Online Supplement to "CHRONOLOGICALLY TRIMMED LS FOR NONLINEAR PREDICTIVE REGRESSIONS WITH PERSISTENCE OF UNKNOWN FORM"

Zhishui Hu
University of Science & Technology of China

Ioannis Kasparis
University of Cyprus

Qiying Wang
The University of Sydney

October 12, 2024

This supplement is organized as follows. Section S1 provides a more detailed review of existing robust inferential methods under temporal dependence. In Section S2, we investigate CTLS inference in multi-covariate models, providing an extension of Theorem 3 given in Section 4 of the main paper. Section S3 provides the proofs of Lemmas 1 and 2 of the main paper. The proofs of the theorems in the main paper are given in Sections S4 to S10. Since Theorems 5 and 6 provide the basic tools for other proofs, we arrange the proofs of Theorem 5 and 6 in Sections S4 and S5, respectively. The proof of Theorem S1 of this supplement is given in Section S11. Section S12 gives some additional simulation results. Throughout the Supplement, unless mentioned explicitly, we use the same notation and equations as those given in the main paper.

S1 A Detailed Review of Existing Inferential Methods

The first approach put forward for addressing the dichotomy in inference, between stationary and nonstationary regimes, relies on so-called conservative methods. In particular, a number of studies develop procedures that yield robust inference in the presence of NI processes, in the context of reduced form regressions where the covariate is predetermined with respect to the regression error. For example, Cavanagh et al. (1995), Campbell and Yogo (2006), Janson and Moreira (2006), Elliott et al. (2015) study parametric models with an NI covariate. The aforementioned papers propose test statistics with limit distributions free of nuisance near-to-unity parameters. This is achieved mainly[1] via conservative inferential methods, e.g. Bonferroni methods or by considering test statistics averaged over a prespecified range for the nuisance parameter space -for a review see Mikusheva (2007) and Phillips (2014, 2015). Although these procedures provide valid inference under local deviations from a unit root, their emphasis is on NI models and may not be valid under large deviations from unity (see Phillips, 2014). Further, their implementation is more involved than that of conventional tests based on studentized regression estimators (i.e., t-/F-tests). This is due to the fact that the related test statistics can be more complex, but more importantly because limit distributions are not conventional (e.g. $\mathbf{N}(0,1)$, χ^2). Therefore, critical values are not readily available from commonly used statistical tables. The implementation of these methods becomes even more difficult in situations where the dimensionality of the nuisance parameter space increases, e.g. when the model involves multiple near unit roots and/or memory parameters, tail parameters (heavy tailed data), TVPs, different types of nonlinearities in the regression function, etc.

The possibility of fractional predictors has received little attention in the literature on robust predictive regressions, despite substantial related work on fractional cointegrated systems, e.g., Robinson and Hualde (2003), Christensen and Nielsen (2006), Hualde and Robinson (2010), and Andersen and Varneskov (2021). The specifications considered in the aforementioned studies are in general structural (i.e. covariates may not be predetermined) and in some cases (e.g., Hualde and Robinson, 2010; Andersen and Varneskov, 2021) both stationary and nonstationary long memory are allowed. These methods are mainly semi-parametric (spectral OLS) with respect to the short memory components of the system and may attain sub-OLS[2] convergence rates due to bandwidth parameters. Regression estimators have mixed Gaussian or Gaussian limit distributions, and therefore inference is conventional in this framework. However, preliminary memory estimators are required, which makes implementation somewhat more involved. Further, although these models are quite general, nonlinearities and nearly integrated arrays are ruled out. For instance, similarly to FMLS (e.g. Phillips, 1995), the spectral LS method of Robinson and Hualde (2003) relies on (fractionally) differencing the data. It is well known that this approach may result in severe size distortions when there are near-to-unity parameters.

The relatively recent work of Phillips and Magdalinos (2009, PM hereafter) provides an alternative approach to inference in regressions with possible nonstationary covariates. This

[1]Janson and Moreira (2006) consider conditional inference rather than conservative tests.
[2]By sub-OLS we mean the OLS rate less an arbitrary slow regularly varying rate.

study proposes instrumentation based on certain linear filtering of the regressors. The resultant IVX instruments exhibit weaker signals than those of the covariates and, as a result, induce mixed normal limit distributions in situations where independent variables are unit roots or NI arrays. IVX instrumentation yields asymptotically vanishing endogeneity, and this is sufficient for a martingale CLT to operate. Hence, contrary to the OLS estimator, in the presence of nonstationary data the IVX estimator has mixed Gaussian limit distribution and studentized IVX estimators have $\mathbf{N}(0,1)$ (t-tests) or χ^2 (F-tests) limit distributions. Conventional and nuisance parameter free inference is achieved for a wide range of persistence in the data at the expense of a slight reduction in the convergence rate. In particular, the IVX estimator attains a sub-OLS convergence rate. PM consider multivariate regressions with mildly and nearly integrated data. The subsequent work of Kostakis, Magdalinos and Stamotogiannis (2015; KMS) extends PM to stationary short memory regressors, and also provides finite sample improvement methods relating to intercept demeaning. For further work on the IVX method, see, e.g. Yang et al. (2020), Demetrescu et al. (2022), Magdalinos (2022), Magdalinos and Petrova (2022) and the references therein.

S2 CTLS Inference in Multi-Covariate Models

In this section we extend CTLS based inference to multi-covariate models as per (4) in the main paper. Recall that the specification of the aforementioned equation is as follows.

$$y_k = \mu + \boldsymbol{\beta}'\mathbf{f}(\mathbf{x}_{k-1}) + e_k,$$

where the covariate and the slope parameter $\boldsymbol{\beta}$ are p-dimensional. As in (5) of the main paper, we may rewrite the model above as

$$y_k = \boldsymbol{\theta}'\mathbf{F}(\mathbf{x}_{k-1}) + e_k, \tag{S2.1}$$

where $\mathbf{F}(\mathbf{x}_{k-1}) = [1, \mathbf{f}(\mathbf{x}_{k-1})']'$ and $\boldsymbol{\theta} = [\mu, \boldsymbol{\beta}']'$. Define $\mathbf{f}_{k-1} = \mathbf{f}(\mathbf{x}_{k-1}), \mathbf{F}_{k-1} = \mathbf{F}(\mathbf{x}_{k-1})$ and

$$\left\{\mathcal{H}_n,\ \hat{\mathcal{V}}_n,\ \mathcal{A}_n\right\} := \left\{\sum_{k=1}^{n} \mathbf{f}_{k-1}\bar{\mathbf{f}}'_{k-1}K_{kn},\ \sum_{k=1}^{n} \check{e}_k^2\mathbf{F}_{k-1}\mathbf{F}'_{k-1}K_{kn}^2,\ \left[-\bar{\mathbf{f}},\ I_p\right]\right\}, \tag{S2.2}$$

where $\check{e}_k := y_k - \hat{\boldsymbol{\theta}}'_{LS}\mathbf{F}_{k-1}$ are the OLS residuals, I_p is the p-dimensional identity matrix, $\overline{\mathbf{f}}_{k-1} = \mathbf{f}_{k-1} - \overline{\mathbf{f}}$, and $\overline{\mathbf{f}} = \sum_{k=2}^{n} \mathbf{f}_{k-1} K_{kn} / \sum_{k=2}^{n} K_{kn}$. For the single restriction hypothesis

$$H_0 : \beta_i = \eta \in \mathbb{R}, \quad i = 1, ..., p, \tag{S2.3}$$

we have the following general formulation of the CTLS t-statistic

$$\hat{T}_i = \frac{\hat{\beta}_i - \eta}{\sqrt{\left[\mathcal{H}_n^{-1}\mathcal{A}_n\hat{\mathcal{V}}_n\mathcal{A}_n'\mathcal{H}_n^{-1}\right]_{ii}}},$$

where recall $[.]_{ii}$ stands for the i^{th} diagonal element of some matrix. We also consider multiple restrictions of the form

$$H_0 : R\boldsymbol{\beta} = \boldsymbol{\eta}, \tag{S2.4}$$

where R is a $q \times p$ $(q \leq p)$ matrix and $\boldsymbol{\eta}$ a predetermined q-dimensional vector. For the latter type of restrictions we consider the CTLS F-statistic

$$\hat{F} = \left[R\hat{\boldsymbol{\beta}} - \boldsymbol{\eta}\right]' \left[\mathcal{H}_n^{-1}\mathcal{A}_n\hat{\mathcal{V}}_n\mathcal{A}_n'\mathcal{H}_n^{-1}\right]^{-1} \left[R\hat{\boldsymbol{\beta}} - \boldsymbol{\eta}\right].$$

In the presence of nonstationary regressors, the CTLS estimators attains multiple convergence rates due to a variation in the degree of persistence, between various covariates, and nonlinearities arising from the regression model (see Theorem 2 in the main paper). This phenomenon requires matrix normalization for various components in $\hat{F}$. Matrix normalization creates technical difficulties due to non commutability of matrix products -for a discussion see Magdalinos and Phillips (2018). To avoid these technical difficulties, under nonstationarity we assume $q = p$ and $R = I_p$. This is general enough to allow for the joint predictability restrictions $\boldsymbol{\beta} = \mathbf{0}$.

We now state the main result for multi-covariate models under the null hypothesis when the regressors are either stationary or nonstationary as discussed in Theorem 3 of the main paper. Its proof is given in Section S11.

Theorem S1. *Suppose that, in addition to the conditions of Theorem 1 or Theorem 2 in the main paper,* $\sup_{k\geq 1} Eu_k^4 < \infty$. *Under* $H_0 : \beta_i = \eta$, *we have*

$$\hat{T}_i \to_d \mathbf{N}(0, 1). \tag{S2.5}$$

Furthermore, under $H_0 : R\boldsymbol{\beta} = \boldsymbol{\eta}$

$$\hat{F} \to_d \chi_q^2. \tag{S2.6}$$

As explained in Remark 8 of the main paper, the requirement $\sup_{k\geq 1} Eu_k^4 < \infty$ can be dispensed with when the regression errors are conditionally homoscedastic, i.e. $\sigma_k^2 = \sigma^2$ for all k.

S3 Proofs of Lemmas 1 and 2

S3.1 Proof of Lemma 1

We only prove (32). The proof of (33) is similar and relatively simple. We shall first assume that there exists an $A > 0$ such that $K(x) = 0$, if $|x| \geq A$ and $K(x)$ is Lipschitz continuous on $\mathbb{R}$. This restriction will be removed later.

Without loss of generality, suppose that $A = 1$. Set $\delta_{1n,j} = [n(\tau_j - 1/c_n)] \vee 1$, $\delta_{2n,j} = [n(\tau_j + 1/c_n)] \vee 1$ and $\delta_{n,j} = [n\tau_j]$. Recall that $\tau_j = j/(l_n + 1)$. Since

$$|c_n(k/n - \tau_j)| < 1 \quad \text{only if} \quad \delta_{1n,j} \leq k \leq \delta_{2n,j}, \quad j = 1, ..., l_n, \tag{S3.1}$$

by letting $R_{1n,j} = \frac{c_n}{n} \sum_{k=\delta_{1n,j}}^{\delta_{2n,j}} v_k \, K\big[c_n(k/n - \tau_j)\big]$ and

$$R_{2n,j} = \frac{c_n}{n} \sum_{k=\delta_{1n,j}}^{\delta_{2n,j}} \big[G\big(X_{n,k}\big) - G\big(X_{n,\delta_{n,j}}\big)\big] \; v_k \; K\big[c_n(k/n - \tau_j)\big],$$

we have

$$\begin{aligned}
S_{n,l_n} &= \frac{1}{l_n} \sum_{j=1}^{l_n} \frac{c_n}{n} \sum_{k=1}^{n} G(X_{n,k}) \, v_k \, K\big[c_n(k/n - \tau_j)\big] \\
&= \frac{1}{l_n} \sum_{j=1}^{l_n} G\big(X_{n,\delta_{n,j}}\big) \frac{c_n}{n} \sum_{k=\delta_{1n,j}}^{\delta_{2n,j}} v_k \; K\big[c_n(k/n - \tau_j)\big] \\
&\quad + \frac{1}{l_n} \sum_{j=1}^{l_n} \frac{c_n}{n} \sum_{k=\delta_{1n,j}}^{\delta_{2n,j}} \big[G\big(X_{n,k}\big) - G\big(X_{n,\delta_{n,j}}\big)\big] \; v_k \; K\big[c_n(k/n - \tau_j)\big] \\
&= \frac{1}{l_n} \sum_{j=1}^{l_n} G\big(X_{n,\delta_{n,j}}\big) \, R_{1n,j} + \frac{1}{l_n} \sum_{j=1}^{l_n} R_{2n,j}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{l_n}\sum_{j=1}^{l_n} G\big(X_{n,\delta_{n,j}}\big)\, A_0 \int K + \frac{1}{l_n}\sum_{j=1}^{l_n} G\big(X_{n,\delta_{n,j}}\big)\, \Big[R_{1n,j} - A_0 \int K\Big] + \frac{1}{l_n}\sum_{j=1}^{l_n} R_{2n,j} \\
&:= \frac{1}{l_n}\sum_{j=1}^{l_n} G\big(X_{n,\delta_{n,j}}\big)\, A_0 \int K + R_{1n} + R_{2n}.
\end{aligned}
$$

Since $\frac{1}{l_n}\sum_{j=1}^{l_n} G\big(X_{n,\delta_{n,j}}\big) = \int_0^1 G(X_{n,[nt]})dt + o_P(1) \to_d \int_0^1 G(\mathcal{X}_t)dt$, it suffices to show that

$$
R_{jn} = o_P(1), \quad j = 1,\ 2. \tag{S3.2}
$$

To prove (S3.2), we start with some preliminaries. Recalling $X_{n,[nt]} \Rightarrow \mathcal{X}_t$ on $D_{\mathbb{R}^q}[0,1]$ and that the limit process $\mathcal{X}_t$ is path continuous, we have $X_{n,[nt]} \Rightarrow \mathcal{X}_t$ on $D_{\mathbb{R}^q}[0,1]$ in the sense of uniform topology. See, for instance, Section 18 of Billingsley (1968). This fact implies that

$$
\limsup_{N\to\infty} \limsup_{n\to\infty} P\left(\max_{1\le k\le n} \|X_{n,k}\| \ge N\right) = 0, \tag{S3.3}
$$

and by the tightness of $\{X_{n,[nt]}\}_{0\le t\le 1}$, for any $\varepsilon > 0$ and $\delta > 0$, there is some $\tilde{\delta} = \tilde{\delta}(\varepsilon,\delta) > 0$ such that

$$
P\left(\sup_{|s-t|\le\tilde{\delta}} \big\|X_{n,[nt]} - X_{n,[ns]}\big\| \ge \delta\right) \le \varepsilon \tag{S3.4}
$$

holds for all sufficiently large n. In view of (S3.4), for any $\delta > 0$, we have

$$
\lim_{n\to\infty} P\left(\max_{1\le j\le l_n} \max_{\delta_{1n,j}\le l\le k\le \delta_{2n,j}} \|X_{n,k} - X_{n,l}\| \ge \delta\right) = 0. \tag{S3.5}
$$

We are now ready to prove (S3.2), starting with $j = 1$.

For any $N > 0$, and any real $\mathbf{x} \in \mathbb{R}^q$ define $G_N(\mathbf{x}) = \xi_N(\mathbf{x})G(\mathbf{x})$ with

$$
\xi_N(\mathbf{x}) := \begin{cases} 1, & \|\mathbf{x}\| \le N, \\ 2 - \|\mathbf{x}\|/N, & N < \|\mathbf{x}\| < 2N, \\ 0, & \|\mathbf{x}\| \ge 2N. \end{cases}
$$

Set

$$
\widetilde{R}_{1n} := \frac{1}{l_n}\sum_{j=1}^{l_n} G_N\left(X_{n,\delta_{n,j}}\right)\left[R_{1n,j} - A_0 \int K\right].
$$

Note that as $n \to \infty$ first and then $N \to \infty$,

$$P(R_{1n} \neq \widetilde{R}_{1n}) \leq P\left(\max_{1\leq k\leq n} \|X_{n,k}\| \geq N\right) \to 0. \tag{S3.6}$$

Moreover,

$$\left\|\widetilde{R}_{1n}\right\| \leq \frac{C_N}{l_n}\sum_{j=1}^{l_n}\left\|R_{1n,j} - A_0\int K\right\|, \tag{S3.7}$$

where $C_N := \sup_{\mathbf{x}\in\mathbb{R}^q}\|G_N(\mathbf{x})\| < \infty$, due to the fact that G_N is continuous with compact support. The result (S3.2) with $j = 1$ will follow if we prove

$$\max_{1\leq j\leq l_n} E\left\|R_{1n,j} - A_0\int K\right\| \to 0, \tag{S3.8}$$

as $n \to \infty$. Indeed, by virtue of (S3.7) and (S3.8), we have $E\left\|\widetilde{R}_{1n}\right\| \to 0$ for each $N \geq 1$. This, together with (S3.6), yields $R_{1n} = o_P(1)$.

Since, as $n \to \infty$,

$$\max_{1\leq j\leq l_n}\left|\frac{c_n}{n}\sum_{k=\delta_{1n,j}}^{\delta_{2n,j}} K\left[c_n(k/n - \tau_j)\right] - \int K\right| \to 0, \tag{S3.9}$$

to prove (S3.8), it suffices to show that $\max_{1\leq j\leq l_n} E\|A_n(\tau_j)\| \to 0$, where

$$A_n(\tau_j) = \frac{c_n}{n}\sum_{k=\delta_{1n,j}}^{\delta_{2n,j}} (v_k - A_0)\ K\left[c_n(k/n - \tau_j)\right].$$

Let $\gamma = \gamma_n$ be integers such that $\gamma \to \infty$ and $\gamma\, c_n/n \to 0$, $T_{1n,j} = [\delta_{1n,j}/\gamma]$ and $T_{2n,j} = [\delta_{2n,j}/\gamma]$. Noting (S3.1), we may write

$$\begin{aligned}
\|A_n(\tau_j)\| &= \left\|\frac{c_n}{n}\sum_{k=\delta_{1n,j}}^{\delta_{2n,j}} (v_k - A_0)\, K\big[c_n(k/n - \tau_j)\big]\right\| \\
&= \frac{c_n}{n}\left\|\sum_{s=T_{1n,j}}^{T_{2n,j}}\ \sum_{k=s\gamma}^{(s+1)\gamma} (v_k - A_0)\, K\left[c_n(k/n - \tau_j)\right]\right\| + o_P(1) \\
&\leq \frac{\gamma c_n}{n}\sum_{s=T_{1n,j}}^{T_{2n,j}} K\left[c_n(s\gamma/n - \tau_j)\right]\frac{1}{\gamma}\left\|\sum_{k=s\gamma}^{(s+1)\gamma}(v_k - A_0)\right\|
\end{aligned}$$

$$
\begin{aligned}
&+\frac{c_n}{n}\sum_{s=T_{1n,j}}^{T_{2n,j}}\sum_{k=s\gamma}^{(s+1)\gamma}\|v_k-A_0\|\;|K\left[c_n(k/n-\tau_j)\right]-K\left[c_n(s\gamma/n-\tau_j)\right]|+o_P(1)\\
:=\;& A_{1n}(\tau_j)+A_{2n}(\tau_j)+o_P(1).
\end{aligned}
$$

Recall that $\sup_{k\geq 1}E\|v_k\|<\infty$ by condition (b). Therefore, it follows from the Lipschitz continuity of $K(x)$ that

$$
EA_{2n}(\tau_j)\leq C\,\frac{\gamma c_n}{n}\,\frac{c_n}{n}\sum_{k=\delta_{1n,j}}^{\delta_{2n,j}}E\|v_k-A_0\|\leq C\,\frac{\gamma c_n}{n}\to 0,
$$

uniformly in $1\leq j\leq l_n$. Similarly, in view of condition (b) and the fact that $\max_{1\leq j\leq l_n}\Big|A_{3n}(\tau_j)-\int K\Big|\to 0$ we have

$$
\max_{1\leq j\leq l_n}EA_{1n}(\tau_j)\leq\max_{\gamma\leq s\leq n-\gamma}E\left\|\frac{1}{\gamma}\sum_{k=s}^{s+\gamma}v_k-A_0\right\|\max_{1\leq j\leq l_n}A_{3n}(\tau_j)\to 0,
$$

where

$$
A_{3n}(\tau_j)=\frac{\gamma c_n}{n}\sum_{s=T_{1n,j}}^{T_{2n,j}}K\left[c_n(s\gamma/n-\tau_j)\right].
$$

In view of the results above, (S3.5) holds true, and this completes the proof of $R_{1n}=o_P(1)$.

Next, we show $R_{2n}=o_P(1)$. Let $\widetilde{R}_{2n}:=\frac{1}{l_n}\sum_{j=1}^{l_n}\widetilde{R}_{2n,j}$, where

$$
\widetilde{R}_{2n,j}=\frac{c_n}{n}\sum_{k=\delta_{1n,j}}^{\delta_{2n,j}}\left[G_N\big(X_{n,k}\big)-G_N\big(X_{n,\delta_{n,j}}\big)\right]\;v_k\;K\big[c_n(k/n-\tau_j)\big].
$$

In view of (S3.6), we have

$$
P(R_{2n}\neq\widetilde{R}_{2n})\leq P\Big(\max_{1\leq k\leq n}\|X_{n,k}\|\geq N\Big)\to 0,
$$

as $n\to\infty$ first and then $N\to\infty$. For the asymptotic negligibility of R_{2n} it suffices to show that $\widetilde{R}_{2n}=o_P(1)$, for each fixed $N\geq 1$.

By definition, $G_N(\mathbf{x})$ is continuous with compact support. Hence, for any $\epsilon>0$ there exists a $\delta_\epsilon>0$ such that $\|G_N(\mathbf{x})-G_N(\mathbf{y})\|\leq\epsilon$ whenever $\|\mathbf{x}-\mathbf{y}\|\leq\delta_\epsilon$. Write

$$
\Omega_{\delta_\epsilon}=\{\omega:\max_{1\leq j\leq l_n}\max_{\delta_{1n,j}\leq l\leq k\leq\delta_{2n,j}}\|X_{n,k}-X_{n,l}\|\leq\delta_\epsilon\}.
$$

By virtue of the facts above and (S3.9), it is readily seen that

$$
\begin{aligned}
& \max_{1\le j\le l_n} E\left[\left\|\widetilde{R}_{2n,j}\right\| I(\Omega_{\delta_\epsilon})\right] \\
\le\ & E\Big\{\max_{1\le j\le l_n}\max_{\delta_{1n,j}\le l\le k\le \delta_{2n,j}} \|G_N(X_{n,k})-G_N(X_{n,l})\|\ \frac{c_n}{n}\sum_{k=\delta_{1n,j}+1}^{\delta_{2n,j}} \|v_k\|\,|K\big[c_n(k/n-\tau_j)\big]|\Big\} \\
\le\ & \epsilon \sup_{k\ge 1} E\,\|v_k\|\ \frac{c_n}{n}\sum_{k=\delta_{1n,j}+1}^{\delta_{2n,j}} K\big[c_n(k/n-\tau_j)\big] \le C_N\epsilon,
\end{aligned}
$$

where C_N is a constant depending only on N. Now, for any $\eta_1>0$ and $\eta_2>0$, let $\epsilon=\eta_1\eta_2$ and n_0 be large enough so that, for all $n\ge n_0$ [recall (S3.5)],

$$
P\left(\max_{1\le j\le l_n}\max_{\delta_{1n,j}\le l\le k\le \delta_{2n,j}} \|X_{n,k}-X_{n,l}\|\ge \delta_\epsilon\right)\le \eta_2.
$$

Hence, for all $n\ge n_0$,

$$
P\left(\left\|\widetilde{R}_{2n}\right\|\ge \eta_1\right)\le P(\bar{\Omega}_{\delta_\epsilon})+\eta_1^{-1}\,\frac{1}{l_n}\sum_{j=1}^{l_n} E\left[\left\|\widetilde{R}_{2n,j}\right\| I(\Omega_{\delta_\epsilon})\right]\ \le\ C_N\,\eta_2,
$$

where $\bar{\Omega}_{\delta_\epsilon}$ denotes the complementary set of Ω_{δ_ϵ} and C_N is a constant depending only on N. This yields $\widetilde{R}_{2n}=o_P(1)$, for each fixed $N\ge 1$, and completes the proof of $R_{2n}=o_P(1)$.

Finally, we remove the restriction on K and then conclude the proof of Lemma 1. If K has a compact support, there exists $A_1>0$ such that $K(x)=0$ holds for all $|x|\ge A_1$. If K is eventually monotonic (without loss of generality, we assume $K\ge 0$), for any $\epsilon>0$, we can also choose a constant $A_{1\epsilon}>0$ such that $K(x)$ is monotonic on $(-\infty,-A_{1\epsilon})$ and $(A_{1\epsilon},\infty)$ and $\int_{|x|>A_{1\epsilon}} K(x)dx<\epsilon$. As a consequence, it follows from $\int K<\infty$ that for any $\epsilon>0$ and $A\ge \max\{A_1,A_{1\epsilon}\}+1$, there exists a $K_{\epsilon,A}(x)$ such that

$$
\int |K-K_{\epsilon,A}|\le 2\epsilon, \tag{S3.10}
$$

where $K_{\epsilon,A}(x)=0$ if $|x|\ge A$ and $K_{\epsilon,A}(x)$ is Lipschitz continuous on $\mathbb{R}$. It has been shown in the first part that, for any $\epsilon>0$ and $A\ge\max\{A_1,A_{1\epsilon}\}+1$,

$$
\begin{aligned}
& \frac{1}{l_n}\sum_{j=1}^{l_n}\frac{c_n}{n}\sum_{k=1}^{n} G(X_{n,k})\,v_k\,K_{\epsilon,A}\left[c_n(k/n-\tau_j)\right] \\
&\qquad = \int_0^1 G(X_{n,[nt]})dtA_0\int K_{\epsilon,A}+o_P(1)\to_d \int_0^1 G(\mathcal{X}_t)dtA_0\int K_{\epsilon,A}.
\end{aligned}
$$

To show (32) it suffices proving that as $n \to \infty$ first and then $\epsilon \to 0$ (implying $A \to \infty$),

$$S_{n,\epsilon} := \frac{1}{l_n} \sum_{j=1}^{l_n} \frac{c_n}{n} \sum_{k=1}^{n} G(X_{n,k})\, v_k\, \widetilde{K}\big[c_n(k/n - \tau_j)\big] = o_P(1), \tag{S3.11}$$

where $\widetilde{K}(x) = K(x) - K_{\epsilon,A}(x)$.

The proof of (S3.11) is similar to that of (S3.2). For any $\epsilon > 0$, set A as in (S3.10). First, note that as in (S3.9),

$$\sup_{1\le j\le l_n} \Big| \frac{c_n}{n} \sum_{k=1}^{n} \big|\widetilde{K}\big[c_n(k/n - \tau_j)\big]\big| I(c_n|k/n - \tau_j| \le A) - \int_{-A}^{A} |\widetilde{K}(x)| dx \Big| \to 0,$$

when $n \to \infty$. Hence, for n sufficiently large,

$$A_{1j} := \frac{c_n}{n} \sum_{k=1}^{n} \big|\widetilde{K}\big[c_n(k/n - \tau_j)\big]\big| I(c_n|k/n - \tau_j| \le A) \le \int |\widetilde{K}(x)| dx + \epsilon \le 3\epsilon,$$

uniformly in $1 \le j \le l_n$. On the other hand, it follows from the monotonicity of $K(x)$ on $(-\infty, -A)$ and (A, ∞) that, whenever n is sufficiently large,

$$\begin{aligned} A_{2j} &:= \frac{c_n}{n} \sum_{k=1}^{n} \big|\widetilde{K}\big[c_n(k/n - \tau_j)\big]\big| I(c_n|k/n - \tau_j| > A) \\ &= \frac{c_n}{n} \sum_{k=1}^{n} K\big[c_n(k/n - \tau_j)\big] I(c_n|k/n - \tau_j| > A) \\ &\le \int_{|x|>A-c_n/n} K(x) dx \le \int_{|x|>\max\{A_1, A_{1\epsilon}\}} K(x) dx < \epsilon, \end{aligned}$$

uniformly in $1 \le j \le l_n$. Using these facts, when n is sufficiently large, we have

$$\frac{1}{l_n} \sum_{j=1}^{l_n} \frac{c_n}{n} \sum_{k=1}^{n} \Big| \widetilde{K}\left[c_n(k/n - \tau_j)\right] \Big| \le \frac{1}{l_n} \sum_{j=1}^{l_n} (A_{1j} + A_{2j}) \le 4\epsilon.$$

Now, for any $\delta > 0$, let

$$S_{n,\epsilon,N} := \frac{1}{l_n} \sum_{j=1}^{l_n} \frac{c_n}{n} \sum_{k=1}^{n} G_N(X_{n,k})\, v_k\, \widetilde{K}\left[c_n(k/n - \tau_j)\right].$$

Using the fact that $G_N(\mathbf{x})$ is uniformly bounded, we have

$$P\left(\|S_{n,\epsilon}\| \ge \delta\right) \le P\left(S_{n,\epsilon} \ne S_{n,\epsilon,N}\right) + P\left(\|S_{n,\epsilon,N}\| \ge \delta\right)$$

$$
\begin{aligned}
&\leq P\left(\max_{1\leq k\leq n}\|X_{n,k}\|\geq N\right)+\delta^{-1}E\,\|S_{n,\epsilon,N}\| \\
&\leq P\left(\max_{1\leq k\leq n}|\,\|X_{n,k}\|\,|\geq N\right)+\delta^{-1}C_N\sup_k E\,\|v_k\| \\
&\qquad\cdot\frac{1}{l_n}\sum_{j=1}^{l_n}\frac{c_n}{n}\sum_{k=1}^{n}\left|\widetilde{K}\left[c_n(k/n-\tau_j)\right]\right| \\
&\leq P\left(\max_{1\leq k\leq n}\|X_{n,k}\|\geq N\right)+C_{1N}\,\epsilon\delta^{-1}\to 0,
\end{aligned}
$$

as $n\to\infty$ first, $N\to\infty$ second and then $\epsilon\to 0$. This proves (S3.11) and hence completes the proof of Lemma 1. □

S3.2 Proof of Lemma 2

We first prove (35) and, without loss of generality, assume $K\geq 0$. Using similar arguments as in the proof of (S3.2) or (S3.11), it suffices to show that, as $n\to\infty$,

$$
I_n:=\frac{c_n}{n}\sum_{k=1}^{n}\frac{1}{l_n}\sum_{1\leq i<j\leq l_n}K\big[c_n(k/n-\tau_i)\big]K\big[c_n(k/n-\tau_j)\big]\to 0.
$$

Set $\eta_{n,i,j}:=\frac{1}{2}n(\tau_i+\tau_j)$. Note that $c_n(k/n-\tau_i)\geq c_n(j-i)/(2(l_n+1))$, if $k\geq\eta_{n,i,j}$, and $|c_n(k/n-\tau_j)|\geq c_n(j-i)/(2(l_n+1))$, if $k\leq\eta_{n,i,j}$. In view of the fact that $K(x)\leq 1/|x|$ for x sufficiently large[3], we have

$$
\begin{aligned}
I_n &= \frac{1}{l_n}\sum_{1\leq i<j\leq l_n}\frac{c_n}{n}\sum_{k=1}^{n}K\big[c_n(k/n-\tau_i)\big]K\big[c_n(k/n-\tau_j)\big] \\
&\leq \frac{C}{l_n}\sum_{1\leq i<j\leq l_n}\frac{l_n+1}{c_n(j-i)}\,\frac{c_n}{n}\sum_{k=1}^{n}\big(K\big[c_n(k/n-\tau_i)\big]+K\big[c_n(k/n-\tau_j)\big]\big) \\
&\leq \frac{C}{c_n}\sum_{1\leq i<j\leq l_n}\frac{1}{j-i}\leq C\,l_n\log l_n/c_n\to 0,
\end{aligned}
$$

as required.

The proof of (34) is similar to that of (35) and, hence, the details are omitted. The result of (36) follows easily from (34) and (35). Finally, (37) follows from similar arguments as those

[3] Since $\int K<\infty$ and $K\geq 0$ is eventually monotonic, we have that K is decreasing on (A_1,∞) for some $A_1>0$, and

$$
xK(x)/2\leq\int_{x/2}^{x}K(t)dt\to 0,\quad x\to+\infty.
$$

Similarly $\lim_{x\to-\infty}xK(x)=0$. Hence, $K(x)\leq 1/|x|$ when x is sufficiently large.

used in the proof of (S3.2) and the fact that as $n \to \infty$,

$$
\begin{aligned}
& \big(\frac{c_n}{n}\big)^2 \sum_{k=1}^{n} \Big(\frac{1}{\sqrt{l_n}} \sum_{j=1}^{l_n} K\big[c_n(k/n-\tau_j)\big]\Big)^4 \\
\le\ & 2\,\big(\frac{c_n}{n}\big)^2 \sum_{k=1}^{n} \Big(\frac{1}{\sqrt{l_n}} \sum_{j=1}^{l_n} K^2\big[c_n(k/n-\tau_j)\big]\Big)^2 \\
& +8\,\big(\frac{c_n}{n}\big)^2 \sum_{k=1}^{n} \Big(\frac{1}{l_n} \sum_{1\le i<j\le l_n} K\big[c_n(k/n-\tau_i)\big] K\big[c_n(k/n-\tau_j)\big]\Big)^2 \\
\le\ & 2\,C^2\big(\frac{c_n}{n}\big)^2 \sum_{k=1}^{n} \Big(\frac{1}{\sqrt{l_n}} \sum_{j=1}^{l_n} K\big[c_n(k/n-\tau_j)\big]\Big)^2 + 8I_n^2 \to 0,
\end{aligned}
$$

due to (35) and (36). □

S4 Proofs of Theorem 5

We only consider M_{1n,l_n}, i.e., (26), since the limit result for $S_{1n,l_n}^{(m)}$ given in (25) follows easily from Lemma 1 with $G(x) \equiv I_{p+1}$ and $v_k = \mathbf{F}(\mathbf{x}_{k-1})\mathbf{F}(\mathbf{x}_{k-1})'\sigma_k^m$.

Set $Q_{k,n} := \sqrt{\frac{c_n}{n}}\alpha'\mathbf{F}(\mathbf{x}_{k-1})\sigma_k \frac{1}{\sqrt{l_n}} \sum_{j=1}^{l_n} K\left[c_n(k/n-\tau_j)\right]$, where $\alpha \in \mathbb{R}^{p+1}$. Using (35) in Lemma 2 with $G(x) \equiv 1$ and $v_k \equiv [\alpha'\mathbf{F}(\mathbf{x}_{k-1})\sigma_k]^2$, we have

$$
\begin{aligned}
\sum_{k=1}^{n} Q_{k,n}^2 &= \frac{c_n}{n} \sum_{k=1}^{n} \big[\alpha'\mathbf{F}(\mathbf{x}_{k-1})\sigma_k\big]^2 \frac{1}{l_n} \sum_{j=1}^{l_n} K^2\left[c_n(k/n-\tau_j)\right] + o_P(1) \\
&= E\big[\alpha'\mathbf{F}(\mathbf{x}_1)\sigma_2\big]^2 \int K^2 + o_P(1),
\end{aligned}
\tag{S4.1}
$$

where the second equation follows from Lemma 1, with $K(x)$ replaced by $K^2(x)$, and $A_0 = E\left[\alpha'\mathbf{F}(\mathbf{x}_1)\sigma_2\right]^2$. In view of (S4.1), it follows from the classical martingale limit theorem (e.g., Hall and Heyde (1980), Theorem 3.2 or Wang (2014), Theorem 2.1) that to prove (26), it suffices to show that

$$
\max_{1\le k\le n} |Q_{k,n}| = o_P(1).
$$

Note that for any $A > 0$,

$$
\max_{1\le k\le n} |Q_{k,n}| \ \le\ \left\{\sum_{k=1}^{n} Q_{k,n}^2 I\left\{\|\mathbf{F}(\mathbf{x}_{k-1})\sigma_k\| > A\right\}\right\}^{1/2} + \left\{\sum_{k=1}^{n} Q_{k,n}^4 I\left\{\|\mathbf{F}(\mathbf{x}_{k-1})\sigma_k\| \le A\right\}\right\}^{1/4}
$$

$$=: \quad II_{1n}(A)^{1/2} + II_{2n}(A)^{1/4}.$$

Similar arguments used in (S4.1) show that the first term above is

$$\begin{aligned} II_{1n}(A) &\leq \|\alpha\|^2 \frac{c_n}{n} \sum_{k=1}^{n} \|\mathbf{F}(\mathbf{x}_{k-1})\sigma_k\|^2 I\{\|\mathbf{F}(\mathbf{x}_{k-1})\sigma_k\| > A\} \frac{1}{l_n} \sum_{j=1}^{l_n} K^2 [c_n(k/n - \tau_j)] + o_P(1) \\ &= \|\alpha\|^2 E \|\mathbf{F}(\mathbf{x}_1)\sigma_2\|^2 I\{\|\mathbf{F}(\mathbf{x}_1)\sigma_2\| > A\} \int K^2 + o_P(1) = o_P(1), \end{aligned}$$

where we take $n \to \infty$ first and then $A \to \infty$. On the other hand, using (37) in Lemma 2 with $G(x) \equiv 1$ and $v_k \equiv 1$, the second term

$$II_{2n}(A) \leq \|\alpha\|^4 A^4 \left(\frac{c_n}{n}\right)^2 \sum_{k=1}^{n} \left\{ \frac{1}{\sqrt{l_n}} \sum_{j=1}^{l_n} K [c_n(k/n - \tau_j)] \right\}^4 = o_P(1),$$

for each $A \geq 1$, as $n \to \infty$. Combining these facts, we establish (26). The proof of Theorem 5 is now complete. □

S5 Proof of Theorem 6

We only consider M_{2n,l_n}, i.e., (28). The result for $\left[S^{(m)}_{2n,l_n}, S^{(m)}_{3n,l_n}\right]$ given in (27) follows directly from Lemma 1 with $G(X_{n,k}) \equiv \mathbf{F}(X_{n,k}) \mathbf{F}(X_{n,k})'$ or $G(X_{n,k}) \equiv \mathbf{Q}(X_{n,k})$, and $v_k \equiv \sigma_k^m$, $m = 0$ or 2.

Set $Q_{k,n} := \sqrt{\frac{c_n}{n}} \alpha' \mathbf{F}(X_{n,k-1}) \sigma_k \frac{1}{\sqrt{l_n}} \sum_{j=1}^{l_n} K[c_n(k/n - \tau_j)]$, where $\alpha \in \mathbb{R}^{p+1}$. Noting that $\int_0^1 \mathbf{F}(X_{n,[nt]}) dt$ is a continuous functional of $X_{n,[nt]}$, the limit result of (28), jointly with (27), will follow if we prove that, for any $\alpha \in \mathbb{R}^{p+1}$.

$$\left[X_{n,[nt]}, \sum_{k=1}^{n} Q_{k,n} u_k \right] \Rightarrow \left[\mathcal{X}_t, \ \mathbf{MN}\left(0, E(\sigma_1^2) \int_0^1 [\alpha' \mathbf{F}(\mathcal{X}_t)]^2 dt \int K^2 \right) \right] \tag{S5.1}$$

on $D_{\mathbb{R}^p \times \mathbb{R}}[0, 1]$. First, note that by using (35) with $v_k \equiv \sigma_k^2$ and $G(.) \equiv \alpha' \mathbf{F}(.)$ first, and then (32),

$$\begin{aligned} \sum_{k=1}^{n} Q_{k,n}^2 &= \frac{c_n}{n} \sum_{k=1}^{n} [\alpha' \mathbf{F}(X_{n,k-1})]^2 \sigma_k^2 \frac{1}{l_n} \sum_{j=1}^{l_n} K^2 [c_n(k/n - \tau_j)] + o_P(1) \\ &= E(\sigma_1^2) \int_0^1 [\alpha' \mathbf{F}(X_{n,[nt]})]^2 dt \int K^2 + o_P(1). \end{aligned} \tag{S5.2}$$

It follows from **A3**(a) and the continuous mapping theorem that

$$\left[\frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor nt\rfloor}\xi_k,\ \frac{1}{\sqrt{n}}\sum_{k=1}^{\lfloor nt\rfloor}\xi_{-k},\ X_{n,[nt]},\ \sum_{k=1}^{n}Q_{k,n}^2\right]$$
$$\Rightarrow\ \left[B_{1t}, B_{2t}, \mathcal{X}_t, E\left(\sigma_1^2\right)\int_0^1\left[\alpha'\mathbf{F}(\mathcal{X}_t)\right]^2 dt\int K^2\right],$$

on $D_{\mathbb{R}^p\times\mathbb{R}^p\times\mathbb{R}^p\times\mathbb{R}}[0,1]$. Recall **A1** and that $Q_{k,n}$ is a functional of $\xi_k, \xi_{k-1}, ...$ By Theorem 2.1 of Wang (2014) or Theorem 3.14 of Wang (2015), the limit result of (S5.1) will follow if we prove

$$\max_{1\le k\le n}|Q_{k,n}| = o_P(1), \tag{S5.3}$$

and

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}|Q_{k,n}| = o_P(1), \tag{S5.4}$$

which is what we set out to do next. In view of the continuity of $||\mathbf{F}(.)||^4$, it follows from (37) with $v_k \equiv \sigma_k^4$ that

$$\begin{aligned}\left[\max_{1\le k\le n}|Q_{k,n}|\right]^4 &\le \sum_{k=1}^{n}Q_{k,n}^4\\ &\le \|\alpha\|^4\left(\frac{c_n}{n}\right)^2\sum_{k=1}^{n}\|\mathbf{F}(X_{n,k-1})\|^4\sigma_k^4\left(\frac{1}{\sqrt{l_n}}\sum_{j=1}^{l_n}K\left[c_n(k/n-\tau_j)\right]\right)^4 = o_P(1),\end{aligned}$$

yielding (S5.3). Similarly, using the fact that $l_n/c_n \to 0$ and (34) in Lemma 2, we have

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{k=1}^{n}|Q_{k,n}| &\le \|\alpha\|\frac{1}{\sqrt{n}}\sqrt{\frac{c_n}{n}}\sum_{k=1}^{n}\|\mathbf{F}(X_{n,k-1})\|\,|\sigma_k|\frac{1}{\sqrt{l_n}}\sum_{j=1}^{l_n}K\left[c_n(k/n-\tau_j)\right]\\ &= \|\alpha\|\sqrt{\frac{l_n}{c_n}}\frac{c_n}{n}\sum_{k=1}^{n}\|\mathbf{F}(X_{n,k-1})\|\,|\sigma_k|\frac{1}{l_n}\sum_{j=1}^{l_n}K\left[c_n(k/n-\tau_j)\right]\\ &= O_P\left(\sqrt{\frac{l_n}{c_n}}\right) = o_P(1),\end{aligned}$$

which shows (S5.4). The proof of Theorem 6 is complete. □

S6 Proof of Theorem 7

We start with the following lemma.

Lemma S1. *Suppose that:*

(a) *for each* $i = 1, ..., q$, $g_i : \mathbb{R} \to \mathbb{R}$ *is an AHF function with limit homogeneous function* H_{g_i} *and asymptotic order* π_{g_i}*;*

(b) $\mathbf{x}_k = [x_{k,1}, ..., x_{k,q}]'$, $\mathcal{X}_t$ *is a* $\mathbb{R}^q$*-valued continuous process, and in* $D_{\mathbb{R}^q}[0,1]$ *there are deterministic sequences* $d_{in} \to \infty$, $i = 1, ..., q$ *such that* $X_{n,[nt]} \Rightarrow \mathcal{X}_t$, *where*

$$X_{n,k} = diag\{d_{1n}, ..., d_{qn}\}^{-1}\mathbf{x}_k;$$

(c) *either* $e_k^2 = \sigma_k^m$, *where* $m = 0, 2$, *and Assumptions* $\mathbf{A3}(b)$ *and* $\mathbf{A4}^*$ *hold, or*

$e_k = \sigma_k u_k$ *with* $\sup_{k\geq 1} Eu_k^4 < \infty$, *and Assumptions* $\mathbf{A1}(c)$, $\mathbf{A3}(b)$ *and* $\mathbf{A4}^*$ *hold;*

(d) $h : \mathbb{R}^q \to \mathbb{R}$ *is a continuous function, and there exist* $c_0 > 0, \alpha > 0$ *and* $\nu \geq 0$ *so that for every* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^q$,

$$|h(\mathbf{x} + \mathbf{y}) - h(\mathbf{x})| \leq c_0 \|\mathbf{y}\|^\alpha (1 + \|\mathbf{x}\| + \|\mathbf{y}\|)^\nu. \tag{S6.1}$$

Then as $n \to \infty$,

$$\frac{c_n}{nl_n}\sum_{k=1}^n e_k^2\, K_{kn} h(\widetilde{\mathbf{x}}_{n,k-1}) = \frac{c_n}{nl_n}\sum_{k=1}^n e_k^2\, K_{kn} h(\widetilde{\mathbf{y}}_{n,k-1}) + o_P(1),$$

where

$$\widetilde{\mathbf{x}}_{nk} := \left[\frac{g_1(x_{k,1})}{\pi_{g_1}(d_{1n})}, \cdots, \frac{g_q(x_{k,q})}{\pi_{g_q}(d_{qn})}\right] \quad \textit{and} \quad \widetilde{\mathbf{y}}_{nk} := \left[H_{g_1}\left(\frac{x_{k,1}}{d_{1n}}\right), \cdots, H_{g_q}\left(\frac{x_{k,q}}{d_{qn}}\right)\right].$$

Proof. We only prove Lemma S1 with $e_k = \sigma_k u_k$. The proof for $e_k^2 = \sigma_k^m$ is similar but simpler. Let $\tilde{\mathbf{z}}_{nk} = \tilde{\mathbf{x}}_{nk} - \tilde{\mathbf{y}}_{nk}$. It follows from the definition AHF that

$$\|\tilde{\mathbf{z}}_{nk}\| \leq a_n \sum_{i=1}^q (1 + |x_{k,i}/d_{in}|^{\delta_{g_i}}) \leq 2a_n q(1 + \|X_{n,k}\|)^\delta,$$

where $a_n = \max_{1\le j\le q} \frac{a_{g_j}(d_{jn})}{\pi_{g_j}(d_{jn})} \to 0$ and $\delta = \max_{1\le j\le q} g_j$. Observe that there is a continuous function $h_0 : \mathbb{R}^q \to \mathbb{R}^q$ such that $\tilde{\mathbf{y}}_{nk} = h_0(X_{n,k})$. Therefore, by the condition (d), we have

$$\frac{c_n}{nl_n}\sum_{k=1}^n e_k^2\, K_{kn}|h(\tilde{\mathbf{x}}_{n,k-1}) - h(\tilde{\mathbf{y}}_{n,k-1})| = o_P(1)\cdot \frac{c_n}{nl_n}\sum_{k=1}^n e_k^2\, K_{kn} h_1(X_{n,k-1}), \tag{S6.2}$$

where

$$h_1(\mathbf{x}) = (1+\|\mathbf{x}\|)^{\alpha\delta}[(1+\|\mathbf{x}\|)^{\delta} + \|h_0(\mathbf{x})\|]^{\nu}$$

is continuous. Recall that $E\left(u_k^2|\mathcal{F}_{k-1}\right) = 1$, and σ_k is $\mathcal{F}_{k-1}$ measurable and stationary. It is readily seen from **A3**(b) and $\sup_{k\ge 1} Eu_k^4 < \infty$ that

$$\begin{aligned}\max_{m\le j\le n-m} E\left|\frac{1}{m}\sum_{k=j+1}^{j+m} e_k^2 - E\left(\sigma_1^2\right)\right| &= E\left|\frac{1}{m}\sum_{k=1}^{m} e_k^2 - E\left(\sigma_1^2\right)\right| \\ &\le E\left|\frac{1}{m}\sum_{k=1}^{m}\sigma_k^2\left[u_k^2 - E\left(u_k^2|\mathcal{F}_{k-1}\right)\right]\right| + E\left|\frac{1}{m}\sum_{k=1}^{m}\left[\sigma_k^2 - E\left(\sigma_k^2\right)\right]\right| \to 0,\end{aligned}$$

for any $0 < m = m_n \to \infty$. Hence, condition (b) of Lemma 1 is satisfied with $v_k \equiv e_k^2$ and $A_0 \equiv E\left(\sigma_1^2\right)$. The desired result follows from (S6.2) and (32) in Lemma 1. □

We now turn to the proof of Theorem 7. It suffices to prove the $o_P(1)$ approximations in (30)-(31). The weak convergence results in the aforementioned equations are a direct consequence of Theorem 6. Further, the proof of (29) is identical to that for (30).

The proof of the $o_P(1)$ approximation in (30) is simple. Indeed, by recalling that $K_{kn} = \sum_{j=1}^{l_n} K\left[c_n(k/n - \tau_j)\right]$ it follows from the condition (b) in Theorem 7 that

$$\begin{aligned}&\frac{c_n}{nl_n}\sum_{k=1}^n \mathcal{L}_n^{-1}\mathbf{F}(\mathbf{x}_{k-1})\mathbf{F}(\mathbf{x}_{k-1})'\mathcal{L}_n^{-1}\,\sigma_k^m K_{kn} \\ &\qquad =: \frac{c_n}{nl_n}\sum_{k=1}^n H_{\mathbf{F}}(X_{n,k-1})H_{\mathbf{F}}(X_{n,k-1})'\sigma_k^m\, K_{kn} + \Delta_{1n},\end{aligned}$$

where Δ_{1n} is a $(p+1)\times(p+1)$ matrix that is determined by the definition AHF. It follows from Lemma S1 with $e_k^2 \equiv \sigma_k^m$ that $\Delta_{1n} = o_P(1)$. Therefore, (30) follows directly from Theorem 6.

We next prove (31). We write

$$\sqrt{\frac{c_n}{nl_n}}\sum_{k=1}^n \mathcal{L}_n^{-1}\mathbf{F}(\mathbf{x}_{k-1})\sigma_k u_k\; K_{kn} = \sqrt{\frac{c_n}{nl_n}}\sum_{k=1}^n H_{\mathbf{F}}(X_{n,k-1})\sigma_k\, K_{kn}u_k + \Delta_{2n},$$

where Δ_{2n} is a $(p+1)$-dimensional vector. In particular, the first element of Δ_{2n} is zero, and the $j+1$ element

$$[\Delta_{2n}]_j = \sqrt{\frac{c_n}{nl_n}} \pi_{f_j}^{-1}(d_{jn}) \sum_{k=1}^{n} R_{f_j}(d_{jn}; x_{k-1,j}) \sigma_k K_{kn} u_k, \ , j = 1, ..., p,$$

with R_{f_j} given in the definition AHF. For $A > 0$, set

$$[\Delta_{2n}]_j (A) = \sqrt{\frac{c_n}{nl_n}} \pi_{f_j}(d_{jn})^{-1} \sum_{k=1}^{n} R_{f_j}(d_{jn}; x_{k-1,j}) I \{|x_{k-1,j}/d_{jn}| \le A\} \sigma_k K_{kn} u_k.$$

Note that as $n \to \infty$ first and then $A \to \infty$

$$P\left([\Delta_{2n}]_j \ne [\Delta_{2n}]_j (A)\right) \le P\left(\max_{1 \le k \le n} |x_{k-1,j}/d_{jn}| > A\right) \to 0. \quad \text{(S6.3)}$$

For any $\epsilon > 0$ and $A > 0$, we have

$$P\left(\left|[\Delta_{2n}]_j\right| \ge \epsilon\right) \le P\left([\Delta_{2n}]_j \ne [\Delta_{2n}]_j (A)\right) + \epsilon^{-2} E\left[[\Delta_{2n}]_j (A)\right]^2. \quad \text{(S6.4)}$$

Furthermore, for any $A > 0$, as $n \to \infty$ we have

$$\begin{aligned} E\left[[\Delta_{2n}]_j (A)\right]^2 &= \frac{c_n}{nl_n} \pi_{f_j}^{-2}(d_{jn}) \sum_{k=1}^{n} E\Big(R_{f_j}(d_{jn}; x_{k-1,j})^2 I \{|x_{k-1,j}/d_{jn}| \le A\} \sigma_k^2\Big) K_{kn}^2 \\ &\le \left[\frac{a_{f_j}(d_{jn})}{\pi_{f_j}(d_{jn})}\right]^2 P_{f_i}^2(A) \frac{c_n}{nl_n} \sum_{k=1}^{n} E(\sigma_k^2) K_{kn}^2 \to 0, \end{aligned} \quad \text{(S6.5)}$$

where P_{f_i} is given in the definition AHF, and we have used (36) of Lemma 2 with $G(x) \equiv 1$ and $v_k \equiv E(\sigma_k^2)$. In view of (S6.3)-(S6.5), as $n \to \infty$

$$P\left(\left|[\Delta_{2n}]_j\right| \ge \epsilon\right) \to 0,$$

for all $j = 1, ..., p+1$. The proof of Theorem 7 is now complete. □

S7 Proof of Theorem 1

The CTLS estimator for $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = \left[\sum_{k=1}^{n} \mathbf{Z}_{kn} \bar{\mathbf{f}}_{k-1}'\right]^{-1} \sum_{k=1}^{n} \mathbf{Z}_{kn} \bar{y}_k$$

$$
\begin{aligned}
&= \left[\sum_{k=1}^{n} \mathbf{Z}_{kn} \overline{\mathbf{f}}_{k-1}'\right]^{-1} \sum_{k=1}^{n} \mathbf{Z}_{kn} \left[\overline{\mathbf{f}}_{k-1}' \boldsymbol{\beta} + e_k - \frac{\sum_{s=1}^{n} e_s K_{sn}}{\sum_{s=1}^{n} K_{sn}}\right] \\
&= \boldsymbol{\beta} + \left[\sum_{k=1}^{n} \mathbf{Z}_{kn} \overline{\mathbf{f}}_{k-1}'\right]^{-1} \cdot \sum_{k=1}^{n} \mathbf{Z}_{kn} \left[e_k - \frac{\sum_{s=1}^{n} e_s K_{sn}}{\sum_{s=1}^{n} K_{sn}}\right],
\end{aligned}
$$

which gives

$$
\sqrt{\frac{nl_n}{c_n}} \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) = \left[\frac{c_n}{nl_n} \sum_{k=1}^{n} \mathbf{Z}_{kn} \overline{\mathbf{f}}_{k-1}'\right]^{-1} \cdot \sqrt{\frac{c_n}{nl_n}} \sum_{k=1}^{n} \mathbf{Z}_{kn} \left[e_k - \frac{\sum_{s=1}^{n} e_s K_{sn}}{\sum_{s=1}^{n} K_{sn}}\right]. \tag{S7.1}
$$

It follows from Theorem 5 that

$$
\begin{aligned}
\frac{c_n}{nl_n} \sum_{k=1}^{n} \mathbf{Z}_{kn} \overline{\mathbf{f}}_{k-1}' &= \frac{c_n}{nl_n} \sum_{k=1}^{n} \mathbf{Z}_{kn} \mathbf{f}_{k-1}' - \frac{c_n}{nl_n} \sum_{k=1}^{n} \mathbf{Z}_{kn} \frac{\frac{c_n}{nl_n} \sum_{s=1}^{n} \mathbf{f}_{s-1}' K_{sn}}{\frac{c_n}{nl_n} \sum_{s=1}^{n} K_{sn}} \\
&= \frac{c_n}{nl_n} \sum_{k=1}^{n} \mathbf{f}_{k-1} \mathbf{f}_{k-1}' K_{kn} - \frac{c_n}{nl_n} \sum_{k=1}^{n} \mathbf{f}_{k-1} K_{kn} \frac{\frac{c_n}{nl_n} \sum_{s=1}^{n} \mathbf{f}_{s-1}' K_{sn}}{\frac{c_n}{nl_n} \sum_{s=1}^{n} K_{kn}} \\
&\to_P \int K \cdot \left[E\left\{\mathbf{f}(\mathbf{x}_1)\mathbf{f}(\mathbf{x}_1)'\right\} - E\left\{\mathbf{f}(\mathbf{x}_1)\right\} E\left\{\mathbf{f}(\mathbf{x}_1)'\right\}\right] \\
&= \int K \cdot \Phi_1.
\end{aligned} \tag{S7.2}
$$

Moreover,

$$
\begin{aligned}
&\sqrt{\frac{c_n}{nl_n}} \sum_{k=1}^{n} \mathbf{Z}_{kn} \left[e_k - \frac{\sum_{s=1}^{n} e_s K_{sn}}{\sum_{s=1}^{n} K_{sn}}\right] \\
&\quad = \left[-\frac{1}{\frac{c_n}{nl_n} \sum_{k=1}^{n} K_{kn}} \frac{c_n}{nl_n} \sum_{k=1}^{n} \mathbf{Z}_{kn},\ I_p\right] \cdot \sqrt{\frac{c_n}{nl_n}} \sum_{k=1}^{n} K_{kn} \begin{bmatrix} 1 \\ \mathbf{f}_{k-1} \end{bmatrix} e_k \\
&\quad =: \mathbf{A}_n \cdot \sqrt{\frac{c_n}{nl_n}} \sum_{k=1}^{n} K_{kn} \mathbf{F}_{k-1} e_k.
\end{aligned}
$$

Again by Theorem 5 we get

$$
\mathbf{A}_n \to_P \mathbf{A} := \left[-E\mathbf{f}(\mathbf{x}_1),\ I_p\right],
$$

and

$$
\sqrt{\frac{c_n}{nl_n}} \sum_{k=2}^{n} K_{kn} \mathbf{F}_{k-1} e_k \to_d \mathbf{N}\left(\mathbf{0}, \int K^2 \cdot E\left(\sigma_2^2 \mathbf{F}(\mathbf{x}_1)\mathbf{F}(\mathbf{x}_1)'\right)\right).
$$

In view of these facts, we have

$$\sqrt{\frac{c_n}{nl_n}}\sum_{k=1}^{n}\mathbf{Z}_{kn}\left[e_k-\frac{\sum_{s=1}^{n}e_sK_{sn}}{\sum_{s=1}^{n}K_{sn}}\right]=\mathbf{A}_n\cdot\sqrt{\frac{c_n}{nl_n}}\sum_{k=1}^{n}K_{kn}\mathbf{F}_{k-1}e_k$$
$$\to_d \mathbf{N}\left(0,\int K^2\cdot\mathbf{A}E(\sigma_2^2\mathbf{F}(\mathbf{x}_1)\mathbf{F}(\mathbf{x}_1)')\mathbf{A}'\right)=_d \mathbf{N}\left(0,\int K^2\cdot\Phi_0\right), \quad \text{(S7.3)}$$

where we have used the fact that $\mathbf{A}E(\sigma_2^2\mathbf{F}(\mathbf{x}_1)\mathbf{F}(\mathbf{x}_1)')\mathbf{A}'=\Phi_0$. The desired result follows by combining (S7.1), (S7.2) and (S7.3). □

S8 Proof of Theorem 2

Similarly to (S7.1), we may write

$$\sqrt{\frac{nl_n}{c_n}}\mathcal{D}_n\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)=\left[\frac{c_n}{nl_n}\mathcal{D}_n^{-1}\sum_{k=1}^{n}\mathbf{Z}_{kn}\bar{\mathbf{f}}_{k-1}'\mathcal{D}_n^{-1}\right]^{-1}\mathcal{D}_n^{-1}\sqrt{\frac{c_n}{nl_n}}\sum_{k=1}^{n}\mathbf{Z}_{kn}\left[e_k-\frac{\sum_{s=1}^{n}e_sK_{sn}}{\sum_{s=1}^{n}K_{sn}}\right]. \quad \text{(S8.1)}$$

It follows from Theorem 7 that

$$\frac{c_n}{nl_n}\mathcal{D}_n^{-1}\sum_{k=1}^{n}\mathbf{Z}_{kn}\bar{\mathbf{f}}_{k-1}'\mathcal{D}_n^{-1}$$
$$=\frac{c_n}{nl_n}\mathcal{D}_n^{-1}\sum_{k=1}^{n}\mathbf{Z}_{kn}\mathbf{f}_{k-1}'\mathcal{D}_n^{-1}-\frac{c_n}{nl_n}\mathcal{D}_n^{-1}\sum_{k=1}^{n}\mathbf{Z}_{kn}\frac{\frac{c_n}{nl_n}\sum_{s=1}^{n}\mathbf{f}_{s-1}'K_{sn}}{\frac{c_n}{nl_n}\sum_{s=1}^{n}K_{sn}}\mathcal{D}_n^{-1}$$
$$=\frac{c_n}{nl_n}\mathcal{D}_n^{-1}\sum_{k=1}^{n}\mathbf{f}_{k-1}\mathbf{f}_{k-1}'K_{kn}\mathcal{D}_n^{-1}-\frac{c_n}{nl_n}\mathcal{D}_n^{-1}\sum_{k=1}^{n}\mathbf{f}_{k-1}K_{kn}\frac{\frac{c_n}{nl_n}\sum_{s=1}^{n}\mathbf{f}_{s-1}'K_{sn}}{\frac{c_n}{nl_n}\sum_{s=1}^{n}K_{kn}}\mathcal{D}_n^{-1}$$
$$\to_d\int K\left[\int_0^1 H_{\mathbf{f}}(\mathcal{X}_t)H_{\mathbf{f}}(\mathcal{X}_t)'dt-\int_0^1 H_{\mathbf{f}}(\mathcal{X}_t)dt\int_0^1 H_{\mathbf{f}}(\mathcal{X}_t)'dt\right]$$
$$=\int K\cdot\int_0^1\tilde{H}_{\mathbf{f}}(\mathcal{X}_t)\tilde{H}_{\mathbf{f}}(\mathcal{X}_t)'dt=\int K\cdot\Phi_2. \quad \text{(S8.2)}$$

Further, by letting $\mathcal{L}_n=diag\{1,\mathcal{D}_n\}$, we have

$$\mathcal{D}_n^{-1}\sqrt{\frac{c_n}{nl_n}}\sum_{k=1}^{n}\mathbf{Z}_{kn}\left[e_k-\frac{\sum_{s=1}^{n}e_sK_{sn}}{\sum_{s=1}^{n}K_{sn}}\right]$$
$$=\left[-\frac{1}{\frac{c_n}{nl_n}\sum_{k=1}^{n}K_{kn}}\mathcal{D}_n^{-1}\frac{c_n}{nl_n}\sum_{k=1}^{n}\mathbf{Z}_{kn},\ I_p\right]\sqrt{\frac{c_n}{nl_n}}\mathcal{L}_n^{-1}\sum_{k=1}^{n}K_{kn}\begin{bmatrix}1\\ \mathbf{f}_{k-1}\end{bmatrix}e_k$$
$$=:\mathbf{B}_n\sqrt{\frac{c_n}{nl_n}}\mathcal{L}_n^{-1}\sum_{k=1}^{n}K_{kn}\mathbf{F}_{k-1}e_k.$$

Again by Theorem 7 we get jointly with (S8.2)

$$\mathbf{B}_n \to_d \mathbf{B} := \left[-\int_0^1 H_{\mathbf{f}}(\mathcal{X}_t)dt,\ I_p \right], \tag{S8.3}$$

and

$$\sqrt{\frac{c_n}{nl_n}}\mathcal{L}_n^{-1}\sum_{k=1}^n K_{kn}\mathbf{F}_{k-1}e_k \to_d \mathbf{MN}\left(\mathbf{0}, \int K^2 \cdot E(\sigma_1^2)\int_0^1 \begin{bmatrix} 1 & H_{\mathbf{f}}(\mathcal{X}_t)' \\ H_{\mathbf{f}}(\mathcal{X}_t) & H_{\mathbf{f}}(\mathcal{X}_t)H_{\mathbf{f}}(\mathcal{X}_t)' \end{bmatrix} dt \right).$$

Therefore,

$$\begin{aligned}
\mathcal{D}_n^{-1}\sqrt{\frac{c_n}{nl_n}}\sum_{k=1}^n \mathbf{Z}_{kn}\left[e_k - \frac{\sum_{s=1}^n e_s K_{sn}}{\sum_{s=1}^n K_{sn}}\right] &= \mathbf{B}_n\sqrt{\frac{c_n}{nl_n}}\mathcal{L}_n^{-1}\sum_{k=1}^n K_{kn}\mathbf{F}_{k-1}e_k \\
&\to_d \mathbf{MN}\left(\mathbf{0}, \int K^2 \cdot E(\sigma_1^2)\mathbf{B}\int_0^1 \begin{bmatrix} 1 & H_{\mathbf{f}}(\mathcal{X}_t)' \\ H_{\mathbf{f}}(\mathcal{X}_t) & H_{\mathbf{f}}(\mathcal{X}_t)H_{\mathbf{f}}(\mathcal{X}_t)' \end{bmatrix} dt\mathbf{B}' \right) \\
&=_d \mathbf{MN}\left(\mathbf{0}, \int K^2 \cdot E(\sigma_1^2)\Phi_2\right).
\end{aligned} \tag{S8.4}$$

In view of (S8.1), (S8.2) and (S8.4) the CTLS estimator for $\boldsymbol{\beta}$

$$\sqrt{\frac{nl_n}{c_n}}\mathcal{D}_n\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \to_d \mathbf{MN}\left(\mathbf{0}, \frac{\int K^2}{\left(\int K\right)^2} \cdot E(\sigma_1^2)\Phi_2^{-1}\right),$$

as required. □

S9 Proof of Theorem 3

See Theorem S1 and hence the details are omitted. □

S10 Proof of Theorem 4 and additional explanation for (15)

We start with some preliminary results. We first derive the pseudo-true limits of the OLS and CTLS estimators under misbalancing (MB) assuming that the conditions of Theorem 4 hold. Recall that d_n denotes the normalizing sequence of Assumption A3(a) for the case $p = 1$. Set $f_k := f(x_{k-1})$, $f_{M,k} := f_M(x_{k-1})$, and $Q_n := diag\{\sqrt{n}, \sqrt{n}\pi_{f_M}(d_n)\}$. Furthermore, for a sequence a_k let $\tilde{a}_k := a_k - n^{-1}\sum_{j=1}^n a_j$. Similarly, for a function $A(t)$, $\tilde{A}(t) := A(t) - \int_0^1 A(s)ds$.

OLS under MB. Let $\left[\tilde{\mu}, \tilde{\beta}\right]$ be the OLS estimator for the parameters of (10) when the fitted model is given by (14). Suppose that the conditions of Theorem 4 hold. Then by Assumption **A3**(a), there is a sequence d_n such that $d_n^{-1}x_{[nt]} \Rightarrow \mathcal{X}_t$ in $D[0,1]$. In view of this and the continuity of the limit homogeneous functions of f and f_M, standard arguments (e.g. Park and Phillips, 2001; Theorem 5.2) give

$$\frac{1}{\pi_f(d_n)\sqrt{n}} Q_n \begin{bmatrix} \tilde{\mu} \\ \tilde{\beta} \end{bmatrix}$$

$$= \left[Q_n^{-1} \sum_k \begin{bmatrix} 1 \\ f_{M,k} \end{bmatrix} \begin{bmatrix} 1 \\ f_{M,k} \end{bmatrix}' Q_n^{-1} \right]^{-1} \frac{1}{\pi_f(d_n)\sqrt{n}} Q_n^{-1} \sum_k \begin{bmatrix} 1 \\ f_{M,k} \end{bmatrix} \beta f_k + o_P(1)$$

$$\to_d \left\{ \int_0^1 \begin{bmatrix} 1 & H_{f_M}(\mathcal{X}_t) \\ H_{f_M}(\mathcal{X}_t) & H_{f_M}^2(\mathcal{X}_t) \end{bmatrix} dt \right\}^{-1} \beta \int_0^1 \begin{bmatrix} H_f(\mathcal{X}_t) \\ H_{f_M}(\mathcal{X}_t) H_f(\mathcal{X}_t) \end{bmatrix} dt$$

$$= \left[\int_0^1 \tilde{H}_{f_M} H_{f_M} \right]^{-1} \begin{bmatrix} \int_0^1 H_f \int_0^1 H_{f_M}^2 - \int_0^1 H_{f_M} \int_0^1 H_{f_M} H_f \\ \int_0^1 \tilde{H}_{f_M} H_f \end{bmatrix} \beta =: \begin{bmatrix} \mu_* \\ \beta_* \end{bmatrix}.$$

In fact, the following joint weak limit holds

$$\frac{1}{\pi_f(d_n)} \tilde{\mu} \to_d \mu_* \quad \text{and} \quad \frac{\pi_{f_M}(d_n)}{\pi_f(d_n)} \tilde{\beta} \to_d \beta_*.$$

CTLS under MB. Similarly, consider the CTLS estimator (in this case, the CTLS instruments are: $Z_{kn} = f_{M,k} K_{kn}$). Using similar arguments as those used above together with Theorem 7 we get

$$\begin{aligned} \frac{\pi_{f_M}(d_n)}{\pi_f(d_n)} \hat{\beta} &= \frac{\frac{1}{\pi_{f_M}(d_n)\pi_f(d_n)}}{\frac{1}{\pi_{f_M}^2(d_n)}} \hat{\beta} = \beta \frac{\frac{c_n}{n l_n \pi_{f_M}(d_n)\pi_f(d_n)} \sum_k Z_{kn} \bar{f}_k}{\frac{c_n}{n l_n \pi_{f_M}^2(d_n)} \sum_k Z_{kn} \bar{f}_{M,k}} + o_P(1) \\ &\to_d \beta \frac{\int_0^1 \tilde{H}_f(\mathcal{X}_t) H_{f_M}(\mathcal{X}_t) dt}{\int_0^1 \tilde{H}_{f_M}(\mathcal{X}_t) H_{f_M}(\mathcal{X}_t) dt} = \beta_*. \end{aligned}$$

OLS estimator for σ^2 under MB. Next, we consider the variance estimator for σ^2 based on the OLS residuals $\check{e}_k$. Write

$$\begin{aligned} \pi_f^{-2}(d_n)\check{\sigma}^2 &= \pi_f^{-2}(d_n) n^{-1} \sum_k \check{e}_k^2 = \pi_f^{-2}(d_n) n^{-1} \sum_k \left(f_k - \tilde{\mu} - \tilde{\beta} f_{M,k} \right)^2 + o_P(1) \\ &= n^{-1} \sum_k \left(\pi_f^{-1}(d_n) f_k - \pi_f^{-1}(d_n)\tilde{\mu} - \frac{\pi_{f_M}(d_n)}{\pi_f(d_n)} \tilde{\beta} \pi_{f_M}^{-1}(d_n) f_{M,k} \right)^2 + o_P(1) \end{aligned}$$

$$\rightarrow_d \quad \int_0^1 [H_f - \mu_* - \beta_* H_{f_M}]^2 =: \sigma_*^2. \tag{S10.1}$$

Hence, we roughly have the following approximation

$$\check{\sigma}^2 \approx \pi_f^2(d_n)\sigma_*^2 + \sigma^2. \tag{S10.2}$$

CTLS t-statistics. Next, we introduce some additional notation. For any asymptotic homogeneous function $g : \mathbb{R} \to \mathbb{R}$ (and some kernel function K) set

$$\mathbf{A}_g := \left[-\int_0^1 H_g(\mathcal{X}_t)dt, 1\right], \; V_g := \begin{bmatrix} 1 & \int_0^1 H_g(\mathcal{X}_t)dt \\ \int_0^1 H_g(\mathcal{X}_t)dt & \int_0^1 H_g^2(\mathcal{X}_t)dt \end{bmatrix} \int K^2.$$

Note that for some AHF function f, the definitions above and straightforward calculations yield

$$\mathbf{A}_f V_f \mathbf{A}_f' = \int K^2 \int_0^1 \tilde{H}_f H_f. \tag{S10.3}$$

Furthermore, for any function $g : \mathbb{R} \to \mathbb{R}$ set $g_k = g(x_{k-1})$ and define

$$\mathcal{A}_{g,n} := [-\bar{g}, 1], \; \mathcal{V}_{g,n} := \sum_{k=1}^n K_{kn}^2 \begin{bmatrix} 1 & g_k \\ g_k & g_k^2 \end{bmatrix}, \; \check{\sigma}^2 = n^{-1}\sum_{k=1}^n \check{e}_k^2,$$

where $\bar{g} = \sum_{k=1}^n K_{kn} g(x_{k-1}) / \sum_{k=1}^n K_{kn}$ (cf. (7)), and $\check{e}_k$ are the OLS residuals based on the fitted regression function $g(x_{k-1})$. In the following, we assume $g = f$ (correct functional form) and $g = f_M$ (misbalanced model).

Without loss of generality, we shall consider t-statistics that utilize the studentization of (13). Hence, under correct functional form the CTLS t-statistic is

$$\hat{T} = \sum_{k=1}^n K_{kn} \bar{f}_k \frac{\hat{\beta} - \beta_0}{\sqrt{\check{\sigma}^2 \mathcal{A}_{f,n} \mathcal{V}_{f,n} \mathcal{A}_{f,n}'}},$$

where $\hat{\beta}$ is the CTLS estimator of β in (10). Under misbalancing, the CTLS t-statistic is of the form

$$\hat{T}_M = \sum_{k=1}^n K_{kn} \bar{f}_{M,k} \frac{\hat{\beta} - \beta_0}{\sqrt{\check{\sigma}^2 \mathcal{A}_{f_M,n} \mathcal{V}_{f_M,n} \mathcal{A}_{f_M,n}'}},$$

with $\hat{\beta}$ being the CTLS estimator based on the fitted model (14).

We now turn to the proof of Theorem 4. We start with (15). Using Theorem 7 the CTLS t-statistic under misbalancing is

$$
\begin{aligned}
\sqrt{\frac{c_n}{nl_n}}\hat{T}_M &:= \sqrt{\frac{c_n}{nl_n}}\sum_{k=1}^{n} K_{kn}\bar{f}_{M,k}\frac{\hat{\beta}}{\sqrt{\breve{\sigma}^2\mathcal{A}_{f_M,n}\mathcal{V}_{f_M,n}\mathcal{A}'_{f_M,n}}} \\
&= \sqrt{\frac{c_n}{nl_n}}\left[\frac{\beta\sum_k Z_{kn}\bar{f}_k}{\sqrt{\breve{\sigma}^2\mathcal{A}_{f_M,n}\mathcal{V}_{f_M,n}\mathcal{A}'_{f_M,n}}} + \frac{\sum_k Z_{kn}\bar{u}_k}{\sqrt{\breve{\sigma}^2\mathcal{A}_{f_M,n}\mathcal{V}_{f_M,n}\mathcal{A}'_{f_M,n}}}\right] \\
&= \frac{\frac{c_n}{nl_n\pi_{f_M}(d_n)\pi_f(d_n)}\beta\sum_k Z_{kn}\bar{f}_k}{\sqrt{\pi_f^{-2}\ \breve{\sigma}^2\frac{c_n}{nl_n\pi^2_{f_M}(d_n)}\mathcal{A}_{f_M,n}\mathcal{V}_{f_M,n}\mathcal{A}'_{f_M,n}}} \\
&\quad + \sqrt{\frac{c_n}{nl_n}}\frac{1}{\pi_f(d_n)}\frac{\sqrt{\frac{c_n}{nl_n\pi^2_{f_M}(d_n)}}\sum_k Z_{kn}\bar{u}_k}{\sqrt{\pi_f^{-2}\breve{\sigma}^2\frac{c_n}{nl_n\pi^2_{f_M}(d_n)}\mathcal{A}_{f_M,n}\mathcal{V}_{f_M,n}\mathcal{A}'_{f_M,n}}} \\
&\to_d \frac{\beta\int K\cdot\int_0^1\tilde{H}_f H_{f_M}}{\sqrt{\sigma_*^2\mathbf{A}_{f_M}V_{f_M}\mathbf{A}'_{f_M}}} + O_P\left(\sqrt{\frac{c_n}{nl_n}}\frac{1}{\pi_f(d_n)}\right), \qquad \text{(S10.4)}
\end{aligned}
$$

where σ_*^2 is defined in (S10.1). (15) follows directly from (S10.4).

Next, we show (16). Recall that the divergence rate under the correct specification is $\pi_f(d_n)\sqrt{\frac{nl_n}{c_n}}$. In fact, under correct FF and under H_1 we have

$$
\begin{aligned}
\sqrt{\frac{c_n}{\pi_f^2(d_n)nl_n}}\hat{T} &:= \sqrt{\frac{c_n}{\pi_f^2(d_n)nl_n}}\sum_{k=1}^{n} K_{kn}\bar{f}_k\frac{\hat{\beta}}{\sqrt{\breve{\sigma}^2\mathcal{A}_{f,n}\mathcal{V}_{f,n}\mathcal{A}'_{f,n}}} \\
&= \sqrt{\frac{c_n}{\pi_f^2(d_n)nl_n}}\sum_{k=1}^{n} K_{kn}\bar{f}_k\frac{\hat{\beta}-\beta}{\sqrt{\breve{\sigma}^2\mathcal{A}_{f,n}\mathcal{V}_{f,n}\mathcal{A}'_{f,n}}} \\
&\quad + \frac{c_n}{nl_n\pi_f^2(d_n)}\sum_k Z_{kn}\bar{f}_k\frac{\beta}{\sqrt{\breve{\sigma}^2\frac{c_n}{nl_n\pi_f^2}\mathcal{A}_{f,n}\mathcal{V}_{f,n}\mathcal{A}'_{f,n}}} \\
&\to_d O_P\left(\sqrt{\frac{c_n}{\pi_f^2(d_n)nl_n}}\right) + \frac{\beta\int K\cdot\int_0^1\tilde{H}_f(\mathcal{X}_t)H_f(\mathcal{X}_t)dt}{\sqrt{\sigma^2\mathbf{A}_f V_f\mathbf{A}'_f}}. \qquad \text{(S10.5)}
\end{aligned}
$$

Result (16) follows directly from (S10.3), (S10.4) and (S10.5). This completes the proof of Theorem 4. □

We finally consider supporting arguments for (17). First, note that (S10.2) together with (S10.4) postulate the following approximate behavior.

$$\hat{T}_M \approx \pi_f(d_n)\sqrt{\frac{nl_n}{c_n}}\frac{\beta\int K\cdot\int_0^1 \tilde{H}_f H_{f_M}}{\sqrt{(\pi_f^2(d_n)\sigma_*^2+\sigma^2)\mathbf{A}_{f_M}V_{f_M}\mathbf{A}'_{f_M}}}. \tag{S10.6}$$

Furthermore, by (S10.5) we have

$$\hat{T} \approx \pi_f(d_n)\sqrt{\frac{nl_n}{c_n}}\frac{\beta\int K\cdot\int_0^1 \tilde{H}_f H_f}{\sqrt{\sigma^2\mathbf{A}_f V_f\mathbf{A}'_f}}. \tag{S10.7}$$

Combining (S10.5) and (S10.6), the ratio of the two test statistics is

$$\begin{aligned}
\hat{T}/\hat{T}_M &\approx \frac{\sqrt{\left(\pi_f^2(d_n)\sigma_*^2+\sigma^2\right)\mathbf{A}_{f_M}V_{f_M}\mathbf{A}'_{f_M}}\int_0^1 \tilde{H}_f H_f}{\sqrt{\sigma^2\mathbf{A}_f V_f\mathbf{A}'_f}\int_0^1 \tilde{H}_{f_M}H_f} \\
&= \sqrt{\frac{\pi_f^2(d_n)\sigma_*^2+\sigma^2}{\sigma^2}}\frac{\sqrt{\int_0^1 \tilde{H}_{f_M}H_{f_M}}\int_0^1 \tilde{H}_f H_f}{\sqrt{\int_0^1 \tilde{H}_f H_f}\int_0^1 \tilde{H}_{f_M}H_{f_M}} \\
&= \sqrt{\frac{\pi_f^2(d_n)\sigma_*^2+\sigma^2}{\sigma^2}}\frac{\sqrt{\int_0^1 \tilde{H}_{f_M}H_{f_M}\int_0^1 \tilde{H}_f H_f}}{\int_0^1 \tilde{H}_{f_M}H_f} \geq \sqrt{\frac{\pi_f^2(d_n)\sigma_*^2+\sigma^2}{\sigma^2}},
\end{aligned}$$

where the lower bound above is due to the Cauchy-Schwarz inequality.

S11 Proof of Theorem S1

We only prove (S2.6), as the proof of (S2.5) is similar.

We first assume that the conditions of Theorem 2 hold, together with $\sup_{k\geq 1} Eu_k^4 < \infty$. Define $\mathcal{L}_n := diag\left\{1, \pi_{f_1}(d_{1n}), ..., \pi_{f_p}(d_{pn})\right\}$ and $\mathcal{V}_n = \sum_{k=1}^n e_k^2\mathbf{F}_{k-1}\mathbf{F}'_{k-1}K_{kn}^2$. Since the OLS residuals in model (S2.1) satisfy

$$\check{e}_k^2 = \left[\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{LS}\right)'\mathbf{F}_{k-1}\right]^2 + 2\left[\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{LS}\right)'\mathbf{F}_{k-1}\right]e_k + e_k^2,$$

it follows that

$$\hat{\mathcal{V}}_n = \sum_{k=1}^n \check{e}_k^2\mathbf{F}_{k-1}\mathbf{F}'_{k-1}K_{kn}^2 = \mathcal{V}_n + R_{1n} + 2R_{2n} \tag{S11.1}$$

where

$$
\begin{aligned}
R_{1n} &= \sum_{k=1}^{n} \left[\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{LS} \right)' \mathbf{F}_{k-1} \right]^2 \mathbf{F}_{k-1} \mathbf{F}'_{k-1} K_{kn}^2, \\
R_{2n} &= \sum_{k=1}^{n} e_k \left[\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{LS} \right)' \mathbf{F}_{k-1} \right] \mathbf{F}_{k-1} \mathbf{F}'_{k-1} K_{kn}^2.
\end{aligned}
$$

We next show that, for $j = 1$ and 2,

$$
\left\| \frac{c_n}{n l_n} \mathcal{L}_n^{-1} R_{jn} \mathcal{L}_n^{-1} \right\| = o_P(1). \tag{S11.2}
$$

In fact, given that the covariates satisfy an FCLT and the regression function is continuous, standard arguments (see, e.g., Park and Phillips, 2001; Chang, Park and Phillips, 2001) give

$$
\sqrt{n} \mathcal{L}_n \left(\hat{\boldsymbol{\theta}}_{LS} - \boldsymbol{\theta} \right) = O_P(1). \tag{S11.3}
$$

This result implies that

$$
\begin{aligned}
\left\| \frac{c_n}{n l_n} \mathcal{L}_n^{-1} R_{1n} \mathcal{L}_n^{-1} \right\| &= \left\| \frac{c_n}{n l_n} \sum_{k=1}^{n} \left[\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{LS} \right)' \mathcal{L}_n \mathcal{L}_n^{-1} \mathbf{F}_{k-1} \right]^2 \mathcal{L}_n^{-1} \mathbf{F}_{k-1} \mathbf{F}'_{k-1} \mathcal{L}_n^{-1} K_{kn}^2 \right\| \\
&\leq n^{-1} \left\| \sqrt{n} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{LS} \right)' \mathcal{L}_n \right\|^2 \cdot \frac{c_n}{n l_n} \sum_{k=1}^{n} \left\| \mathcal{L}_n^{-1} \mathbf{F}_{k-1} \right\|^4 K_{kn}^2 \\
&= O_P\left(n^{-1} \right) \cdot \frac{c_n}{n l_n} \sum_{k=1}^{n} \left\| \mathcal{L}_n^{-1} \mathbf{F}_{k-1} \right\|^4 K_{kn}^2 = O_P(n^{-1}),
\end{aligned}
$$

where we have used Lemma S1 with $h(x) = ||x||^4$ and the similar argument as in the proof of Theorem 6 (c.g. the proof of (S5.2)), yielding

$$
\frac{c_n}{n l_n} \sum_{k=1}^{n} \left\| \mathcal{L}_n^{-1} \mathbf{F}_{k-1} \right\|^4 K_{kn}^2 \to_d \int K^2 \cdot \int_0^1 \left\| H_{\mathbf{F}} \left(\mathcal{X}_t \right) \right\|^4 dt.
$$

Hence (S11.2) is true with $j = 1$. Similarly, using (S11.3) again, we have

$$
\begin{aligned}
\left\| \frac{c_n}{n l_n} \mathcal{L}_n^{-1} R_{2n} \mathcal{L}_n^{-1} \right\| &= \frac{2 c_n}{n l_n} \left\| \sum_{k=1}^{n} e_k \left[\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{LS} \right)' \mathbf{F}_{k-1} \right] \mathcal{L}_n^{-1} \mathbf{F}_{k-1} \mathbf{F}'_{k-1} \mathcal{L}_n^{-1} K_{kn}^2 \right\| \\
&\leq 2 n^{-1/2} \left\| \sqrt{n} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{LS} \right)' \mathcal{L}_n \right\| \cdot \frac{c_n}{n l_n} \sum_{k=1}^{n} |e_k| \left\| \mathcal{L}_n^{-1} \mathbf{F}_{k-1} \right\|^3 K_{kn}^2 \\
&\leq O_P\left(n^{-1/2} \right) \cdot \frac{c_n}{n l_n} \sum_{k=1}^{n} \left(e_k^2 + 1 \right) \left\| \mathcal{L}_n^{-1} \mathbf{F}_{k-1} \right\|^3 K_{kn}^2
\end{aligned}
$$

$$= O_P\left(n^{-1/2}\right) \cdot O_P(1) = o_P(1),$$

i.e., (S11.2) is also true with $j = 2$.

In terms of (S11.1) and (S11.2), we claim that, as $n \to \infty$,

$$\frac{c_n}{nl_n}\mathcal{L}_n^{-1}\hat{\mathcal{V}}_n\mathcal{L}_n^{-1} = \frac{c_n}{nl_n}\mathcal{L}_n^{-1}\mathcal{V}_n\mathcal{L}_n^{-1} + o_P(1). \tag{S11.4}$$

Now (S2.6) under the conditions of Theorem 2 is a direct consequence of Theorem 2 and (S11.4). To see this, set $\mathcal{D}_n^* := \sqrt{\frac{nl_n}{c_n}}\mathcal{D}_n$, $\mathcal{L}_n^* := \sqrt{\frac{nl_n}{c_n}}\mathcal{L}_n$, with $\mathcal{D}_n$ defined in Theorem 2 and recall $\mathcal{L}_n = diag\{1, \mathcal{D}_n\}$. Under the null hypothesis,

$$\begin{aligned}\hat{F} &= \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)' \left[\mathcal{H}_n^{-1}\mathcal{A}_n\hat{\mathcal{V}}_n\mathcal{A}_n'\mathcal{H}_n^{-1}\right]^{-1}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \\ &= \left[\mathcal{D}_n^*\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\right]' \mathcal{M}_n^{-1}\left[\mathcal{D}_n^*\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\right]\end{aligned}$$

with

$$\mathcal{M}_n := \left(\mathcal{D}_n^{*-1}\mathcal{H}_n\mathcal{D}_n^{*-1}\right)^{-1} \cdot \left(\mathcal{D}_n^{*-1}\mathcal{A}_n\mathcal{L}_n^*\right) \cdot \mathcal{L}_n^{*-1}\hat{\mathcal{V}}_n\mathcal{L}_n^{*-1} \cdot \mathcal{L}_n^*\mathcal{A}_n'\mathcal{D}_n^{*-1} \cdot \left(\mathcal{D}_n^{*-1}\mathcal{H}_n\mathcal{D}_n^{*-1}\right)^{-1}.$$

Next, note that by (S11.4) and Theorem 2

$$\mathcal{L}_n^{*-1}\hat{\mathcal{V}}_n\mathcal{L}_n^{*-1} = \mathcal{L}_n^{*-1}\mathcal{V}_n\mathcal{L}_n^{*-1} + o_P(1) \to_d E(\sigma_1^2)\int K^2 \cdot \int_0^1 H_{\mathbf{F}}(X_t)\, H_{\mathbf{F}}(X_t)'\, dt. \tag{S11.5}$$

Further, by (S8.2)

$$\mathcal{D}_n^{*-1}\mathcal{H}_n\mathcal{D}_n^{*-1} \to_d \int K \cdot \Phi_2, \tag{S11.6}$$

with Φ_2 defined in Theorem 2. Moreover, using (S8.3)

$$\mathcal{D}_n^{*-1}\mathcal{A}_n\mathcal{L}_n^* = \mathbf{B}_n \to_d \mathbf{B}, \tag{S11.7}$$

with $\mathbf{B}$ defined in (S8.3). Combining (S11.5)-(S11.7),

$$\mathcal{M}_n \to_d E(\sigma_1^2)\frac{\int K^2}{\left(\int K\right)^2}\Phi_2^{-1}\mathbf{B}\int_0^1 H_{\mathbf{F}}(X_t)\, H_{\mathbf{F}}(X_t)'\, dt\, \mathbf{B}'\Phi_2^{-1} = E(\sigma_1^2)\frac{\int K^2}{\left(\int K\right)^2}\Phi_2^{-1}, \tag{S11.8}$$

where we have used the fact that

$$\Phi_2 = \mathbf{B} \int_0^1 H_{\mathbf{F}}(\mathcal{X}_t) H_{\mathbf{F}}(\mathcal{X}_t)' \, dt \, \mathbf{B}',$$

see e.g. (S8.4). In view of (S11.8) and Theorem 2,

$$\mathcal{M}_n^{-1/2} \mathcal{D}_n^* \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \right) \to_d \mathbf{N}(\mathbf{0}, I_q),$$

and the result follows.

We next prove (S2.6) under the conditions of Theorem 1, together with $\sup_{k \geq 1} E u_k^4 < \infty$. By using Theorem 1 and similar arguments as in the first part, it suffices to show that

$$\frac{c_n}{n l_n} \hat{\mathcal{V}}_n \quad \to_P \quad \int K^2 \cdot E\left[\sigma_2^2 \mathbf{F}(\mathbf{x}_1) \mathbf{F}(\mathbf{x}_1)'\right]. \tag{S11.9}$$

Note that (S11.1) still holds. (S11.9) will follow if we prove

$$\left\| \frac{c_n}{n l_n} R_{jn} \right\| \quad = \quad o_P(1), \quad j = 1, 2, \tag{S11.10}$$

and

$$\frac{c_n}{n l_n} \mathcal{V}_n \to_P \int K^2 \cdot E\left[\sigma_2^2 \mathbf{F}(\mathbf{x}_1) \mathbf{F}(\mathbf{x}_1)'\right]. \tag{S11.11}$$

The proof of (S11.10) is simple. In fact, by recalling **A2** (i.e., $\mathbf{x}_k$ and $\mathbf{F}_k = \mathbf{F}(\mathbf{x}_{k-1})$ are stationary with $E \|\mathbf{F}_1\|^2 < \infty$), the OLS estimator in this case satisfies

$$\sqrt{n} \left(\hat{\boldsymbol{\theta}}_{LS} - \boldsymbol{\theta} \right) = O_P(1) \tag{S11.12}$$

and $\max_{1 \leq k \leq n} \left\| n^{-1/2} \mathbf{F}_{k-1} \right\| = o_P(1)$. It follows from these facts that

$$\begin{aligned}
\left\| \frac{c_n}{n l_n} R_{1n} \right\| &= \left\| \frac{c_n}{n l_n} \sum_{k=1}^n \left[\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{LS} \right)' \mathbf{F}_{k-1} \right]^2 \mathbf{F}_{k-1} \mathbf{F}_{k-1}' K_{kn}^2 \right\| \\
&\leq n^{-1} \left\| \sqrt{n} \left(\hat{\boldsymbol{\theta}}_{LS} - \boldsymbol{\theta} \right) \right\|^2 \frac{c_n}{n l_n} \sum_{k=1}^n \|\mathbf{F}_{k-1}\|^4 K_{kn}^2 \\
&\leq \left\| \sqrt{n} \left(\hat{\boldsymbol{\theta}}_{LS} - \boldsymbol{\theta} \right) \right\|^2 \left[\max_{1 \leq k \leq n} \left\| n^{-1/2} \mathbf{F}_{k-1} \right\| \right]^2 \frac{c_n}{n l_n} \sum_{k=1}^n \|\mathbf{F}_{k-1}\|^2 K_{kn}^2 \\
&= o_P(1),
\end{aligned}$$

where, in the last equality, we have used the result:

$$\frac{c_n}{nl_n}\sum_{k=1}^{n}\|\mathbf{F}_{k-1}\|^2 K_{kn}^2 \to_P \int K^2\cdot E\|\mathbf{F}_1\|^2$$

as explained in Remark 14 with $G(x)\equiv 1$. Similarly, we have

$$\begin{aligned}
\frac{1}{2}\left\|\frac{c_n}{nl_n}R_{2n}\right\| &= \frac{c_n}{nl_n}\left\|\sum_{k=1}^{n} e_k\left[\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{LS}\right)'\mathbf{F}_{k-1}\right]\mathbf{F}_{k-1}\mathbf{F}_{k-1}'K_{kn}^2\right\| \\
&\leq \left[\left\|\sqrt{n}\left(\hat{\boldsymbol{\theta}}_{LS}-\boldsymbol{\theta}\right)\right\|\max_{1\leq k\leq n}\left\|n^{-1/2}\mathbf{F}_{k-1}\right\|\right]\frac{c_n}{nl_n}\sum_{k=1}^{n}|e_k|\,\|\mathbf{F}_{k-1}\|^2 K_{kn}^2 \\
&\leq \left[\left\|\sqrt{n}\left(\hat{\boldsymbol{\theta}}_{LS}-\boldsymbol{\theta}\right)\right\|\max_{1\leq k\leq n}\left\|n^{-1/2}\mathbf{F}_{k-1}\right\|\right]\frac{c_n}{nl_n}\sum_{k=1}^{n}\left(1+e_k^2\right)\|\mathbf{F}_{k-1}\|^2 K_{kn}^2 \\
&= o_P(1),
\end{aligned}$$

where, in the last equality, we have used the fact that

$$\frac{c_n}{nl_n}\sum_{k=1}^{n}\left(1+e_k^2\right)\|\mathbf{F}_{k-1}\|^2 K_{kn}^2 \to_P \int K^2\cdot E\left[\left(1+\sigma_2^2\right)\|\mathbf{F}_1\|^2\right].$$

Next, we prove (S11.11). Let $[A]_{rs}$ denote the (r,s) element of a matrix A. In view of Lemmas 1 and 2 with $G(x)=1$ (cf. Remark 14), it suffices to show that, for any $m:=m_n\to\infty$, $n/m_n\to\infty$,

$$\Delta_{rs,n} := \max_{m\leq j\leq n-m}\frac{1}{m}E\left|\sum_{k=j+1}^{j+m}\left[e_k^2\mathbf{F}_{k-1}\mathbf{F}_{k-1}'\right]_{rs}-E\left[\sigma_k^2\mathbf{F}_{k-1}\mathbf{F}_{k-1}'\right]_{rs}\right|\to 0. \quad \text{(S11.13)}$$

Note that $\Delta_{rs,n}\leq R_{3n}+R_{4n}$, where

$$\begin{aligned}
R_{3n} &= \max_{m\leq j\leq n-m}\frac{1}{m}E\left|\sum_{k=j+1}^{j+m}\left[\sigma_k^2\mathbf{F}_{k-1}\mathbf{F}_{k-1}'\right]_{rs}\left[u_k^2-E\left(u_k^2|\mathcal{F}_{k-1}\right)\right]\right| \\
R_{4n} &= \max_{m\leq j\leq n-m}E\left|\frac{1}{m}\sum_{k=j+1}^{j+m}\left\{\left[\sigma_k^2\mathbf{F}_{k-1}\mathbf{F}_{k-1}'\right]_{rs}-E\left[\sigma_k^2\mathbf{F}_{k-1}\mathbf{F}_{k-1}'\right]_{rs}\right\}\right|.
\end{aligned}$$

Since $\sigma_k^2\mathbf{F}_{k-1}\mathbf{F}_{k-1}'$ is strict stationarity and ergodic, it is readily seen that $R_{4n}\to 0$.

To consider R_{1n}, set $\lambda_k := \left[\sigma_k^2\mathbf{F}_{k-1}\mathbf{F}_{k-1}'\right]_{rs}$ and $U_k := \left[u_k^2-E\left(u_k^2|\mathcal{F}_{k-1}\right)\right]$. Then for all

$A > 0$ as $m \to \infty$ first and then $A \to \infty$,

$$\begin{aligned}
R_{3n} &\le \max_{m\le j\le n-m} \frac{1}{m} E\left|\sum_{k=j+1}^{j+m} \lambda_k I\{|\lambda_k| \le A\} U_k\right| + \max_{m\le j\le n-m} \frac{1}{m} E\left|\sum_{k=j+1}^{j+m} \lambda_k I\{|\lambda_k| > A\} U_k\right| \\
&\le A \max_{m\le j\le n-m} \left\{\frac{1}{m^2} E \sum_{k=j+1}^{j+m} U_k^2\right\}^{1/2} \\
&\quad + \max_{m\le j\le n-m} \frac{1}{m} E \sum_{k=j+1}^{j+m} |\lambda_k| I\{|\lambda_k| > A\} \left[E\left(u_k^2|\mathcal{F}_{k-1}\right) + u_k^2\right] \\
&\le A \left\{\frac{1}{m} \sup_k E u_k^4\right\}^{1/2} + 2E|\lambda_1| I\{|\lambda_1| > A\} \to 0.
\end{aligned}$$

Combining all these estimates, we establish (S11.13). This completes the proof. □

S12 Additional Simulation Results

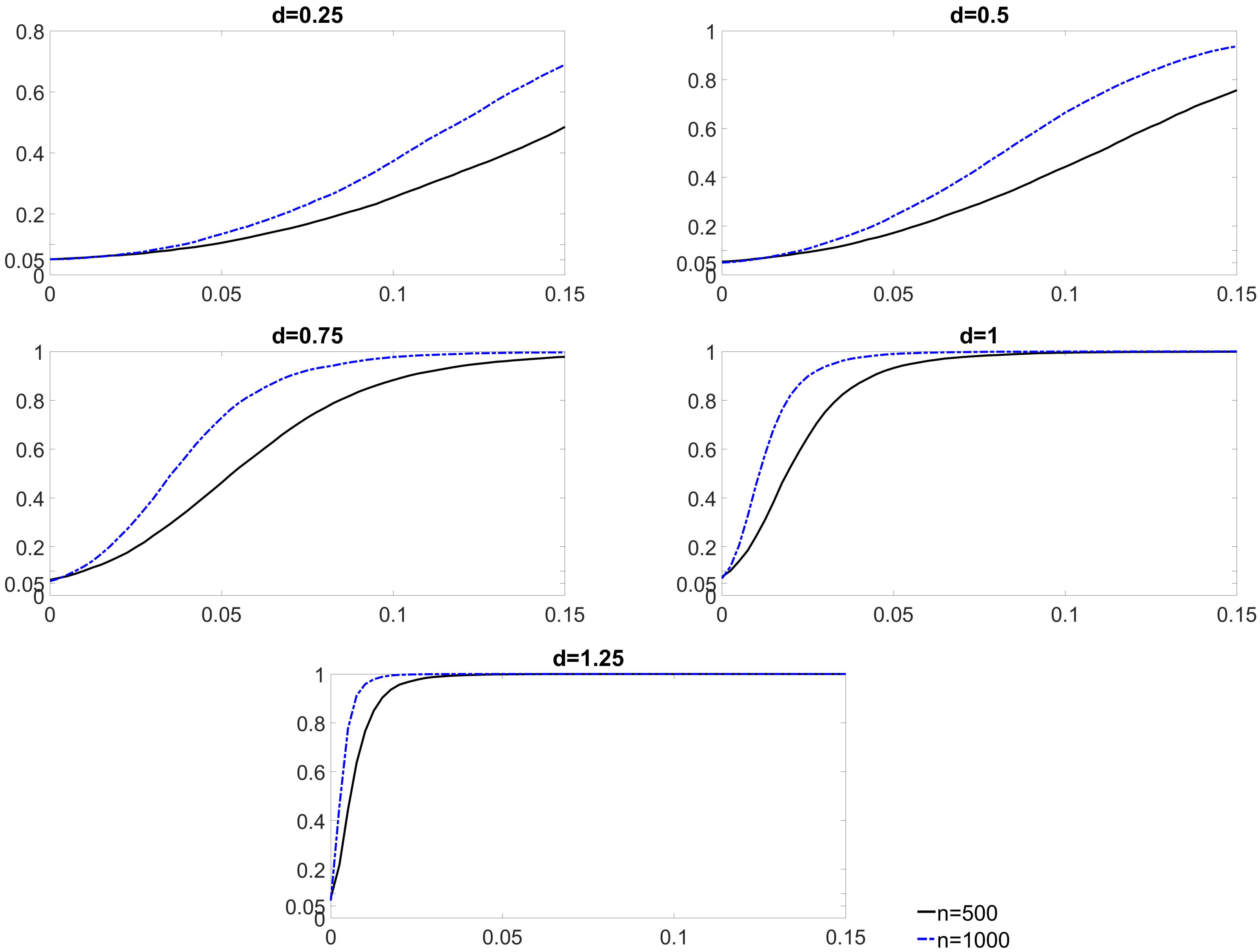
Table S1: Empirical Size of CTLS tests: $\hat{T}$
(nominal size 5%; NI regressor, GARCH(1,1) regression errors)

	$c = 0$					$c = -5$				
ϱ	-0.95	-0.5	0	0.5	0.95	-0.95	-0.5	0	0.5	0.95
n=250	0.083	0.060	0.052	0.062	0.087	0.062	0.054	0.050	0.060	0.069
500	0.077	0.057	0.050	0.061	0.077	0.060	0.051	0.051	0.055	0.059
750	0.070	0.059	0.052	0.058	0.070	0.060	0.056	0.058	0.057	0.061
1000	0.065	0.054	0.048	0.056	0.069	0.054	0.051	0.045	0.049	0.055
	$c = -10$					$c = -20$				
ϱ	-0.95	-0.50	0.00	0.50	0.95	-0.95	-0.50	0.00	0.50	0.95
n=250	0.057	0.053	0.055	0.058	0.062	0.056	0.054	0.056	0.056	0.058
500	0.055	0.049	0.048	0.050	0.055	0.053	0.049	0.047	0.049	0.053
750	0.055	0.055	0.055	0.057	0.055	0.050	0.051	0.053	0.056	0.055
1000	0.052	0.049	0.048	0.046	0.053	0.052	0.048	0.045	0.047	0.053

Table S2: Empirical Size of CTLS tests: $\hat{T}$ (nominal size 5%; fractional regressor, cond. homoscedastic regression errors)

		$d = 0.25$			$d = 0.5$			$d = 0.75$			$d = 0.8$		
	ϱ	-0.95	-0.5	0	-0.95	-0.5	0	-0.95	-0.5	0	-0.95	-0.5	0
CTLS	n=250	0.053	0.052	0.053	0.056	0.051	0.049	0.067	0.053	0.049	0.071	0.056	0.049
	500	0.052	0.052	0.050	0.055	0.050	0.047	0.064	0.053	0.049	0.069	0.054	0.052
	750	0.051	0.055	0.056	0.056	0.055	0.055	0.066	0.057	0.055	0.067	0.057	0.056
	1000	0.051	0.052	0.049	0.051	0.049	0.049	0.059	0.052	0.050	0.061	0.052	0.050
OLS	n=250	0.050	0.052	0.052	0.074	0.059	0.053	0.158	0.085	0.051	0.184	0.093	0.052
	500	0.052	0.050	0.048	0.072	0.055	0.051	0.161	0.085	0.054	0.184	0.091	0.055
	750	0.052	0.051	0.052	0.068	0.058	0.053	0.155	0.081	0.053	0.178	0.086	0.051
	1000	0.050	0.048	0.049	0.067	0.053	0.047	0.155	0.077	0.049	0.183	0.086	0.049
		$d = 0.9$			$d = 1$			$d = 1.1$			$d = 1.2$		
	ϱ	-0.95	-0.5	0	-0.95	-0.5	0	-0.95	-0.5	0	-0.95	-0.5	0
CTLS	n=250	0.077	0.057	0.051	0.084	0.059	0.051	0.089	0.060	0.052	0.089	0.063	0.053
	500	0.073	0.058	0.052	0.077	0.059	0.054	0.081	0.063	0.054	0.082	0.061	0.051
	750	0.073	0.058	0.054	0.076	0.059	0.052	0.078	0.060	0.051	0.079	0.061	0.051
	1000	0.066	0.054	0.050	0.070	0.054	0.049	0.072	0.055	0.050	0.072	0.054	0.051
OLS	n=250	0.235	0.107	0.053	0.278	0.117	0.053	0.308	0.121	0.052	0.325	0.126	0.052
	500	0.242	0.102	0.054	0.287	0.114	0.054	0.319	0.120	0.055	0.337	0.123	0.056
	750	0.230	0.098	0.052	0.272	0.109	0.051	0.301	0.117	0.051	0.322	0.119	0.053
	1000	0.229	0.102	0.053	0.278	0.111	0.053	0.310	0.118	0.054	0.327	0.120	0.055

Figure S1: Empirical Power of CTLS tests: $\hat{T}$ Plotted against β.
(5% nominal size; $\varrho = -0.95 = -0.95$; fractional regressor, cond. homoscedastic regression errors)



References

[1] Andersen, T.G. and Varneskov, R.T. (2021). Consistent inference for predictive regressions in persistent economic systems. *Journal of Econometrics*, **224**(1), 215–244.

[2] Billingsley P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[3] Campbell, J.Y. and Yogo, M. (2006). Efficient tests of stock return predictability. *Journal of Financial Economics,* **81**, 27–60.

[4] Cavanagh, C.L., Elliot, G. and Stock, J. H. (1995). Inference in models with nearly integrated regressors. *Econometric Theory*, **11**, 1131–1147.

[5] Chang, Y., Park, J.Y. and Phillips P.C.B. (2001). Nonlinear econometric models with cointegrated and deterministically trending regressors. *Econometrics Journal*, **4**, 1–36.

[6] Christensen, B.J. and Nielsen M. Ø. (2006). Asymptotic normality of narrow-band least squares in the stationary fractional cointegration model and volatility forecasting. *Journal of Econometrics*, **133**, 343–371.

[7] Demetrescu, M., Georgiev, I., Rodrigues, P. and Taylor, R. (2022). Testing for episodic predictability in stock returns. *Journal of Econometrics*, **227**(1), 85–113.

[8] Elliott, G., Müller, U.K. and Watson, M.W. (2015) Nearly optimal tests when a nuisance parameter is present under the null hypothesis. *Econometrica*, **83**, 771–811.

[9] Hall, P. and Heyde, C.C. (1980). *Martingale Limit Theory and its Application.* Academic Press, New York.

[10] Hualde, J. and Robinson, P.M. (2010). Semiparametric inference in multivariate fractionally cointegrated systems. *Journal of Econometrics*, **157**(2), 492–511.

[11] Janson, M. and Moreira, J.M. (2006). Optimal inference in regression models with nearly integrated regressors. *Econometrica*, **74**(3), 681–714.

[12] Kostakis, A., Magdalinos, T. and Stamatogiannis, M.P. (2015). Robust econometric inference for stock return predictability. *Review of Financial Studies*, **28**(5), 1506–1553.

[13] Magdalinos, T. (2022). Least squares and IVX limit theory in systems of predictive regressions with GARCH innovations. *Econometric Theory*, **38**(5), 875–912.

[14] Magdalinos, T. and Petrova, K. (2022). Uniform and distribution-free inference with general autoregressive processes, Working Paper 1344, Barcelona School of Economics.

[15] Magdalinos, T. and Phillips, P.C.B. (2018). Wald testing with matrix normalization. Working Paper, University of Southampton.

[16] Mikusheva, A. (2007). Uniform inference in autoregressive models. *Econometrica*, **75**(5), 1411–1452.

[17] Park, J.Y. and Phillips P.C.B. (2001). Nonlinear regressions with integrated time series. *Econometrica*, **69**(1), 117–161.

[18] Phillips, P.C.B. (1995). Fully modified least squares and vector autoregression. *Econometrica*, **63**(5), 1023–1078.

[19] Phillips, P.C.B. (2014). On confidence intervals for autoregressive roots and predictive regressions. *Econometrica*, **82**(3), 1177–1195.

[20] Phillips, P.C.B. (2015). Pitfalls and possibilities in predictive regression. *Journal of Financial Econometrics*, **13**(3), 521–555.

[21] Phillips, P.C.B. and Magdalinos, T. (2009). Econometric inference in the vicinity of unity. *Mimeo*, Singapore Management University.

[22] Robinson, P.M. and Hualde, J. (2003). Cointegration in fractional systems with unknown integration orders. *Econometrica*, **71**(6), 1727–1766.

[23] Wang, Q. (2014). Martingale limit theorem revisited and nonlinear cointegrating regression. *Econometric Theory*, **30**(3), 509–535.

[24] Wang, Q. (2015). *Limit Theorems for Nonlinear Cointegrating Regression*. World Scientific, Singapore.

[25] Yang, B., Long W., Peng, L. and Cai, Z. (2020). Testing the predictability of U.S. housing price index returns based on an IVX-AR model. *Journal of the American Statistical Association*, **115**(532), 1–34.