{1}{T^{1-2\kappa}} \left| \sum_{t=1}^{T} d_t (\vartheta) s_t (\vartheta) \right| = o_p (1),
\]

(S.61)
that the third derivative of
\[
\left( \sum_{t=1}^{T} \phi(L; \varphi) u_t (\delta - \delta_0) h_{t-1,T}(\gamma - \delta, \varphi) \right)^2 = o_p(1).
\]

First, we justify (S.58). Clearly
\[
U(\tau) = \sigma_0^2 \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{\phi(e^{i\lambda}; \varphi)}{\phi(e^{i\lambda}; \varphi_0)} \right|^2 \left| 1 - e^{i\lambda(\delta - \delta_0)} \right| d\lambda - 1 \right),
\]
and we show that \( U(\tau) \) is a strictly convex function at \( \tau_0 \) with a strict local minimum at \( \tau = \tau_0 \). Noting that
\[
\int_{-\pi}^{\pi} e^{iq\lambda} d\lambda = 0 \text{ for any } q = \pm 1, \pm 2, \ldots,
\]
and \( \int_{-\pi}^{\pi} \log (2 - 2 \cos \lambda) d\lambda = 0 \), it is straightforward to show that \( \partial U(\tau_0)/\partial \tau = 0 \). Similarly, using again (S.63),
\[
\frac{\partial^2 U(\tau_0)}{\partial \tau \partial \tau'} = \begin{pmatrix}
\int_{-\pi}^{\pi} \log^2 (2 - 2 \cos \lambda) d\lambda & \int_{-\pi}^{\pi} \frac{\delta\phi(e^{i\lambda}; \varphi_0)/\delta\varphi'}{\phi(e^{i\lambda}; \varphi_0)} \log (2 - 2 \cos \lambda) d\lambda \\
\int_{-\pi}^{\pi} \frac{\delta\phi(e^{i\lambda}; \varphi_0)/\delta\varphi'}{\phi(e^{i\lambda}; \varphi_0)} \log (2 - 2 \cos \lambda) d\lambda & \int_{-\pi}^{\pi} \frac{\delta^2\phi(e^{i\lambda}; \varphi_0)/\delta\varphi^2}{\phi(e^{i\lambda}; \varphi_0)} \log (2 - 2 \cos \lambda) d\lambda
\end{pmatrix}
\]
\[
= \begin{pmatrix}
\frac{2\pi^3}{3} & -4\pi \sum_{j=1}^{\infty} b_j'(\varphi_0)/j \\
-4\pi \sum_{j=1}^{\infty} b_j(\varphi_0) b_j'(\varphi_0) & \frac{4\pi}{j} \sum_{j=1}^{\infty} b_j(\varphi_0) b_j'(\varphi_0)
\end{pmatrix},
\]
which by A4(iii) is positive definite, to complete the proof of strict convexity of \( U(\tau) \) at \( \tau_0 \). Thus, by continuity there exists a point \( \tau^* \) such that \( \|\tau_0 - \tau^*\| = \varepsilon T^{-1}\) and \( \inf_{\tau} U(\tau) = U(\tau^*) \). Then, noting that \( U(\tau_0) = 0 \) and \( \partial U(\tau_0)/\partial \tau = 0 \), by Taylor’s expansion,
\[
U(\tau^*) \geq \frac{1}{2} (\tau^* - \tau_0)' \frac{\partial^2 U(\tau_0)}{\partial \tau \partial \tau'} (\tau^* - \tau_0) - |w_T|,
\]
where it can be shown that \( w_T = O(T^{-3\alpha}) \). Here, the main issue is to justify that the third derivative of \( U(\tau) \) evaluated at an arbitrarily small neighborhood of \( \tau_0 \) is bounded, but this follows straightforwardly from A4(ii). Additionally,
\[
(\tau^* - \tau_0)' \frac{\partial^2 U(\tau_0)}{\partial \tau \partial \tau'} (\tau^* - \tau_0) \geq \lambda \|\tau^* - \tau_0\|^2,
\]
where \( \lambda \) denotes the minimum eigenvalue of the matrix \( \partial^2 U(\tau_0)/\partial \tau \partial \tau' \), so by (S.64) and noting that \( \lambda \) is strictly positive, for a sufficiently small \( \varepsilon > 0 \),
\[
U(\tau^*) > \frac{\varepsilon}{\varepsilon^2} \|\tau^* - \tau_0\|^2,
\]
which justifies (S.58). The proofs of (S.59)–(S.62) are omitted as, for small enough \( \kappa \), they follow by almost identical arguments to those of (24)–(27).

Next, the proof of (S.56) for \( i = 4 \) is omitted because it is basically identical to those of (19)–(21) for \( i = 4 \). The only difference is that now \( \varepsilon T^{-\kappa} \leq \| \tau - \tau_0 \| \leq \varepsilon \) instead of \( \| \tau - \tau_0 \| \geq \varepsilon \), but this does not make any difference. This completes the justification of (S.54).

Finally, we prove (S.55). For the same reason as in the proof of (68), we need to prove that

\[
\inf_{\varphi \in \mathcal{N}_r \cap M_2^*} \frac{T^{2\kappa}}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_t^2 (\varphi) > \varepsilon, \tag{S.65}
\]

\[
\sup_{\varphi \in \mathcal{N}_r \cap M_2^*} \frac{T^{2\kappa}}{T^{2(\gamma_0 - \delta) - 1}} \left| \sum_{t=1}^{T} d_t (\varphi) s_{1t} (\tau) \right| = o_p(1), \tag{S.66}
\]

\[
\sup_{\varphi \in \mathcal{N}_r \cap M_2^*} \frac{T^{2\kappa}}{T^{2(\gamma_0 - \delta) - 1}} \left( \sum_{j=1}^{T} s_{1j} (\tau) h_{j-1,T} (\gamma - \delta, \varphi) \right)^2 = o_p(1). \tag{S.67}
\]

As in (34), the proof of (S.65) follows by Lemma 9, whereas the proofs of (S.66) and (S.67) hold as in (70) and (71) for \( \kappa > 0 \) sufficiently small.

S.7. Additional proof details for (83). Recalling \( d_t (\varphi_0) = 0 \) and noting \( \partial^2 d_t (\varphi_0) / \partial \tau \partial \tau' = 0 \), it is sufficient to prove that the second-order and cross-derivatives satisfy

\[
\frac{1}{T} \sum_{t=1}^{T} \frac{\partial s_{1t} (\tau_0) \partial s_{2t} (\varphi_0)}{\partial \tau \partial \tau'} = o_p(1), \quad \frac{1}{T^{\gamma_0 - \delta_0}} \sum_{t=1}^{T} \frac{\partial s_{1t} (\tau_0) \partial s_{2t} (\varphi_0)}{\partial \gamma} = o_p(1),
\]

\[
\frac{1}{T^{\gamma_0 - \delta_0}} \sum_{t=1}^{T} \frac{\partial s_{2t} (\varphi_0) \partial s_{2t} (\varphi_0)}{\partial \gamma} = o_p(1), \quad \frac{1}{T^{\gamma_0 - \delta_0}} \sum_{t=1}^{T} \frac{\partial s_{2t} (\varphi_0) \partial d_t (\varphi_0)}{\partial \gamma} = o_p(1),
\]

\[
\frac{1}{T^{2(\gamma_0 - \delta_0) - 1}} \sum_{t=1}^{T} \left( \frac{\partial s_{2t} (\varphi_0)}{\partial \gamma} \right)^2 = o_p(1), \quad \frac{1}{T^{2(\gamma_0 - \delta_0) - 1}} \sum_{t=1}^{T} \frac{\partial s_{2t} (\varphi_0) \partial d_t (\varphi_0)}{\partial \gamma} = o_p(1),
\]

\[
\frac{1}{T} \sum_{t=1}^{T} s_t (\varphi_0) \frac{\partial^2 s_{2t} (\varphi_0)}{\partial \tau \partial \tau'} = o_p(1), \quad \frac{1}{T^{\gamma_0 - \delta_0}} \sum_{t=1}^{T} s_t (\varphi_0) \frac{\partial^2 s_{2t} (\varphi_0)}{\partial \tau \partial \gamma} = o_p(1),
\]

\[
\frac{1}{T^{2(\gamma_0 - \delta_0) - 1}} \sum_{t=1}^{T} s_t (\varphi_0) \frac{\partial^2 s_{2t} (\varphi_0)}{\partial \gamma^2} = o_p(1), \quad \frac{1}{T^{\gamma_0 - \delta_0}} \sum_{t=1}^{T} s_t (\varphi_0) \frac{\partial^2 d_t (\varphi_0)}{\partial \tau \partial \gamma} = o_p(1).
\[
\frac{1}{T^{2(\gamma_0-\delta_0)-1}} \sum_{t=1}^{T} s_t(\vartheta_0) \frac{\partial^2 d_t(\vartheta_0)}{\partial \gamma^2} = o_p(1).
\]

The proofs of these results are heavily based on the arguments employed in the proofs of (S.33) and (S.34), and are therefore omitted.

**S.8. Additional proof details for (87) and (88).** Clearly (87) follows if

(S.68)
\[
\sum_{t=1}^{T} c_{t-1}(\gamma_0 - \delta, \varphi) k_{t-1,T}(\gamma - \delta, \varphi) - \sum_{t=1}^{T} c_{t-1}(\gamma_0 - \delta, \varphi_0) k_{t-1,T}(\gamma_0 - \delta, \varphi_0) = o_p(\log T),
\]

(S.69)
\[
\frac{1}{\log T} \sum_{t=1}^{T} c_{t-1}(\gamma_0 - \delta, \varphi_0) k_{t-1,T}(\gamma_0 - \delta, \varphi_0) \rightarrow -1.
\]

First, (S.68) can be easily justified by applying Lemmas S.3 and S.5, noting that

\[
k_{t-1,T}(\gamma - \delta, \varphi) = \frac{c_{t-1}(\gamma - \delta, \varphi)}{\sum_{t=1}^{T} c_{t-1}^2(\gamma - \delta, \varphi)} - \frac{2c_{t-1}(\gamma - \delta, \varphi) \sum_{t=1}^{T} c_{t-1}(\gamma - \delta, \varphi) c_{t-1}(\gamma - \delta, \varphi)}{\left(\sum_{t=1}^{T} c_{t-1}^2(\gamma - \delta, \varphi)\right)^2}.
\]

Next, the left-hand side of (S.69) is

\[
-\frac{1}{\log T} \sum_{t=1}^{T} c_{t-1}(\gamma_0 - \delta_0, \varphi_0) c_{t-1}(\gamma_0 - \delta_0, \varphi_0)
= -\frac{1}{\log T} \frac{\sum_{t=1}^{T} \pi_{t-1}(\gamma_0 - \delta_0) \pi_{t-1}(\gamma_0 - \delta_0)}{\sum_{t=1}^{T} \pi_{t-1}^2(\gamma_0 - \delta_0)} + o(1),
\]

as in (S.49) and (S.51). Thus (S.69) follows immediately, noting (S.50), (S.53) and that, by (S.105) and Lemma 10, for \(d > 1/2,\)

\[
\frac{1}{T^{2d-1} \log T} \sum_{t=1}^{T} \log (d + t - 1) \pi_{t-1}^2(d) \rightarrow \frac{1}{\Gamma^2(d)(2d-1)}.
\]

Next, the left-hand side of (88) is

(S.70)
\[
\frac{T^{\gamma_0-\delta_0-1/2}}{\log T} \sum_{t=1}^{T} \phi(L; \varphi_0) \{u_t^I(t > 0)\} k_{t-1,T}(\gamma_0 - \delta_0, \varphi_0)
\]
so, in view of Theorem 1(ii) and (89), the result holds on showing
\[ \text{(S.71)} \]
\[ \frac{T^{\gamma_0 - \delta_0 - 1/2}}{\log T} \sum_{t=1}^{T} (\phi(L; \hat{\varphi}) u_t(\hat{\delta} - \delta_0) - \phi(L; \varphi_0) \{u_t I(t > 0)\}) k_{t-1,T}^{(\gamma_0 - \delta_0, \varphi_0)} \]
\[ \text{(S.72)} \]
\[ \frac{T^{\gamma_0 - \delta_0 - 1/2}}{\log T} \sum_{t=1}^{T} \phi(L; \varphi_0) \{u_t I(t > 0)\} (k_{t-1,T}^{(\hat{\gamma} - \hat{\delta}, \varphi)} - k_{t-1,T}^{(\gamma_0 - \delta_0, \varphi_0)}) \]
\[ \frac{T^{\gamma_0 - \delta_0 - 1/2}}{\log T} \sum_{t=1}^{T} \phi(L; \varphi_0) u_t(\delta - \delta_0) - \phi(L; \varphi_0) \{u_t I(t > 0)\} \]
\[ \text{(S.73)} \]
\[ \times (k_{t-1,T}^{(\hat{\gamma} - \hat{\delta}, \varphi)} - k_{t-1,T}^{(\gamma_0 - \delta_0, \varphi_0)}). \]

First, using summation by parts and (S.94), (S.95) in the proof of Lemma 7 and noting that, as in Lemmas S.3 and S.4,
\[ \text{(S.74)} \]
\[ k_{t-1,T}^{(\gamma_0 - \delta_0, \varphi_0)} = O_p(T^{\gamma_0 - \delta_0 - 1} T^{1 - 2(\gamma_0 - \delta_0)}), \]
it follows that (S.70) is \( O_p(\log^{-1} T) \). Next, by (S.74) and Lemma S.5 with \( \kappa = 1/2 \) (because \( \hat{\tau} \) is \( T^{1/2} \)-consistent), (S.71) is \( O_p(T^{1/2 - \kappa} \log^{-1} T) = O_p(T \log^{-1} T) \). Next, by summation by parts, (S.94), (S.95), the mean value theorem and Lemmas 1 and S.5, it can be easily shown that (S.72) 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 (S.71) and (S.72), it is straightforward to show that (S.73) is \( o_p(1) \), to conclude the proof of (88).

**S.9. Additional proof details for (93).** First, for any \( \epsilon > 0 \), clearly
\[ \Pr \left( \| T^{1/2} \partial S_T(\hat{\theta}) / \partial \tau \| \geq \epsilon \right) = \Pr \left( \| T^{1/2} \partial S_T(\hat{\theta}) / \partial \tau \| \geq \epsilon, \| \hat{\tau} - \tau_0 \| < \epsilon \right) \]
\[ + \Pr \left( \| T^{1/2} \partial S_T(\hat{\theta}) / \partial \tau \| \geq \epsilon, \| \hat{\tau} - \tau_0 \| \geq \epsilon \right) \]
\[ \leq \Pr \left( \| T^{1/2} \partial S_T(\hat{\theta}) / \partial \tau \| \geq \epsilon, \| \hat{\tau} - \tau_0 \| < \epsilon \right) \]
\[ + \Pr (\| \hat{\tau} - \tau_0 \| \geq \epsilon), \]
so, in view of Theorem 1(ii) and (89), the result holds on showing
\[ \text{(S.75)} \]
\[ \sup_{\theta \in \mathcal{M}_c} \frac{1}{T^{1/2}} \sum_{t=1}^{T} (d_t(\theta) - s_{2t}(\theta)) \left( \frac{\partial d_t(\theta)}{\partial \tau} - \frac{\partial s_{2t}(\theta)}{\partial \tau} \right) = o_p(1), \]
\[ \text{(S.76)} \]
\[ \sup_{\theta \in \mathcal{M}_c} \frac{1}{T^{1/2}} \sum_{t=1}^{T} s_{1t}(\tau) \left( \frac{\partial d_t(\theta)}{\partial \tau} - \frac{\partial s_{2t}(\theta)}{\partial \tau} \right) = o_p(1), \]
The proof of (S.75) follows upon showing that, for any \( \theta > 0 \) and \( \varepsilon \) such that \( 0 < \varepsilon < \theta \),

\[
\sup_{\vartheta \in M} \left| d_t(\vartheta) \right| = O(t^{\max\{\gamma_0 - \delta_0 + \varepsilon - 1, -1 - \varsigma\} + T^{2\theta} t^{-1/2 - \theta}}),
\]

(S.78)

\[
\sup_{\vartheta \in M} \left| s_{2t}(\vartheta) \right| = O_p(T^{2\theta} t^{-1/2 - \theta}),
\]

(S.79)

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

(S.80)

\[
\sup_{\vartheta \in M} \left\| \frac{\partial s_{2t}(\vartheta)}{\partial \tau} \right\| = O_p(T^{4\theta} t^{-1/2 - \theta}),
\]

(S.81)

and then letting \( \theta \) be sufficiently small. We only show (S.80) and (S.81) because the proofs for (S.78) and (S.79) are very similar but simpler. First, by (S.106) of Lemma S.3

\[
\sup_{\vartheta \in M} \left| c_t(\gamma_0 - \delta, \varphi) \right| = O(t^{\max\{\gamma_0 - \delta_0 + \varepsilon - 1, -1 - \varsigma\}}),
\]

(S.82)

and by a simple modification of that result

\[
\sup_{\vartheta \in M} \left\| \frac{\partial c_t(\gamma_0 - \delta, \varphi)}{\partial \tau} \right\| = O(t^{\max\{\gamma_0 - \delta_0 + \varepsilon - 1, -1 - \varsigma\} \log t}).
\]

(S.83)

Then (S.80) follows by direct application of (94) of Lemma 1, (S.82) and (S.83), noting that the bound in (94) also applies if the derivative is taken with respect to \( \tau \).

To prove (S.81) we apply (S.46), where the \( \sup_{\vartheta \in M} \) 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 (94) of Lemma 1 and (106) of Lemma 5, noting that \( \delta_0 - \delta \leq \varepsilon \) and that these bounds also apply if the derivatives are taken with respect to \( \tau \). For the second term on the right-hand side of (S.46), noting that \( \sum_{j=0}^{t-1} s_{1j}(\tau) = \sum_{j=0}^{t-1} c_j(\delta_0 - \delta, \varphi) \sum_{l=1}^{t-j} u_l \), it is straightforward to show that, by (S.106) of Lemma S.3,

\[
\sup_{\vartheta \in M} \left\| \sum_{j=1}^{t} \frac{\partial s_{1j}(\tau)}{\partial \tau} \right\| = O_p(t^{1/2 + \varepsilon / 2} \log t).
\]

(S.84)

Therefore, using summation by parts as in the proof of Lemma 5 and by (94) of Lemma 1, the \( \sup_{\vartheta \in M} \) of the absolute value of the second term on
the right-hand side of (S.46) is $O_p(T^{2\theta} t^{-1/2-\theta})$, to justify (S.81) and hence (S.75).

Finally, (S.76) and (S.77) can be established by using summation by parts followed by direct application of the results in (S.78), (S.80), (S.84), and Lemma 2, noting also that by previous arguments it can be easily shown that

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

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

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

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

The left-hand side of (94) is bounded by (S.85)

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

Suppose first that $m = 0$. Using the definition (17) and applying (S.106) of Lemma S.3 and (S.110) of Lemma S.4, the first term of (S.85) is bounded by

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

so the bound in (94) applies to the first term of (S.85) (although it is not tight). Next, the second term of (S.85) is bounded by

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

By (S.107) of Lemma S.3 the numerator is $O(t^{-1}(t/T)^{1/2-\theta})$ and by Lemma S.4 the denominator is bounded from below by $e\theta^{-1/2} T^{-1/2}$. Thus (94) for $m = 0$ follows.
Next, for the derivative we find
(S.86)
\[
\frac{\partial h_{t-1, T}(d, \varphi)}{\partial d} = \frac{c_{l-1}^{(1)}(d, \varphi)}{\left(\sum_{j=1}^{T} c_{j-1}^2(d, \varphi)\right)^{1/2}} - \frac{h_{t-1, T}(d, \varphi) \sum_{j=1}^{T} c_{j-1}(d, \varphi) c_{j-1}^{(1)}(d, \varphi)}{\sum_{j=1}^{T} c_{j-1}^2(d, \varphi)}.
\]

First we show (94). Proceeding as in the proof for \( m = 0 \), taking into account the extra log-term arising from (S.106) in Lemma S.3, the first term of (S.86) is \( O(t^{-1/2} (T/t)^{\theta} \log T) \), so the bound in (94) applies. Next, using again (S.106) in Lemma S.3 and also (94) for \( m = 0 \), the second term of (S.86) is \( O(t^{-1/2} (T/t)^{\theta} T^{2g} \sum_{j=1}^{T} j^{-1-2g} \log j) \), so the bound in (94) applies for \( m = 1 \).

Next, we show (95). Clearly
(S.87)
\[
\frac{\partial h_{t-1, T}(d, \varphi)}{\partial d} = \frac{\partial T^{-d} c_{l-1}(d, \varphi)}{\partial d} \left(\sum_{j=1}^{T} T^{-2d} c_{j-1}^2(d, \varphi)\right)^{1/2}
- \frac{T^{-d} c_{l-1}(d, \varphi) \sum_{j=1}^{T} T^{-d} c_{j-1}(d, \varphi) \partial T^{-d} c_{j-1}(d, \varphi)}{\left(\sum_{j=1}^{T} T^{-2d} c_{j-1}^2(d, \varphi)\right)^{3/2}}.
\]

First,
(S.88)
\[
\sup_{d \geq 1/2 + \theta, \varphi \in \Psi} \left| \frac{\partial T^{-d} c_{l-1}(d, \varphi)}{\partial d} \right| \leq \sup_{d \geq 1/2 + \theta, \varphi \in \Psi} \left| \frac{\partial T^{-d} c_{l-1}(d, \varphi)}{\partial d} \right| \left(\inf_{d \geq 1/2 + \theta, \varphi \in \Psi} \sum_{j=1}^{T} T^{-2d} c_{j-1}^2(d, \varphi)\right)^{1/2}
= O\left(t^{-1/2} \left(\frac{t}{T}\right)^{\theta} (1 + |\log (t/T)|)\right),
\]
by (S.107) of Lemma S.3 and (S.111) of Lemma S.4. Similarly, like in (S.87),
(S.89)
\[
\sup_{d \geq 1/2 + \theta, \varphi \in \Psi} \left| \frac{T^{-d} c_{l-1}(d, \varphi)}{\left(\sum_{j=1}^{T} T^{-2d} c_{j-1}^2(d, \varphi)\right)^{1/2}} \right| = O\left(t^{-1/2} \left(\frac{t}{T}\right)^{\theta}\right),
\]
so by (S.87) and (S.88) it is straightforward to show that
(S.90)
\[
\sup_{d \geq 1/2 + \theta, \varphi \in \Psi} \left| \frac{T^{-d} c_{l-1}(d, \varphi) \sum_{j=1}^{T} T^{-d} c_{j-1}(d, \varphi) \partial T^{-d} c_{j-1}(d, \varphi)}{\left(\sum_{j=1}^{T} T^{-2d} c_{j-1}^2(d, \varphi)\right)^{3/2}} \right| = O\left(t^{-1/2} \left(\frac{t}{T}\right)^{\theta}\right),
\]
to conclude the proof of (95).

The proofs of (96)–(98) are omitted because they follow by identical arguments, noting that

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

**S.11. Proof of Lemma 2.** First we show (99). Write $\phi(L;\varphi) u_t(-d) = \sum_{j=0}^{t-1} c_j(d,\varphi) u_{t-j}$ and apply summation by parts,

$$\sum_{j=0}^{t-1} c_j(d,\varphi) u_{t-j} = c_{t-1}(d,\varphi) \sum_{j=0}^{t-1} u_{t-j} - \sum_{j=0}^{t-2} (c_{j+1}(d,\varphi) - c_j(d,\varphi)) \sum_{l=0}^{j} u_{t-l}.$$

Noting that $c_{j+1}(d,\varphi) - c_j(d,\varphi) = c_{j+1}(d-1,\varphi)$, the right-hand side of (S.89) is bounded by

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

Under our conditions, $E \left| \sum_{t=1}^{T} u_t \right| = O(t^{1/2})$, so, in view of (S.106) of Lemma S.3, the expectation of the left-hand side of (99) is bounded by

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

to conclude the proof of (99). The proof of (100) is omitted because it is almost identical to that for (99).

**S.12. Proof of Lemma 3.** The results follow by direct application of (94), (95) of Lemma 1 and (S.106), (S.107) of Lemma S.3.

**S.13. Proof of Lemma 4.** By summation by parts we find

$$\left| \sum_{t=1}^{T} \phi(L;\varphi) u_t(\delta - \delta_0) c_{t-1}(\gamma_0 - \delta,\varphi) \right| \leq |c_{T-1}(\gamma_0 - \delta,\varphi)| \left| \phi(L;\varphi) u_T(\delta - \delta_0 - 1) \right|$$

$$+ \sum_{t=1}^{T-1} c_t(\gamma_0 - \delta - 1,\varphi) \phi(L;\varphi) u_t(\delta - \delta_0 - 1),$$

(S.91)
noting that for any \( k \geq 1 \) it holds that \( \sum_{l=1}^{k} \phi(L; \varphi)u_l(\delta - \delta_0) = \phi(L; \varphi)u_k(\delta - \delta_0 - 1) \) and \( \pi_{j+1}(d) - \pi_j(d) = \pi_{j+1}(d - 1) \), and therefore

\[
(S.92) \quad c_t(d, \varphi) - c_{t-1}(d, \varphi) = c_t(d - 1, \varphi) .
\]

The result (104) then follows by application of (99) of Lemma 2 and (S.106) of Lemma S.3, while the result (105) follows by application of (100) of Lemma 2 and (S.107) of Lemma S.3.

**S.14. Proof of Lemma 5.** By summation by parts as in (S.91), we find

\[
\left| \sum_{t=1}^{T} \phi(L; \varphi)u_t(\delta - \delta_0) h_{t-1,T}(\gamma - \delta, \varphi) \right| \leq |h_{T-1,T}(\gamma - \delta, \varphi)| |\phi(L; \varphi)u_T(\delta - \delta_0 - 1)| + \sum_{t=1}^{T-1} |h_{t,T}(\gamma - \delta, \varphi) - h_{t-1,T}(\gamma - \delta, \varphi)| |\phi(L; \varphi)u_t(\delta - \delta_0 - 1)| .
\]

First, application of (94), (96) of Lemma 1 together with (99), (100) of Lemma 2 implies (106) and (107). Next, (108) and (109) follow from (95), (97) of Lemma 1 and (99), (100) of Lemma 2.

**S.15. Proof of Lemma 6.** Letting \( d_t(\tau, \gamma) = d_t(\vartheta) \), noting (29) and that \( d_t(\tau, \gamma_0) = 0 \), by the mean value theorem,

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

where \(|\gamma - \gamma_0| \leq |\gamma - \gamma_0|\). Then we find the bound

\[
\left| \frac{\partial}{\partial \gamma} \sum_{t=1}^{T} d_t(\vartheta, \gamma) \phi(L; \varphi)u_t(\delta - \delta_0) \right| \leq \sum_{t=1}^{T} \phi(L; \varphi)u_t(\delta - \delta_0) \frac{\partial h_{t-1,T}(\gamma - \delta, \varphi)}{\partial \gamma} \sum_{j=1}^{T} c_{j-1}(\gamma_0 - \delta, \varphi) h_{j-1,T}(\gamma - \delta, \varphi) + \sum_{t=1}^{T} \phi(L; \varphi)u_t(\delta - \delta_0) h_{t-1,T}(\gamma - \delta, \varphi) \sum_{j=1}^{T} c_{j-1}(\gamma_0 - \delta, \varphi) \frac{\partial h_{j-1,T}(\gamma - \delta, \varphi)}{\partial \gamma} .
\]

The results (111)–(113) now all follow by direct application of (101), (102) of Lemma 3 with \( \theta < g - 1/2 \) and (106), (107) of Lemma 5. Results (114) and (115) are derived straightforwardly from (103) of Lemma 3 and (108), (109) of Lemma 5.

\[ \frac{1}{T} \sum_{t=1}^{T} s_t^2 (\vartheta_0) = \frac{1}{T} \sum_{t=1}^{T} (\phi(L; \varphi_0) \{ u_t \mathbb{I}(t > 0) \})^2 \]

(S.93)

\[ - \frac{1}{T} \left( \sum_{t=1}^{T} \phi(L; \varphi_0) \{ u_t \mathbb{I}(t > 0) \} h_{t-1,T} (\gamma_0 - \delta, \varphi_0) \right)^2 . \]

Next, note that

(S.94)

\[ \phi(L; \varphi_0) \{ u_t \mathbb{I}(t > 0) \} = \varepsilon_t - \sum_{j=t}^{\infty} \phi_j (\varphi_0) u_{t-j}, \]

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

(S.95)

\[ \sum_{j=t}^{\infty} \phi_j (\varphi_0) u_{t-j} = O_p(t^{-1/2-\varsigma}). \]

Then, in view of (S.94), (S.95), by Assumptions A1 and A2 and simple application of Lemma 1, the second term on the right-hand side of (S.93) is

\[ O_p(T^{-2} \theta - 1) \]

by choosing \( \theta < 1/2 \). Then the required result holds by (24).

S.17. Proof of Lemma 8. Letting \( \alpha > 0 \) be arbitrarily small (in particular \( \alpha < (\varsigma - 1/2)/3 \), which implies \( \alpha < 1/2 \) and defining \( \Phi_1 = \{ \vartheta \in \Xi : \gamma - \delta \leq 1/2 - \alpha \} \), \( \Phi_2 = \{ \vartheta \in \Xi : 1/2 - \alpha \leq \gamma - \delta \leq 1/2 + \alpha \} \), and \( \Phi_3 = \{ \vartheta \in \Xi : \gamma - \delta \geq 1/2 + \alpha \} \), the result holds on showing

(S.96)

\[ \lim_{T \to \infty} \inf_{\gamma_0 - \delta \geq 1/2 + g, |\gamma - \gamma_0| \geq g} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_t^2 (\vartheta) > \epsilon, \]

for \( j = 1, 2, 3 \). We first deal with \( j = 1, 2 \). Clearly

\[ \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_t^2 (\vartheta) = \frac{\mu_0^2}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} c_{t-1}^2 (\gamma_0 - \delta, \varphi) \]

\[ - \frac{\mu_0^2}{T^{2(\gamma_0 - \delta) - 1}} \left( \sum_{t=1}^{T} c_{t-1} (\gamma_0 - \delta, \varphi) h_{t-1,T} (\gamma - \delta, \varphi) \right)^2 , \]

so because \( |\gamma - \gamma_0| \geq g \) and by application of Lemma S.4, noting that \( \gamma_0 - \delta \geq 1/2 + g > 1/2, \mu_0 \neq 0 \), and (8), (S.96) for \( j = 1, 2 \) holds on showing

(S.97)

\[ \lim_{T \to \infty} \inf_{\gamma_0 - \delta \geq 1/2 + g} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} d_{t-1}^2 (\gamma_0 - \delta) > \epsilon, \]
First, we show (S.97). By (S.105) in Lemma S.2,

\[
\sup_{\{\gamma_0 - \delta \geq 1/2 + g\} \cap \Phi_4} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \left( \sum_{t=1}^{T} c_{t-1} (\gamma_0 - \delta, \varphi) h_{t-1,T} (\gamma - \delta, \varphi) \right)^2 = o(1).
\]

By Lemma S.2, (S.105) in Lemma S.2, the left-hand side of (S.98) is bounded by

\[
\inf_{\gamma_0 - \delta \geq 1/2 + g} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} \pi_{t-1}^2 (\gamma_0 - \delta) \geq \epsilon \inf_{\gamma_0 - \delta \geq 1/2 + g} \frac{1}{T} \sum_{t=1}^{T} \left( \frac{t}{T} \right)^{2(\gamma_0 - \delta) - 2} = \frac{1}{2g} + o(1),
\]

so that (S.97) holds by taking limits as \( T \to \infty \). Next, we show (S.98) for \( j = 1 \). By (S.110) of Lemma S.4, the left-hand side of (S.98) is bounded by

\[
\sup_{\{\gamma_0 - \delta \geq 1/2 + g\} \cap \Phi_4} \left( T^{1/2 - (\gamma_0 - \delta)} \sum_{t=1}^{T} c_{t-1} (\gamma_0 - \delta, \varphi) c_{t-1} (\gamma - \delta, \varphi) \right)^2.
\]

By Lemma S.3, (S.100) is \( O(T^{-\alpha}) = o(1) \) to conclude the proof of (S.98), and therefore that of (S.96), for \( j = 1 \). Regarding \( j = 2 \), the left-hand side of (S.98) is bounded by

\[
\left( \sup_{\{\gamma_0 - \delta \geq 1/2 + g\} \cap \Phi_2} T^{-(\gamma_0 - 2\delta + \gamma - 1)} \sum_{t=1}^{T} c_{t-1} (\gamma_0 - \delta, \varphi) c_{t-1} (\gamma - \delta, \varphi) \right)^2.
\]

where the denominator can be made arbitrarily large by setting \( \alpha \) close enough to zero, see (S.111) of Lemma S.4. By (S.107) of Lemma S.3 the square-root of the numerator of (S.101) is \( O(T^{-1/2 - \gamma + \alpha}) = O(1) \). This completes the proof of (S.98), and hence that of (S.96), for \( j = 2 \).

Finally we show (S.96) for \( j = 3 \). By very similar steps to those in the proof of Lemma S.4, 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 (\varphi) = \frac{\mu_0^2 \varphi^2 (1; \varphi)}{T^{2(\gamma_0 - \delta) - 1}} \left( \sum_{t=1}^{T} \pi_{t-1}^2 (\gamma_0 - \delta) - \frac{\left( \sum_{t=1}^{T} \pi_{t-1} (\gamma_0 - \delta) \pi_{t-1} (\gamma - \delta) \right)^2}{\sum_{t=1}^{T} \pi_{t-1}^2 (\gamma - \delta)} \right)

+ qT (\gamma_0, \gamma, \delta, \varphi),
\]

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

\[
\inf_{\{\gamma_0 - \delta \geq 1/2 + g\} \cap \Phi_3} \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \left[ \sum_{t=1}^{T} \pi_{t-1}^2 (\gamma_0 - \delta) - \frac{\left( \sum_{t=1}^{T} \pi_{t-1} (\gamma_0 - \delta) \pi_{t-1} (\gamma - \delta) \right)^2}{\sum_{t=1}^{T} \pi_{t-1}^2 (\gamma - \delta)} \right]
\]

\[
\geq \epsilon \inf_{\{\gamma_0 - \delta \geq 1/2 + g\}} \frac{1}{\Gamma^2 (\gamma_0 - \delta)} \left[ \frac{1}{T^{2(\gamma_0 - \delta) - 1}} \sum_{t=1}^{T} \pi_{t-1}^2 (\gamma_0 - \delta - 1) - \frac{1}{\Gamma^2 (\gamma_0 - \delta)} \sum_{t=1}^{T} \pi_{t-1}^2 (\gamma_0 + \gamma - 2\delta - 1) \right]^2 - o(1)
\]

\[
\geq \epsilon \inf_{\gamma_0 - \delta \geq 1/2 + g} \frac{g^2}{\Gamma^2 (\gamma_0 - \delta) 2g (\alpha + g)^2} - o(1),
\]

which is positive and bounded away from zero, to complete the proof of (S.96) for \( j = 3 \).

**S.18. Proof of Lemma 9.** First, \( \frac{\partial d_T(\vartheta)}{\partial \gamma} \) equals

\[
- \mu_0 \sum_{j=1}^{T} c_{j-1}^2 (\gamma - \delta, \varphi) \sum_{j=1}^{T} c_{j-1} (\gamma_0 - \delta, \varphi) c_{j-1} (\gamma - \delta, \varphi)
\]

\[
- \mu_0 \sum_{j=1}^{T} c_{j-1} (\gamma - \delta, \varphi) \sum_{j=1}^{T} c_{j-1}^2 (\gamma_0 - \delta, \varphi) c_{j-1}^2 (\gamma - \delta, \varphi)
\]

\[
+ \frac{2\mu_0 \sum_{j=1}^{T} c_{j-1} (\gamma - \delta, \varphi) \sum_{j=1}^{T} c_{j-1} (\gamma_0 - \delta, \varphi) c_{j-1} (\gamma - \delta, \varphi) \sum_{j=1}^{T} c_{j-1} (\gamma - \delta, \varphi) c_{j-1}^2 (\gamma - \delta, \varphi)}{(\sum_{j=1}^{T} c_{j-1}^2 (\gamma - \delta, \varphi))^2}.
\]

Noting that in \( \cup_{i=1}^{d} \mathcal{H}_i \), \( \gamma_0 - \delta \geq 1 + \eta \) and \( \gamma - \delta \geq 1 + \eta - \varrho \), proceeding as in the proof of Lemma S.4,

\[
\frac{1}{T^{\gamma_0 - \delta}} \frac{\partial d_T(\tau, \gamma)}{\partial \gamma} =
\]

\[
- \frac{\mu_0 \phi(1; \varphi)}{T^{\gamma_0 - \delta}} \left( \sum_{j=1}^{T} \pi_{j-1} (\gamma - \delta) \sum_{j=1}^{T} \pi_{j-1} (\gamma_0 - \delta) \pi_{j-1} (\gamma - \delta) \right) - \frac{\sum_{j=1}^{T} \pi_{j-1}^2 (\gamma - \delta) \sum_{j=1}^{T} \pi_{j-1} (\gamma_0 - \delta) \pi_{j-1} (\gamma - \delta)}{\sum_{j=1}^{T} \pi_{j-1}^2 (\gamma - \delta)}
\]

\[
+ \frac{\sum_{j=1}^{T} \pi_{j-1} (\gamma - \delta) \sum_{j=1}^{T} \pi_{j-1} (\gamma_0 - \delta) \pi_{j-1}^2 (\gamma - \delta)}{\sum_{j=1}^{T} \pi_{j-1}^2 (\gamma - \delta)}
\]
\[ -2 \sum_{j=1}^{T} \pi_{j-1} (\gamma - \delta) \sum_{j=1}^{T} \pi_{j-1} (\gamma - \delta) \sum_{j=1}^{T} \pi_{j-1} (\gamma - \delta) \pi_{j-1}^{(1)} (\gamma - \delta) \] 
\[ \frac{2}{(\sum_{j=1}^{T} \pi_{j-1} (\gamma - \delta))^2} \] 
(S.102) 
\[ + g_{1,T}(\vartheta), \]

where \( \sup_{\vartheta} |g_{1,T}(\vartheta)| = O(1) \). Now, substituting (S.50) into (S.102) (noting that the contribution of \( \psi(d) \) cancels), approximating \( \psi(d + j) \) by \( \log j \), introducing \( \log T \) terms (which cancel) and using (S.105) in Lemma S.2, it can be shown that

\[ \frac{1}{T^{\gamma_0 - \delta}} \frac{\partial \tilde{d}_{T} (\tau, \gamma)}{\partial \gamma} = - \frac{\mu_0 \phi(1; \varphi)}{\Gamma (\gamma_0 - \delta) T^{\gamma_0 - \delta}} \frac{\sum_{j=1}^{T} \log (j/T) \sum_{j=1}^{T} j^{\gamma_0 + \gamma - 2\delta}}{\sum_{j=1}^{T} j^{2\gamma - 2\delta - 2}} \] 
\[ + \sum_{j=1}^{T} \frac{j^{\gamma - 1} \sum_{j=1}^{T} \log (j/T) j^{\gamma_0 + \gamma - 2\delta}}{\sum_{j=1}^{T} j^{2\gamma - 2\delta - 2}} \] 
(S.103) 
\[ - 2 \sum_{j=1}^{T} \frac{j^{\gamma - 1} \sum_{j=1}^{T} j^{\gamma_0 + \gamma - 2\delta - 2} \sum_{j=1}^{T} \log (j/T) j^{2\gamma - 2\delta - 2}}{\sum_{j=1}^{T} j^{2\gamma - 2\delta - 2}} \] 
\[ + g_{2,T}(\vartheta), \]

where \( \sup_{\vartheta} |g_{2,T}(\vartheta)| = O(1) \). Finally, (116) then follows by approximating sums by integrals, see Lemma 10.

Define \( \tilde{d}_{T} (\tau, \gamma) = \tilde{d}_{T} (\vartheta) \). Because \( \tilde{d}_{T} (\tau, \gamma_0) = 0 \), the mean value theorem yields \( \tilde{d}_{T} (\tau, \gamma) = (\gamma - \gamma_0) \partial \tilde{d}_{T} (\tau, \gamma)/\partial \gamma \), where \( |\gamma - \gamma_0| \leq |\gamma - \gamma_0| \), so the left-hand side of (117) can be bounded from below by

\[ \lim_{T \to \infty} \inf_{\vartheta} T^{2\kappa_1} (\gamma - \gamma_0)^2 \left( \frac{1}{T^{\gamma_0 - \delta}} \frac{\partial \tilde{d}_{T} (\tau, \gamma)}{\partial \gamma} \right)^2 \geq \lim_{T \to \infty} \inf_{\vartheta} T^{2\kappa_1} (\gamma - \gamma_0)^2 \inf_{\vartheta} \left( \frac{1}{T^{\gamma_0 - \delta}} \frac{\partial \tilde{d}_{T} (\tau, \gamma)}{\partial \gamma} \right)^2 \] 
\[ = \xi^2 \lim_{T \to \infty} \inf_{\vartheta} \left( \frac{1}{T^{\gamma_0 - \delta}} \frac{\partial \tilde{d}_{T} (\tau, \gamma)}{\partial \gamma} \right)^2. \]

Thus, setting \( \xi \) large enough, (117) follows if

\[ (S.104) \lim_{T \to \infty} \inf_{\vartheta} \left( \frac{1}{T^{\gamma_0 - \delta}} \frac{\partial \tilde{d}_{T} (\tau, \gamma)}{\partial \gamma} \right)^2 > \epsilon, \]

which, noting that \( \gamma_0 - \delta \geq 1 + \eta \), is a consequence of (8) and (116) because

\[ \inf_{\vartheta} \left( 2 (\gamma - \delta)^2 - 2 (\gamma - \delta) + 1 - (\gamma_0 - \delta) \right) = \inf_{\vartheta} \left( 2 (\gamma - \delta)^2 - 3 (\gamma - \delta) + 1 - (\gamma_0 - \gamma) \right) \] 
\[ \geq 2 (\eta - \delta)^2 + \eta - 2 \delta > 0. \]
S.19. 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

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

Approximating the sum by an integral we find

$$\frac{1}{T} \sum_{t=1}^{T} \log \left( \frac{t}{T} \right) t^{d-1} \sim \int_{0}^{1} \log (x) x^{d-1} dx = B(d, 1) (\psi(d) - \psi(d + 1)),$$

see p. 535 of [2], where $B(x, y) = \Gamma(x) \Gamma(y) / \Gamma(x + y)$ 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 [1]. Similarly, $T^{-d} \sum_{t=1}^{T} \log^2 \left( \frac{t + a}{T} \right) t^{d-1}$ can be approximated by

$$\frac{1}{T} \sum_{t=1}^{T} \log^2 \left( \frac{t}{T} \right) t^{d-1} \sim \int_{0}^{1} \log^2 (x) x^{d-1} dx$$

$$= B(d, 1) \left( (\psi(d) - \psi(d + 1))^2 + \psi'(d) - \psi'(d + 1) \right),$$

see p. 538 of [2], 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 [1].

S.20. Technical lemmas.

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

**Proof.** The proof of Lemma S.1 is given in Lemma B.4 of [6].

**Lemma S.2.** Let $j \geq 1$ and $K$ denote any compact subset of $\mathbb{R}\setminus\mathbb{N}_0$. Then

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

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

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

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

**Proof.** The proof of Lemma S.2 is given in Lemma B.3 of [6] and Lemma A.5 of [7].

**Lemma S.3.** Under Assumptions A1, A3, uniformly in \( t = 1, \ldots, T \) and \( T \geq 1 \), for \( m \geq 0 \),

\[
\sup_{d \geq g, \varphi \in \Psi} \left| \frac{\partial^m c_t(d, \varphi)}{\partial d^m} \right| = O\left( t^{\max\{g-1, -1\}} (\log t)^m \right),
\]

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

**Proof.** The proof of S.3 is almost identical to that of Lemma 1 of [5] and is therefore omitted.

**Lemma S.4.** Under Assumptions A1 and A3,

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

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

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

\[
\inf_{d_1 \leq d_2 \leq T, \varphi \in \Psi} \sum_{t=1}^{T} c_{t-1}^2(d, \varphi) \geq \frac{1}{T^{2d-1}} \sum_{t=1}^{T} \pi_{t-1}^2(d) \geq \frac{\epsilon}{\alpha} + o(1),
\]

for some \( \epsilon > 0 \), which does not depend on \( \alpha \) or \( T \).

**Proof.** By summation by parts,

\[
\sum_{t=1}^{T} c_{t-1}^2(d, \varphi) \geq \frac{\phi^2(1; \varphi)}{T^{2d-1}} \sum_{t=1}^{T} \pi_{t-1}^2(d) - 2\phi(1; \varphi) \sum_{t=1}^{T} \pi_{t-1}^2(d) \sum_{k=t}^{\infty} \phi_k(\varphi).
\]

First we show (S.108). Clearly

\[
\sum_{t=1}^{T} c_{t-1}^2(d, \varphi) \geq \frac{\phi^2(1; \varphi)}{T^{2d-1}} \sum_{t=1}^{T} \pi_{t-1}^2(d) - 2\phi(1; \varphi) \sum_{t=1}^{T} \pi_{t-1}^2(d) \sum_{k=t}^{\infty} \phi_k(\varphi).
\]
\[-2\phi(1; \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} (\varphi) + 2 \sum_{t=1}^{T} \pi_{t-1} (d) \sum_{j=t}^{\infty} \phi_{j} (\varphi) \sum_{k=0}^{t-2} \pi_{k+1} (d - 1) \sum_{l=0}^{k} \phi_{t-1-l} (\varphi) \).

(S.113)

Noting (7), the fourth term on the right-hand side of (S.113) is of smaller order than the third term. Then the proof of (S.108) follows on showing

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

(S.114)

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

(S.115)

First, by (7) and Lemma S.2, the left-hand side of (S.114) is bounded by

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

so (S.114) holds. Similarly, the left-hand side of (S.115) 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 \( \eta \) small, noting that \( \varsigma > 1/2 \), where the second inequality is due to Lemma S.1, to conclude the proof of (S.108). Next, in view of (S.112), the proof of (S.109) follows by (S.114), (S.115) and

\[
\sup_{d \geq 1/2 + \eta, \varphi \in \Psi} \frac{1}{T^{2d-1}} \sum_{t=1}^{T} \pi_{t-1}^{2} (d) \left( \sum_{k=t}^{\infty} \phi_{k} (\varphi) \right)^{2} = o(1),
\]
The proof of (S.110) is immediate because which follow by straightforward arguments using (S.95) and Lemma S.1. The proof of (S.110) is immediate because
\[
\inf_{d_1 \leq d \leq d_2, \varphi \in \Psi} \sum_{t=1}^{T} \left( \sum_{k=0}^{d-1} \pi_{k} \right) \phi_{t-1} = \text{O}(1),
\]
which follow by straightforward arguments using (S.95) and Lemma S.1. Then (S.119) is bounded by (S.112), which is bounded by (S.119). For the proof of (S.111), it clearly holds that (S.110) is bounded from below by (S.111). Then, as in (S.112), the right-hand side of (S.116) is bounded from below by (S.117).

\[
\epsilon \inf_{d \geq 1/2 - \alpha} \frac{1}{T^{d-1}} \sum_{t=1}^{T} \left( \sum_{k=0}^{d} \pi_{k+1} \phi_{t-1} \right)^2 = \text{O}(1),
\]

which follows by straightforward arguments using (S.95) and Lemma S.1. Then (S.119) is bounded from below by (S.112), which is bounded by (S.119). For the proof of (S.111), it clearly holds that (S.110) is bounded from below by (S.111). Then, as in (S.112), the right-hand side of (S.116) is bounded from below by (S.117).

\[
\epsilon \inf_{d \geq 1/2 - \alpha} \frac{1}{T^{d-1}} \sum_{t=1}^{T} \left( \sum_{k=0}^{d} \pi_{k+1} \phi_{t-1} \right)^2 = \text{O}(1),
\]

which follows by straightforward arguments using (S.95) and Lemma S.1. Then (S.119) is bounded from below by (S.112), which is bounded by (S.119). For the proof of (S.111), it clearly holds that (S.110) is bounded from below by (S.111). Then, as in (S.112), the right-hand side of (S.116) is bounded from below by (S.117).

\[
\epsilon \inf_{d \geq 1/2 - \alpha} \frac{1}{T^{d-1}} \sum_{t=1}^{T} \left( \sum_{k=0}^{d} \pi_{k+1} \phi_{t-1} \right)^2 = \text{O}(1),
\]

which follows by straightforward arguments using (S.95) and Lemma S.1. Then (S.119) is bounded from below by (S.112), which is bounded by (S.119). For the proof of (S.111), it clearly holds that (S.110) is bounded from below by (S.111). Then, as in (S.112), the right-hand side of (S.116) is bounded from below by (S.117).

\[
\epsilon \inf_{d \geq 1/2 - \alpha} \frac{1}{T^{d-1}} \sum_{t=1}^{T} \left( \sum_{k=0}^{d} \pi_{k+1} \phi_{t-1} \right)^2 = \text{O}(1),
\]

which follows by straightforward arguments using (S.95) and Lemma S.1. Then (S.119) is bounded from below by (S.112), which is bounded by (S.119). For the proof of (S.111), it clearly holds that (S.110) is bounded from below by (S.111). Then, as in (S.112), the right-hand side of (S.116) is bounded from below by (S.117).

\[
\epsilon \inf_{d \geq 1/2 - \alpha} \frac{1}{T^{d-1}} \sum_{t=1}^{T} \left( \sum_{k=0}^{d} \pi_{k+1} \phi_{t-1} \right)^2 = \text{O}(1),
\]
Taylor expansion, \(S.2\). Next, by straightforward application of Lemma \(S.2\) and a second-order \(A1-A4\), the left-hand side of (\(S.123\)) is of smaller order, whereas by (\(12\)), the first one is of smaller order, whereas by (\(12\)), the first one is \(O_p\left(T^{-\kappa}\right)\). In view of (\(S.116\), (\(S.118\), and (\(S.119\)), this proves (\(S.111\)).

**Lemma S.5.** Let \(\hat{\tau} - \tau_0 = O_p\left(T^{-\kappa}\right)\) for \(\kappa > 0\). Then, under Assumptions \(A1-A4\),

\[
\begin{align*}
  c_t(\hat{\delta}, \hat{\varphi}) &= c_t(\delta_0, \varphi_0) + O_p\left(T^{-\kappa}t^{\max\{\delta_0-1,1-c\}} \log^2 t\right), \\
  c_t^{(1)}(\hat{\delta}, \hat{\varphi}) &= c_t^{(1)}(\delta_0, \varphi_0) + O_p\left(T^{-\kappa}t^{\max\{\delta_0-1,1-c\}} \log^3 t\right),
\end{align*}
\]

and, uniformly in \(t = 1, ..., T\),

\[
\phi(L; \hat{\varphi}) u_t(\hat{\delta} - \delta_0) = \sum_{j=0}^{t-1} \phi_j(\hat{\varphi}) u_{t-j}(\hat{\delta} - \delta_0) = \sum_{j=0}^{t-1} \phi_j(\varphi_0) u_{t-j} + O_p\left(T^{-\kappa}\right).
\]

**Proof.** First we show (\(S.120\)). Clearly

\[
\begin{align*}
  c_t(\hat{\delta}, \hat{\varphi}) - c_t(\delta_0, \varphi_0) &= \sum_{j=0}^{t} (\phi_j(\hat{\varphi}) - \phi_j(\varphi_0)) \pi_{t-j}(\delta_0) + \sum_{j=0}^{t} (\pi_{t-j}(\hat{\delta}) - \pi_{t-j}(\delta_0)) \phi_j(\varphi_0) \\
  &= \sum_{j=0}^{t} (\phi_j(\hat{\varphi}) - \phi_j(\varphi_0)) (\pi_{t-j}(\hat{\delta}) - \pi_{t-j}(\delta_0))
\end{align*}
\]

Fix \(\epsilon < 1/2\). Then

\[
\phi_j(\hat{\varphi}) - \phi_j(\varphi_0) = (\phi_j(\hat{\varphi}) - \phi_j(\varphi_0)) (\mathbb{I} (\|\hat{\varphi} - \varphi_0\| < \epsilon) + \mathbb{I} (\|\hat{\varphi} - \varphi_0\| \geq \epsilon))
\]
so by the mean value theorem the left-hand side of (\(S.124\)) is bounded by

\[
\sup_{\|\varphi - \varphi_0\| < \epsilon} \left\| \frac{\partial \phi_j(\varphi)}{\partial \varphi} \right\| \|\hat{\varphi} - \varphi_0\| + K \sup_{\varphi_0} |\phi_j(\varphi)| \|\hat{\varphi} - \varphi_0\|^{N} 
\]

for any arbitrarily large fixed number \(N\). Then by (\(7\)) and the \(T^{-\kappa}\)-consistency of \(\hat{\tau}\), the second term in (\(S.125\)) is of smaller order, whereas by (\(12\)), the first one is \(O_p(T^{-\kappa}t^{-1-c})\). This implies that the first term on the right-hand side of (\(S.123\)) is \(O_p(T^{-\kappa}t^{\max\{\delta_0-1,1-c\}} \log t)\) by Lemmas \(S.1\) and \(S.2\). Next, by straightforward application of Lemma \(S.2\) and a second-order Taylor expansion, \(\pi_{t-j}(\hat{\delta}) - \pi_{t-j}(\delta_0) = O_p((t-j)^{\delta_0-1} (\log t)^{T^{-\kappa}})\), so by
(7) and Lemma S.1 the second term on the right-hand side of (S.123) is $O_p(T^{-\kappa}t^{\max\{0,1-\kappa\}} \log^2 t)$. Finally, combining the arguments for the first two terms, the third term on the right-hand side of (S.123) is of smaller order, to conclude the proof of (S.120). The proof of (S.121) is omitted. It is almost identical to that of (S.120) with the only difference that the coefficients $\pi^{(1)}_t(\cdot)$ instead of $\pi_t(\cdot)$ lead to an extra $\log t$ factor, see Lemma S.2. Finally, we show (S.122). Clearly the left-hand side of (S.122) is

\[
\sum_{j=0}^{t-1} \phi_j(\varphi_0) u_{t-j} + \sum_{j=0}^{t-1} (\phi_j(\hat{\varphi}) - \phi_j(\varphi_0)) u_{t-j} + \sum_{j=0}^{t-1} \phi_j(\hat{\varphi}) (u_{t-j}(\delta - \delta_0) - u_{t-j}).
\]

Using the mean value theorem as in (S.124) and (S.125) and summation by parts, it can be shown that the second term in (S.126) is $O_p(T^{-\kappa})$. Similarly, by Lemma C.5 of [8] and (7), the third term in (S.126) is also $O_p(T^{-\kappa})$, to conclude the proof of (S.122).

\[\Box\]

REFERENCES