TI 2023-079/III
Tinbergen Institute Discussion Paper

Bootstrapping trending time-varying coefficient panel models with missing observations

Yicong Lin[1]
Bernhard van der Sluis[2]
Marina Friedrich[3]

1 Vrije Universiteit Amsterdam and Tinbergen Institute

2 Erasmus University Rotterdam

3 Vrije Universiteit Amsterdam and Tinbergen Institute

Bootstrapping trending time-varying coefficient panel models with missing observations

Yicong Lin*[1,2], Bernhard van der Sluis[3] and Marina Friedrich[1,2]

[1]Vrije Universiteit Amsterdam

[2]Tinbergen Institute

[3]Erasmus University Rotterdam

December, 2023

Abstract

We study a class of trending panel regression models with time-varying coefficients that incorporate cross-sectional and serial dependence, as well as heteroskedasticity. Our models also allow for missing observations in the dependent variable. We introduce a local linear dummy variable estimator capable of handling missing observations and derive its asymptotic properties. A key ingredient in our theoretical framework is a generic uniform convergence result for near-epoch processes in kernel estimation for large panels ($N, T \to \infty$). The resulting limiting distribution reflects the pattern of missing values and depends on various nuisance parameters. An autoregressive wild bootstrap (AWB) is proposed to construct confidence intervals and bands. The AWB accommodates missing observations and automatically replicates all the nuisance parameters, demonstrating good finite sample performance. We apply our methods to investigate (i) the relationship between $PM_{2.5}$ and mortality and (ii) common trends in atmospheric ethane emissions in the Northern Hemisphere. Both examples yield statistical evidence for time variation.

Keywords: autoregressive wild bootstrap, confidence bands, cross-sectional and serial dependence, time-varying, missing observations.

*Corresponding author: Department of Econometrics and Data Science, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, the Netherlands. E-mail address: yc.lin@vu.nl.

1 Introduction

Common trends and time-varying relations are often observed in panel data with a long time span. Accordingly, recent empirical studies frequently adopt panel regression models that allow trending intercepts and time-varying coefficients to evolve smoothly and deterministically (Silvapulle et al., 2017; Cai et al., 2018; Hailemariam et al., 2019; Liddle et al., 2020; Uddin et al., 2020; Awaworyi Churchill et al., 2021; Ren et al., 2022; Sun et al., 2023). This formulation offers the advantage of straightforward empirical interpretation without compromising flexibility in the model specification. However, despite the widespread use of these panel models in empirical studies, there exists limited theoretical guidance on conducting inference. This shortfall constitutes a significant gap between empirical and theoretical studies in two key aspects.

First, existing asymptotic approximations depend heavily on various nuisance parameters that are difficult to estimate, including second-order bias terms and long-run covariance matrices (Li et al., 2011; Chen and Huang, 2018). These nuisance parameters pose a challenge when applying the asymptotic results for inference. Therefore, the previously mentioned empirical studies resort to a naive wild bootstrap, generating bootstrap samples series by series, for constructing pointwise confidence intervals at each time point. However, the theoretical justification for employing any bootstrap method to construct confidence intervals in these models is absent from the literature. This naive method falls short in capturing cross-sectional and serial dependence (Gonçalves and Perron, 2020), making it unsuitable for many macro-level applications due to the pervasive presence of dependence in the data. As such, there is a pressing need for developing a more sophisticated bootstrap method and corresponding theoretical foundations.

Second, in various applications, particularly in climate sciences and environmental economics, the prevalence of missing observations is evident (Keef et al., 2009). This can be attributed, for instance, to measurements being impeded by unfavorable weather conditions (Friedrich et al., 2020). Common practice involves multiple imputation. Nevertheless, particularly for climatic data, multiple imputation is challenging, given its complex dependence structure across units (Bashir and Wei, 2018; Li et al., 2018; Cahan et al., 2023). The impact of imputed data on statistical inference remains unclear as well. Missing observations add additional complexity, as quantifying the uncertainty about the occurrence of missing data becomes crucial for ensuring valid inference. In our setting, there are currently no available methods that allow inference in the presence of missing observations. This limitation hinders the applicability of these models, especially in climate studies. For example, researchers investigating the common trend of hydrologic regimes or air pollution emissions face challenges due to the absence of suitable methods for handling missing observations (Bard et al., 2015). This highlights the second urgent need in our research agenda.

We bridge the gap between empirical requirements and existing theoretical studies by introducing a new toolkit for conducting inference. Specifically, we introduce a local linear dummy variable (LLDV) procedure for estimating parameters and a novel residual-based autoregressive wild bootstrap (AWB) scheme for constructing pointwise intervals and simultaneous bands. Both methods are designed to handle missing observations in the dependent variable without imputation. Furthermore, our framework allows for flexible forms of cross-sectional and serial dependence, as well as heteroskedasticity. The AWB is simple to implement but also consistently estimates the nuisance parameters. We theoretically demonstrate its capability to mimic the pattern of missing observations and the structure of dependence, both cross-sectionally and temporally.

The current paper builds upon multiple pioneering works. First, we refine the LLDV estimation proposed by Li et al. (2011) to allow for missing observations. In environmental applications, missing observations mainly occur in the dependent variable. Therefore, we focus on addressing missing observations solely in the dependent variable, given that the missing pattern in explanatory variables can significantly complicate asymptotic analysis. Second, we extend the previously established asymptotic framework presented in Robinson (2012), Li et al. (2011), Chen et al. (2012), Chen and Huang (2018), and Gao et al. (2020). While these seminal works are valuable, each is characterized by at least one of the following restrictions: exclusion of explanatory variables, assumption of time-constant slope coefficients, condition of independence (cross-sectionally and/or serially), consideration only of strictly stationary data, or a lack of the capability to handle missing values. We relax all these requirements. Notably, the pattern of missing observations, which may arise with dependence along both cross-sectional and temporal dimensions, enter into our asymptotic approximations. A crucial element in deriving our asymptotic results is a new uniform deviation bound for near-epoch processes with cross-sectional dependence. This result extends Li et al. (2012) from strictly stationary time series to potentially nonstationary panels and can be of separate interest for asymptotic analysis in large panels. Finally, the proposed AWB can be traced back to the works of Smeekes and Urbain (2014) and Friedrich et al. (2020). It was initially designed for time series data, specifically for multivariate unit root testing and nonparametric trend analysis. Our simulation study confirms the theoretical finding that it has noteworthy potential in panel applications.

We illustrate the proposed methods through two empirical applications. We first investigate the impact of surface particulate matter air pollution ($PM_{2.5}$) on mortality. Our findings reveal a positive and significant impact of $PM_{2.5}$ on mortality, along with an overall increasing trend in mortality over time. In our second application, we employ the proposed method to investigate common trends in atmospheric ethane in the Northern Hemisphere. Our results indicate that the trend reversal pattern identified in the previous literature using a univariate approach extends to a common global trend.

The paper is organized as follows. Section 2 describes the model and the nonparametric estimation with missing observations in the dependent variable. Section 3 establishes the asymptotic results. Section 4 proposes our autoregressive wild bootstrap. A thorough simulation study is conducted in Section 5. Section 6 presents the empirical applications. Section 7 concludes. The notation used throughout this paper is explained in Appendix A. All proofs, along with additional results from simulation and empirical studies, are presented in the supplemental appendix.

2 The model and estimation

Consider the time-varying trending panel regression model originally proposed by Li et al. (2011):

$$y_{it} = \alpha_i + g_t + \sum_{j=1}^{d} \beta_{t,j} x_{it,j} + e_{it} = \alpha_i + g_t + \boldsymbol{x}_{it}'\boldsymbol{\beta}_t + e_{it}, \quad i = 1, \dots, N, \quad t = 1, \dots, T, \tag{2.1}$$

where y_{it} represents the dependent variable of the i-th cross-sectional unit at time t, α_i captures unobservable time-invariant heterogeneous effects that may correlate with the individual-specific explanatory variables $\boldsymbol{x}_{it} = (x_{it,1}, \dots, x_{it,d})'$, commonly known as fixed effects. As detailed in Section 3, we accommodate deterministic trending behaviors (as well as cross-sectional and serial dependence) in $\boldsymbol{x}_{it}$. The term g_t captures the evolving environment shared among the units in the panel and can be interpreted as a global trend. It may serve as a representative term for unobservable variables, such as global crises, technology growth, and rising environmental awareness (Lin and Reuvers, 2022; Friedrich et al., 2023). Unlike traditional two-way fixed effect approaches that consider g_t as a nuisance parameter and seek ways to eliminate it in estimation, we are interested in analyzing the common trending behavior observed in our data. Furthermore, the slope coefficients are stacked in the vector $\boldsymbol{\beta}_t = (\beta_{t,1}, \dots, \beta_{t,d})'$ and are allowed to vary over time. The error process $\{e_{it}\}$ may exhibit heteroskedasticity, and cross-sectional and serial dependence (Assumptions A3 - A4). The model offers a useful interpretation by decomposing y_{it} into a global component g_t and a local component $\boldsymbol{x}_{it}'\boldsymbol{\beta}_t$, making it popular in empirical studies.

We consider flexible functional forms: $g_t = g(t/T)$, where $g(\cdot) = [0, 1] \to \mathbb{R}$ is an unknown, smooth function. Similarly, let $\boldsymbol{\beta}_t = \boldsymbol{\beta}(t/T)$; $\boldsymbol{\beta}(\cdot) = (\beta_1(\cdot), \dots, \beta_d(\cdot))' : [0, 1] \to \mathbb{R}^d$ be a vector of unknown, smooth functions. Our primary focus is estimating the unknown functions $g(\cdot)$ and $\boldsymbol{\beta}(\cdot)$, constructing their confidence intervals/bands, while accounting for potentially missing observations in the dependent variable. We adopt the asymptotic framework as $(N, T) \to \infty$ jointly.

The theoretical properties of models similar to (2.1) have been explored in the literature, albeit with different emphases or restrictions. The pioneering work by Li et al. (2011) requires cross-sectional independence of regressors and errors. Robinson (2012) focuses on the case where

$\boldsymbol{\beta}_t \equiv \mathbf{0}$ with cross-sectional dependence but requires that the error process is uncorrelated over time and homoscedastic. Chen et al. (2012) and Gao et al. (2020) allow for cross-sectional dependence, but the slope coefficients are not allowed to vary over time. Chen and Huang (2018) and Atak et al. (2023) concentrate on testing whether the parameters are time-varying and/or homogeneous across i. Importantly, all the previous papers do not consider missing data, limiting their direct suitability and applicability to many of our datasets of interest.

2.1 Nonparametric estimation with missing values

We first impose the following common condition for the identification of g_t:

$$\sum_{i=1}^{N} \alpha_i = 0. \tag{2.2}$$

Our models explicitly allow for missing observations in $\{y_{it}\}$. Intuitively, even in the presence of missing observations for some unit i at time t, the observed data for units $j \neq i$ still carry the signal of $(g_t, \boldsymbol{\beta}_t)$. Consequently, pooling information from non-missing units enables consistent estimates of the parameters. We define

$$M_{it} = \mathbb{1}\{y_{it} \text{ is observed}\}, \qquad i = 1 \ldots, N, \quad t = 1, \ldots, T. \tag{2.3}$$

We adapt the local linear dummy variable (LLDV) estimation originally proposed in Li et al. (2011) to accommodate missing observations. The adapted LLDV relies on the following approximation:

$$y_{it}^M \approx M_{it}\alpha_i + \boldsymbol{z}_{it}^M(\tau)'\boldsymbol{\theta}(\tau) + M_{it}e_{it}, \qquad \boldsymbol{\theta}(\tau) = \left(g(\tau), \boldsymbol{\beta}(\tau)', hg^{(1)}(\tau), h\boldsymbol{\beta}^{(1)}(\tau)'\right)', \tag{2.4}$$

where $y_{it}^M = M_{it}y_{it}$, $\boldsymbol{z}_{it}^M(\tau) = M_{it}\boldsymbol{z}_{it}(\tau)$, $\boldsymbol{z}_{it}(\tau) = \left(1, \boldsymbol{x}_{it}', \frac{\tau_t-\tau}{h}, \frac{\tau_t-\tau}{h}\boldsymbol{x}_{it}'\right)'$, and $\tau_t = t/T$. Our LLDV estimator, adjusted for missing data, minimizes the following weighted loss criterion:

$$\widehat{\boldsymbol{\theta}}(\tau) = \arg\min_{\boldsymbol{\theta}(\tau)} \sum_{i=1}^{N}\sum_{t=1}^{T} \left[y_{it}^M - M_{it}\alpha_i - \boldsymbol{z}_{it}^M(\tau)'\boldsymbol{\theta}(\tau)\right]^2 K\left(\frac{\tau_t - \tau}{h}\right), \tag{2.5}$$

subject to the identification condition in Eq. (2.2). In Eq. (2.5), $K(\cdot)$ denotes a kernel function and $h \downarrow 0$ is a bandwidth determining the smoothness of the estimators. Without missing observations, we have $M_{it} = 1$, for all $i = 1, \ldots, N$, $t = 1, \ldots, T$, such that the loss criterion (2.5) reduces to the one given in Li et al. (2011). While we would ideally include missing values in the covariates, the pattern of missing data significantly complicates the analysis. Hence, we consider this as an avenue for future research.

The minimization problem (2.5) is computationally efficient and fast, with a closed-form expression provided in (2.10) below. To illustrate, we require additional notation. More specifically, we can express (2.4) in stacked notation as

$$\boldsymbol{y}_i^M \approx \boldsymbol{m}_i \alpha_i + \boldsymbol{Z}_i^M(\tau)\boldsymbol{\theta}(\tau) + \boldsymbol{e}_i^M, \qquad i = 1, \ldots, N, \tag{2.6}$$

where $\boldsymbol{m}_i = (M_{i1}, \ldots, M_{iT})'$, $\boldsymbol{y}_i^M = \operatorname{diag}(\boldsymbol{m}_i)\, \boldsymbol{y}_i$, $\boldsymbol{y}_i = (y_{i1}, \ldots, y_{iT})'$. Moreover, $\boldsymbol{Z}_i^M(\tau) = \operatorname{diag}(\boldsymbol{m}_i)\, \boldsymbol{Z}_i(\tau)$ with $\boldsymbol{Z}_i(\tau) = (\boldsymbol{z}_{i1}(\tau), \ldots, \boldsymbol{z}_{iT}(\tau))'$, and $\boldsymbol{e}_i^M = \operatorname{diag}(\boldsymbol{m}_i)\,(e_{i1}, \ldots, e_{iT})'$. Let $\boldsymbol{k}_h(\tau) = \left[K\left(\frac{\tau_1 - \tau}{h}\right), \ldots, K\left(\frac{\tau_T - \tau}{h}\right)\right]' \in \mathbb{R}^{T \times 1}$, $\boldsymbol{K}_h(\tau) = \operatorname{diag}[\boldsymbol{k}_h(\tau)]$ be a diagonal matrix with elements $\boldsymbol{k}_h(\tau)$ on the diagonal. In the spirit of the Frisch-Waugh-Lovell Theorem, we introduce a procedure minimizing the weighted loss in (2.5) given the identification condition (2.2).[1]

STEP 1 For each $\tau \in [0, 1]$, project $\boldsymbol{K}_h^{1/2}(\tau)\boldsymbol{Z}_i^M(\tau)$ on $\boldsymbol{K}_h^{1/2}(\tau)\boldsymbol{m}_i\alpha_i$, $i = 1, \ldots, N$, and obtain the residuals $\widetilde{\boldsymbol{Z}}_i^M(\tau)$. It leads to

$$\widetilde{\boldsymbol{Z}}_i^M(\tau) = \boldsymbol{K}_h^{1/2}(\tau)\boldsymbol{Z}_i^M(\tau) - \operatorname{diag}(\boldsymbol{m}_i)\boldsymbol{k}_h^{1/2}(\tau)\boldsymbol{k}_h(\tau)'\left[\nu_{\tau,i}\left(\boldsymbol{Z}_i^M(\tau) - \bar{\boldsymbol{Z}}^M(\tau)\right)\right], \tag{2.7}$$

for $i = 1, \ldots, N$, where

$$\nu_{\tau,i} = \begin{cases} \left(\sum_{t=1}^T K\left(\frac{\tau_t - \tau}{h}\right) M_{it}\right)^{-1}, & \text{if } \sum_{t=1}^T K\left(\frac{\tau_t - \tau}{h}\right) M_{it} \neq 0, \\ 0, & \text{otherwise}, \end{cases} \tag{2.8}$$

and $\bar{\boldsymbol{Z}}^M(\tau) = \omega_\tau^{-1}\sum_{i=1}^N \nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)$. Moreover, $\omega_\tau = \sum_{i=1}^N \nu_{\tau,i}$ if $\nu_{\tau,i} \neq 0$ for some i; $\omega_\tau = 1$ if $\nu_{\tau,i} = 0$ for all i.

STEP 2 Project $\boldsymbol{K}_h^{1/2}(\tau)\boldsymbol{y}_i^M$ on $\boldsymbol{K}_h^{1/2}(\tau)\boldsymbol{m}_i\alpha_i$, $i = 1, \ldots, N$, and obtain the residuals $\widetilde{\boldsymbol{y}}_i^M$. It

[1] Find the MATLAB codes for our estimation and bootstrap methods on https://yiconglin.com/code-and-data/.

leads to

$$\widetilde{\boldsymbol{y}}_i^M = \boldsymbol{K}_h^{1/2}(\tau)\boldsymbol{y}_i^M - \operatorname{diag}(\boldsymbol{m}_i)\boldsymbol{k}_h^{1/2}(\tau)\boldsymbol{k}_h(\tau)' \left[\nu_{\tau,i}\left(\boldsymbol{y}_i^M - \bar{\boldsymbol{y}}^M\right)\right], \tag{2.9}$$

for $i = 1, \dots, N$, where $\bar{\boldsymbol{y}}^M = \omega_\tau^{-1} \sum_{i=1}^N \nu_{\tau,i}\boldsymbol{y}_i^M$.

Step 3 Project $\widetilde{\boldsymbol{y}}_i^M$ on $\widetilde{\boldsymbol{Z}}_i^M(\tau)$, $i = 1, \dots, N$, and obtain $\widehat{\boldsymbol{\theta}}(\tau)$ given by

$$\widehat{\boldsymbol{\theta}}(\tau) = \begin{pmatrix} \widehat{g}(\tau) \\ \widehat{\boldsymbol{\beta}}(\tau) \\ \widehat{hg^{(1)}}(\tau) \\ \widehat{h\boldsymbol{\beta}^{(1)}}(\tau) \end{pmatrix} = \left(\sum_{i=1}^N \widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{Z}}_i^M(\tau)\right)^{-1} \left(\sum_{i=1}^N \widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{y}}_i^M\right). \tag{2.10}$$

Step 4 Given $\widehat{\boldsymbol{\theta}}(\tau)$, $\tau \in [0,1]$, we can obtain the estimates of fixed effects. Specifically, let $\widehat{\boldsymbol{\alpha}} = (\widehat{\alpha}_1, \dots, \widehat{\alpha}_N)' = T^{-1}\sum_{t=1}^T \widehat{\boldsymbol{\alpha}}^\dagger(\tau_t)$ be the estimate of $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)'$, where $\widehat{\boldsymbol{\alpha}}^\dagger(\tau) = \left(\widehat{\alpha}_1^\dagger(\tau), \dots, \widehat{\alpha}_N^\dagger(\tau)\right)'$ with $\widehat{\alpha}_i^\dagger(\tau) = \boldsymbol{k}_h(\tau)'\left[\nu_{\tau,i}\widehat{\boldsymbol{e}}_i^M(\tau)\right]$, and

$$\widehat{\boldsymbol{e}}_i^M(\tau) = \left(\boldsymbol{y}_i^M - \bar{\boldsymbol{y}}^M\right) - \left[\boldsymbol{Z}_i^M(\tau) - \bar{\boldsymbol{Z}}^M(\tau)\right]\widehat{\boldsymbol{\theta}}(\tau), \qquad i = 1 \dots, N. \tag{2.11}$$

Several remarks are in order. First, there may be points τ around which no data is observed in an h-neighborhood in small samples. In such cases, we define $\nu_{\tau,i} = 0$ as seen in Eq. (2.8) in Step 1. Second, the procedure is constructed under the identification condition (2.2), leading to $\sum_{i=1}^N \widehat{\alpha}_i \equiv 0$ for any N and T. Finally, it may be empirically relevant to include observable macro variables that are common to all cross-sectional units in Eq. (2.1). We choose to omit them for illustration purposes, but one could adapt the estimation procedure above by extending $\boldsymbol{z}_{it}(\tau)$ to $\boldsymbol{z}_{it}(\tau) = \left(1, \boldsymbol{x}_{it}', \boldsymbol{\omega}_t', \frac{\tau_t-\tau}{h}, \frac{\tau_t-\tau}{h}\boldsymbol{x}_{it}', \frac{\tau_t-\tau}{h}\boldsymbol{\omega}_t'\right)'$ to estimate the corresponding coefficients, where $\boldsymbol{\omega}_t$ denotes the vector of macroeconomic variables.

3 Asymptotic theory

In the following sections, we present the asymptotic results. We begin by discussing the assumptions necessary for establishing our asymptotic theory in Section 3.1. Several of these assumptions are grounded in a crucial uniform deviation result, which will be introduced in Section 3.2. In the same section, we establish the uniform consistency and asymptotic distribution of the LLDV estimator in Eq. (2.10).

3.1 Assumptions underpinning asymptotic theory

Our asymptotic analysis hinges on the following sets of assumptions.

Assumptions: *A1 The kernel function $K(\cdot)$ is positive, symmetric, Lipschitz continuous, and has compact support $[-1, 1]$ with $\mu_0 \equiv \int_{-1}^{1} K(u)\,\mathrm{d}u = 1$.*

A2 The coefficient function $\boldsymbol{\beta}(\cdot) \in \mathcal{C}^3[0, 1]$, namely $\beta_k(\cdot) \in \mathcal{C}^3[0, 1]$, $k = 1, \ldots, d$. Moreover, the global trend function $g(\cdot) \in \mathcal{C}^3[0, 1]$.

Assumption A1 on the kernel function is satisfied by many commonly adopted kernels, such as the Epanechnikov kernel. Assumption A2 imposes a standard smoothness condition on the functions, as in, e.g., Zhou and Wu (2010, Assumption 6).

The following assumptions impose some structure on error processes and explanatory variables. One of the key assumptions necessary to develop the asymptotic theory involves allowing for a general class of innovation processes, known as near-epoch dependent (NED) processes. These processes accommodate various forms of dependence commonly encountered in econometrics, including linear/nonlinear processes (Davidson, 2002) and strong mixing processes. We extend the definition of NED processes from Lu and Linton (2007) and Li et al. (2012), originally applied to strictly stationary time series, to a panel setting with potential nonstationarity. Refer to Definition A.1 in Appendix A for details. A NED process can be viewed as "approximately" mixing along the time dimension, in the sense that it can be well-approximated by a mixing process.

Assumptions: *A3 Innovations: Define $\boldsymbol{\xi}_{\cdot t} = (\boldsymbol{\xi}'_{1t}, \ldots, \boldsymbol{\xi}'_{Nt})'$, $\boldsymbol{\xi}_{it} = (M_{it}, \varepsilon_{it}, \boldsymbol{\zeta}'_{it})'$. Suppose $\{\boldsymbol{\xi}_{\cdot t},\, t \in \mathbb{Z}\}$ is an α-mixing process with mixing coefficients $\alpha(j) \leq C_\alpha j^{-\varphi_\alpha}$, $0 < C_\alpha < \infty$, $\varphi_\alpha > 2 \vee (2+\delta)/\delta$ for some $\delta > 0$. Further assumptions regarding the elements of $\boldsymbol{\xi}_{\cdot t}$ are* *presented separately in A4 - A6 below.*

A4 *Let* $\boldsymbol{e}_{\bullet t} = (e_{1t}, \ldots, e_{Nt})' = \boldsymbol{\sigma}_t \boldsymbol{\varepsilon}_{\bullet t}$, $\boldsymbol{\varepsilon}_{\bullet t} = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$. *The process* $\{\boldsymbol{\varepsilon}_{\bullet t},\ t \in \mathbb{Z}\}$ *is strictly stationary* α*-mixing and satisfies Assumption A3. Moreover,* $\mathbb{E}(\boldsymbol{\varepsilon}_{\bullet t}) = \mathbf{0}$ *and* $\mathbb{E}\,\|\boldsymbol{\varepsilon}_{\bullet t}\|^{p_0} \leq C < \infty$, $p_0 = 2(2+\delta)$ *for the same* $\delta > 0$ *as specified in Assumption A3. The deterministic term* $\boldsymbol{\sigma}_t = \boldsymbol{\sigma}(t/T) = \operatorname{diag}(\sigma_1(t/T), \ldots, \sigma_N(t/T))$, *where* $\sigma_i(\cdot) : [0,1] \to [\sigma_L, \sigma_U] \subset (0, +\infty)$ *is Lipschitz continuous, i.e., for* $i = 1, \ldots, N$, $|\sigma_i(\tau_1) - \sigma_i(\tau_2)| \leq K_\sigma |\tau_1 - \tau_2|$, $K_\sigma > 0$.

A5 *Explanatory variables:* $\boldsymbol{x}_{it} = \boldsymbol{\chi}_i + \boldsymbol{\ell}(t/T) + \boldsymbol{\nu}_{it}$, *where* $\boldsymbol{\ell}(t/T) = (\ell_1(t/T), \ldots, \ell_d(t/T))'$, $\boldsymbol{\chi}_i = (\chi_{i,1}, \ldots, \chi_{i,d})' \in \mathbb{R}^d$. *Moreover,* $\{\boldsymbol{\chi}_i,\ i = 1, \ldots, N\}$ *is independent of* $\{\boldsymbol{\xi}_{\bullet t},\ t \in \mathbb{Z}\}$. *Let* $p_0 = 2(2+\delta)$ *with the same* $\delta > 0$ *in Assumption A3.*

(*a*) *The local trend function* $\boldsymbol{\ell}(\cdot) \in \mathcal{C}^2[0,1]$, *i.e.,* $l_k(\cdot) \in \mathcal{C}^2[0,1]$, $k = 1, \ldots, d$.

(*b*) *The individual levels* $\boldsymbol{\chi}_i$ *are independent across* $i = 1, \ldots, N$, *where* $\max_{1 \leq i \leq N} \|\boldsymbol{\chi}_i\| = O_p(1)$, $\mathbb{E}(\boldsymbol{\chi}_i) = \mathbf{0}_d$, $\mathbb{E}(\boldsymbol{\chi}_i \boldsymbol{\chi}_i') = \boldsymbol{\Sigma}_\chi$ *is positive semidefinite, and* $\mathbb{E}(\|\boldsymbol{\chi}_i\|^{p_0}) \leq C < \infty$.

(*c*) $\left\{(\boldsymbol{\nu}_{1t}, \ldots, \boldsymbol{\nu}_{Nt})',\ t \in \mathbb{Z}\right\}$ *is strictly stationary with* $\mathbb{E}(\boldsymbol{\nu}_{it}) = \mathbf{0}$ *and* $\mathbb{E}(\|\boldsymbol{\nu}_{it}\|^{p_0}) \leq C < \infty$ *Define* $\boldsymbol{\nu}_{it}^{(m)} = \left(\nu_{it,1}^{(m)}, \ldots, \nu_{it,d}^{(m)}\right)'$. *Let* $\left\{\boldsymbol{\nu}_{it} = (\nu_{it,1}, \ldots, \nu_{it,d})',\ t \in \mathbb{Z}\right\}$ *be NED in* L_{p_0} *with respect to* $\{\boldsymbol{\zeta}_{it},\ t \in \mathbb{Z}\}$, *and* $\psi_{i,p_0}(m) = \sup_{t \in \mathbb{Z}} \mathbb{E}\left\|\boldsymbol{\nu}_{it} - \boldsymbol{\nu}_{it}^{(m)}\right\|^{p_0} \leq d_i^{p_0} m^{-\varphi_\nu}$, *where* $\varphi_\nu > 0$, *and* $\sum_{i=1}^N d_i = O(N^{1/2})$ *with* $d_i \geq 0$. *The process* $\{\boldsymbol{\zeta}_{it},\ t \in \mathbb{Z}\}$ *is strictly stationary* α*-mixing and satisfies Assumption A3. Moreover,* $\lim_{N \to \infty} N^{-1} \sum_{i=1}^N \|\mathbb{E}(\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}') - \boldsymbol{\Sigma}_\nu\| = 0$, *where* $\boldsymbol{\Sigma}_\nu \in \mathbb{R}^{d \times d}$ *is positive definite (p.d.).*

Assumption A3 places conditions on the decaying rate of mixing coefficients. It is relatively weak and guarantees only the absolute summability of autocovariances. As seen later, we do not assume the process of missing observations to be strictly stationary, and thus we do not impose strict stationarity of $\{\boldsymbol{\xi}_{\bullet t}\}$. The deterministic process $\{\boldsymbol{\sigma}_t,\ t \geq 1\}$ in Assumption A4 governs the shape of the volatility of $\{\boldsymbol{e}_{\bullet t}\}$. It permits the error process to exhibit heteroskedasticity, allowing for a wide range of unconditional volatility processes, such as smooth trends and fluctuations. One could consider relaxing the Lipschitz continuity condition on the volatility function $\boldsymbol{\sigma}(\cdot)$ by permitting a finite number of discontinuities, as shown in Cavaliere et al. (2010). In Assumption A5, we assume that the explanatory variables $\boldsymbol{x}_{it}$ can be decomposed into an individual-specific (random) component $\boldsymbol{\chi}_i$, a deterministic trending component $\boldsymbol{\ell}(\cdot)$, and a random component $\boldsymbol{\nu}_{it}$ that is NED in L_{p_0}, where $p_0 > 4$. The process $\boldsymbol{\nu}_{it}$ captures the dependence in regressors across cross-sectional and time dimensions, allowing for various dependence structures over time. The condition $\sum_{i=1}^N d_i = O(N^{1/2})$ constrains the strength of cross-sectional dependence in $\boldsymbol{\nu}_{it}$ and enables us to apply the generic convergence result established in Theorem 1 in the next section.

This specification of $\boldsymbol{x}_{it}$ is of practical relevance since it allows regressors to be trending and seasonal, such as climate variables like $\mathrm{PM}_{2.5}$ and precipitation, or economic variables such as GDP. It is similar to Chen et al. (2012, Eq. (1.2)), but we allow for a more flexible process of $\boldsymbol{\nu}_{it}$ (L_{p_0}-NED) instead of strictly stationary α-mixing. Next, we introduce assumptions regarding the pattern of missing observations.

Assumption: *A6* (*a*) *For $i = 1, \dots, N$, $\{M_{it}, t \in \mathbb{Z}\}$ satisfies Assumption A3, with $\mathbb{E}(M_{it}) = \mathbb{P}(M_{it} = 1) = p_i(t/T) \in [p_L, 1]$, $0 < p_L \leq 1$, where $p_i(\cdot) \in \mathcal{C}^2[0,1]$. Moreover, there exist $\bar{p}(\cdot), \bar{q}(\cdot) \in \mathcal{C}^2[0,1]$ such that $\left|N^{-1}\sum_{i=1}^N p_i(\tau) - \bar{p}(\tau)\right| = O(\phi_{p,N})$ and $\left|N^{-1}\sum_{i=1}^N p_i^{-1}(\tau) - \bar{q}(\tau)\right| = O(\phi_{q,N})$, uniformly in $\tau \in [0,1]$, where $\phi_{p,N} \downarrow 0$, $\phi_{q,N} \downarrow 0$, as $N \to \infty$.*

(*b*) *$\mathbb{E}\left(M_{it}M_{j(t+k)}\right) = \mathcal{R}_{i,j}\big(t/T, (t+k)/T\big)$, $k \geq 0$. Suppose $\mathcal{R}_{i,j}(\cdot,\cdot) : [0,1]^2 \to [0,1]$ is Lipschitz continuous uniformly in i, j. Namely, $|\mathcal{R}_{i,j}(\boldsymbol{\tau}_1) - \mathcal{R}_{i,j}(\boldsymbol{\tau}_2)| \leq K_{\mathcal{R}} \|\boldsymbol{\tau}_1 - \boldsymbol{\tau}_2\|$, where $\boldsymbol{\tau}_1, \boldsymbol{\tau}_2 \in [0,1]^2$, and $K_{\mathcal{R}} > 0$ is independent of i, j.*

Assumption A6 is similar to Friedrich et al. (2020, Assumption 4) for nonparametric trending time series models. The missing process can exhibit weak dependence and vary across cross-sectional units. The missing probability can evolve smoothly over time, as observed in our empirical studies and illustrated in Figure 2 of Friedrich et al. (2020). To facilitate the presentation of our theory, we stipulate in Assumption A6(*a*) that the average proportion of missing/observed data can be approximated by a smooth function. This average proportion is allowed to change over time, as illustrated in Figure 3 (Section 6). As mentioned, $\{M_{it},\, t \in \mathbb{Z}\}$ does not need to be strictly stationary. For instance, M_{it} can be independently (across i and over t) Bernoulli distributed with probability $p\,(t/T)$ for all i. A broad class of generating processes satisfies Assumption A6(*a*). If the processes $\{M_{it},\, t \in \mathbb{Z}\}$ are strictly stationary Markov chains, they are α-mixing with coefficients decreasing to zero, at least, exponentially fast (Bradley, 2005, Theorem 3.1). Therefore, they fulfill Assumptions A3 and A6. Furthermore, a class of dynamic time series binary choice models considered in de Jong and Woutersen (2011) also satisfies these assumptions under suitable conditions. Similar to the missing probability, we allow the cross-sectional moments to exhibit smooth time variation, as indicated in Assumption A6(*b*).

Some moment conditions on cross-sectional dependence and requirements for exogeneity between the error process and regressors are necessary to establish the limiting distribution.

Assumption: *A7* (*a*) *Cross-sectional dependence: For some* $m_0 > 1$, $\sup_{t\in\mathbb{Z}} \mathbb{E}\big| \sum_{i=1}^{N} [M_{it} - \mathbb{E}\,(M_{it})]\big|^{m_0} = O(N^{m_0/2})$; $\sup_{t\in\mathbb{Z}} \mathbb{E}\big\| \sum_{i=1}^{N} \left[M_{it}\boldsymbol{\nu}_{it} - \mathbb{E}\,(M_{it}\boldsymbol{\nu}_{it})\right] \big\|^{q_0} = O\left(N^{q_0/2}\right)$, *where* $q_0 \in (1, p_0]$; $\sup_{t\in\mathbb{Z}} \mathbb{E}\big\| \sum_{i=1}^{N} \left[M_{it}\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}' - \mathbb{E}\,(M_{it}\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}')\right] \big\|^{r_0} = O\left(N^{r_0/2}\right)$, $r_0 \in (1, p_0/2]$; $\sup_{t\in\mathbb{Z}} \mathbb{E}\big\| \sum_{i=1}^{N} w_{it} M_{it}\varepsilon_{it}\boldsymbol{q}_{it}\big\|^{p_0/2} = O\left(N^{p_0/4}\right)$, *where* $\boldsymbol{q}_{it} = (1, \boldsymbol{\chi}_i', \boldsymbol{\nu}_{it}')'$, *and* $w_{it} \in [w_L, w_U] \subset (0, \infty)$ *is a nonrandom sequence,* $i = 1, \ldots, N$, $t = 1, \ldots, T$.

(*b*) *Define* $\boldsymbol{V}_i = (\boldsymbol{\nu}_{i1}, \ldots, \boldsymbol{\nu}_{iT})$. *For* $i, j \in \{1, \ldots, N\}$, $s, t \in \{1, \ldots, T\}$, *the following conditions of exogeneity hold almost surely:* $\mathbb{E}\big(\boldsymbol{\nu}_{it} \,\big|\, M_{it}\big) = \mathbb{E}(\boldsymbol{\nu}_{it})$, $\mathbb{E}\big(\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{js}' \,\big|\, M_{it}, M_{js}\big) = \mathbb{E}(\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{js}')$, $\mathbb{E}\big(\boldsymbol{\chi}_i\boldsymbol{\nu}_{it}' \,\big|\, M_{it}\big) = \mathbb{E}\big(\boldsymbol{\chi}_i\boldsymbol{\nu}_{it}'\big)$, $\mathbb{E}\big(\boldsymbol{\chi}_i\boldsymbol{\nu}_{it}'\boldsymbol{\chi}_j\boldsymbol{\nu}_{jt}' \,\big|\, M_{it}, M_{jt}\big) = \mathbb{E}\big(\boldsymbol{\chi}_i\boldsymbol{\nu}_{it}'\boldsymbol{\chi}_j\boldsymbol{\nu}_{jt}'\big)$. *Moreover,* $\mathbb{E}\big(\varepsilon_{it} \,\big|\, M_{it}, \boldsymbol{\chi}_i\big) = 0$, $\mathbb{E}\big(\varepsilon_{it}\varepsilon_{js} \,|\, \boldsymbol{m}_i, \boldsymbol{m}_j, \boldsymbol{V}_i, \boldsymbol{V}_j\big) = \mathbb{E}\,(\varepsilon_{it}\varepsilon_{js})$, $\mathbb{E}\big(\boldsymbol{\nu}_{it}\varepsilon_{it} \,\big|\, M_{it}\big) = \boldsymbol{0}$, $\mathbb{E}\big(\boldsymbol{\nu}_{js}\varepsilon_{it}\varepsilon_{js} \,\big|\, M_{it}, M_{js}\big) = \boldsymbol{0}$, *and* $\mathbb{E}\big(\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{js}'\varepsilon_{it}\varepsilon_{js} \,\big|\, M_{it}, M_{js}\big) = \mathbb{E}\big(\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{js}'\big)\mathbb{E}\big(\varepsilon_{it}\varepsilon_{js}\big)$.

(*c*) $\sum_{i,j=1}^{N} \sum_{k=1}^{T} |\mathbb{E}\,(\varepsilon_{i1}\varepsilon_{jk})| = O(N)$.

(*d*) *For any* $\tau \in (0, 1)$, *there exist p.d. matrices* $\boldsymbol{\Lambda}_\varepsilon(\tau) \in \mathbb{R}^{2\times 2}$ *and* $\boldsymbol{\Lambda}_{\nu\varepsilon}(\tau) \in \mathbb{R}^{d\times d}$ *such that, as* $N \to \infty$, $N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N} \mathcal{R}_{i,j}(\tau,\tau)\boldsymbol{Q}_{i,j}(\tau)\sigma_i(\tau)\sigma_j(\tau)\Omega_\varepsilon(i,j) \to \boldsymbol{\Lambda}_\varepsilon(\tau) := \begin{pmatrix} \Lambda_{\varepsilon,11}(\tau) & \Lambda_{\varepsilon,12}(\tau) \\ \Lambda_{\varepsilon,21}(\tau) & \Lambda_{\varepsilon,22}(\tau) \end{pmatrix}$, $N^{-1}\sum_{i=1}^{N}\sum_{j=1}^{N} \mathcal{R}_{i,j}(\tau,\tau)\sigma_i(\tau)\sigma_j(\tau)\boldsymbol{\Omega}_{\nu\varepsilon}(i,j) \to \boldsymbol{\Lambda}_{\nu\varepsilon}(\tau)$, *and more-*

over, $N^{-1}\sum_{i=1}^{N}\mathcal{R}_{i,i}(\tau,\tau)\sigma_i^2(\tau)\Omega_\varepsilon(i,i) \to \bar{\Lambda}_\varepsilon(\tau)$, *where* $\mathcal{R}_{i,j}(\cdot,\cdot)$ *is defined in A6,* $\boldsymbol{Q}_{i,j}(\tau) = \begin{pmatrix} [p_i(\tau)p_j(\tau)]^{-1} & [p_i(\tau)]^{-1} \\ [p_j(\tau)]^{-1} & 1 \end{pmatrix}$, $\Omega_\varepsilon(i,j) = \sum_{s=-\infty}^{\infty}\mathbb{E}\left(\varepsilon_{it}\varepsilon_{j(t+s)}\right)$, *and finally,* $\boldsymbol{\Omega}_{\nu\varepsilon}(i,j) = \sum_{s=-\infty}^{\infty}\mathbb{E}\big(\varepsilon_{it}\boldsymbol{\nu}_{it}\varepsilon_{j(t+s)}\boldsymbol{\nu}'_{j(t+s)}\big)$.

The moment conditions in Assumption A7(a) characterize the cross-sectional dependence. They are commonly used in the panel model literature, see, e.g., Bai (2009, Assumption C), Chen et al. (2012, Assumption A2), Corradi and Swanson (2014, Assumption AN1), and Gonçalves and Perron (2014, Assumption 2). Under certain conditions, these requirements can easily be verified. For instance, if $\operatorname{cov}(M_{it}, M_{jt}) = 0$, $i \neq j$, we have $\sup_{t\in\mathbb{Z}}\mathbb{E}\left|\sum_{i=1}^{N}[M_{it} - \mathbb{E}(M_{it})]\right|^2 = \sup_{t\geq 1}\sum_{i=1}^{N}\mathbb{V}\mathrm{ar}(M_{it}) \leq N/4$ fulfilling the first term in A7(a) with $m_0 = 2$. Similarly, considering the element $\boldsymbol{\chi}_i$ in $\boldsymbol{q}_{it}$, by the multinomial theorem, for $k = 1,\ldots,d$, and utilizing the independence of $\{\boldsymbol{\chi}_i\}$, it is straightforward to obtain $\sup_{t\in\mathbb{Z}}\left\{\mathbb{E}\big|\sum_{i=1}^{N}w_{it}M_{it}\chi_{i,k}\varepsilon_{it}\big|^4\right\}^{1/4} = \sup_{t\in\mathbb{Z}}\left\{\sum_{i=1}^{N}\mathbb{E}\big(w_{it}M_{it}\chi_{i,k}\varepsilon_{it}\big)^4 + 6\sum_{i=1}^{N-1}\sum_{\ell=i+1}^{N}\mathbb{E}\big(w_{it}w_{\ell t}M_{it}M_{\ell t}\chi_{i,k}^2\chi_{\ell,k}^2\varepsilon_{it}^2\varepsilon_{\ell t}^2\big)\right\}^{1/4} \leq CN^{1/2}$, along with Assumptions A4, A5, and the Cauchy-Schwarz inequality. Alternatively, one may replace these conditions by imposing α-mixing conditions on the cross-sectional dimension. Similar moment conditions can then be obtained using Rosenthal-type inequalities for α-mixing processes (Shao and Yu, 1996, Theorem 4.1). Assumption A7(b) provides standard conditions on exogeneity and is clearly weaker than the often-imposed assumption of independence, as seen in, for example, Sun et al. (2009). Assumption A7(c) resembles, e.g., Dong et al. (2015, Assumption 1) and Gao et al. (2020, Assumption 2). Assumption A7(d) ensures that the long-run covariance matrix exists and is well-defined, which is standard in the literature, see, e.g., Chen et al. (2012, Assumption A4), Chen and Huang (2018, Assumption A.4).

Assumption: *A8 Suppose* $2\varphi_\nu \geq p_0\varphi_\alpha$ *and define* $\eta_{min} = \phi(m_0 \wedge q_0 \wedge r_0)$, *where* m_0, q_0, r_0, p_0 *are defined in Assumption A7. Let* $\phi(x) = \dfrac{(1-1/x)\varphi_\alpha}{2+(1+1/x)\varphi_\alpha}$ *and* $\varpi(x) = \dfrac{4(1+\varphi_\alpha/x)}{2+(1+1/x)\varphi_\alpha}$, $x > 1$, *where* φ_α *is given in Assumption A3. The bandwidth* $h \equiv h(N,T)$ *satisfies*

$$\max\left\{h, \frac{1}{Th^2}, \frac{\ln(NT)}{(NT)^{\eta_{min}}h}, \frac{N^{\varpi(p_0/2)}\ln(NT)}{T^{\phi(p_0/2)}h}\right\} \to 0, \qquad (N,T)\to\infty. \tag{3.1}$$

Moreover, $\left[N/(NT)^{\phi(\kappa)}\right]^{\left(2\varphi_\nu/[(p_0\varphi_\alpha)(1-1/\kappa)]-1\right)/2} h^{1/2} = O\left(\sqrt{\ln(NT)}\right)$, $\kappa \in \{m_0, q_0, r_0, p_0/2\}$.

Finally, Assumption A8 serves as a technical condition on the bandwidth parameter, offering guidance on its practical selection. Depending on the desired strength of conditions imposed on cross-sectional and serial dependence, one could choose $h \sim (NT)^{-c}$ for some $c > 0$ in practice. The condition below Eq. (3.1) corresponds to (3.4) in Theorem 1 in the next section. Since ϕ is a positive-valued and strictly increasing function, it suffices to fulfill this condition by requiring $N/(NT)^{\eta_{min}} \leq C < \infty$, given $2\varphi_\nu \geq p_0\varphi_\alpha$.

3.2 Uniform consistency and the limiting distribution

To develop our asymptotic results for $\widehat{\boldsymbol{\theta}}(\cdot)$ under cross-sectional and serial dependence, we first establish a general uniform convergence result that ensures uniform consistency. This result is noteworthy in its own right, and we present it in Theorem 1.

Theorem 1 *For $p > 1$ and $i = 1, \ldots, N$, let $\mathbb{E}|Y_{it}|^p \leq C < \infty$, and suppose $\{Y_{it},\, t \in \mathbb{Z}\}$ is NED in L_p with respect to $\{\eta_{it},\, t \in \mathbb{Z}\}$, where*

$$\psi_{i,p}(m) = \sup_{t\in\mathbb{Z}} \mathbb{E}\left|Y_{it} - Y_{it}^{(m)}\right|^p \leq d_i^p m^{-\lambda}, \qquad d_i \geq 0, \quad \lambda > 0. \tag{3.2}$$

Suppose $\left\{\boldsymbol{\eta}_t = \left(\eta_{1t}, \eta_{2t}, \ldots, \eta_{Nt}\right)',\, t \in \mathbb{Z}\right\}$ is an α-mixing process (possibly nonstationary) with mixing coefficients $\alpha(j) \leq Aj^{-\beta}$ for some $0 < A, \beta < \infty$. Moreover, assume the kernel function $K(\cdot)$ is positive, symmetric, Lipschitz continuous, and has compact support $[-u_0, u_0]$, $u_0 > 0$, with $\int_{-u_0}^{u_0} K(u)\,\mathrm{d}u = 1$. For $x > 1$, define $\phi_\beta(x) = \dfrac{(1-1/x)\beta}{2+(1+1/x)\beta}$. For some $1 < q \leq p$, if

$$\sup_{t\in\mathbb{Z}} \mathbb{E}\left|\sum_{i=1}^{N} [Y_{it} - \mathbb{E}(Y_{it})]\right|^q = O(N^{q/2}), \qquad \sum_{i=1}^{N} d_i = O(N^{1/2}), \tag{3.3}$$

and

$$\left[\frac{N}{(NT)^{\phi_\beta(q)}}\right]^{\left(2\lambda/[p\beta(1-1/q)]-1\right)/2} h^{1/2} = O\left(\sqrt{\ln(NT)}\right), \tag{3.4}$$

then we have

$$\sup_{\tau\in[0,1]}\left|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\frac{\tau_t-\tau}{h}\right)^k K\left(\frac{\tau_t-\tau}{h}\right)\left[Y_{it}-\mathbb{E}\left(Y_{it}\right)\right]\right| = O_p\left(\sqrt{\frac{\ln(NT)}{(NT)^{\phi_\beta(q)}h}}\right), \tag{3.5}$$

where $k\geq 0$ is an integer.

Differing from our concentration probability (3.5), Gao et al. (2023) recently established a central limit theorem (CLT) for (parametric) panel data models that consider both serial correlation and cross-sectional dependence using the notion of physical dependence. While Yao and Jiang (2012) explore panel extensions of Hoeffding's inequality, similar to our result, they require cross-sectional independence. Our Theorem 1 can be viewed as an extension of the uniform convergence results found in Hansen (2008), Jiang (2009), Kristensen (2009), and Li et al. (2012) from (stationary/nonstationary) time series to possibly nonstationary panel settings. Similar to ϕ in Assumption A8, ϕ_β is positive-valued and strictly increasing. It reflects a tradeoff between the strength of the assumptions on dependence and moment conditions and the speed of convergence. The value of $\phi_\beta(q)$ can be made close to 1 by, for instance, considering $2\lambda = p\beta$ and an α-mixing process with an exponentially decaying rate ($\beta\to\infty$), as well as a sufficiently large value of q, but it remains strictly smaller than 1. For $N=1$, the rate in Eq. (3.5) yields $\sqrt{\ln(T)/(T^{\phi_\beta(q)}h)}$. This rate is slower than the optimal rate achievable in the time series literature, which is $\sqrt{\ln(T)/(Th)}$. Thus, we consider the rate in Eq. (3.5) to be sub-optimal.

Building upon Theorem 1, we are now ready to establish the first set of asymptotic results. We will first provide the results and then offer further comments below Corollary 1. The following proposition demonstrates the uniform consistency of our estimator under missing observations.

Proposition 1 *Recall ϕ from Assumption A8. Under Assumptions A1 - A8, we obtain*

$$\sup_{\tau\in[0,1]}\left\|\widehat{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}(\tau)\right\| = O_p\left(h^2+\sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}h}}\right), \qquad (N,T)\to\infty. \tag{3.6}$$

Next, we establish the pointwise limiting distribution. A clear exposition of the result requires further notation.

(a) Quantities associated with $K(\cdot)$: $\mu_k = \int_{-1}^{1} u^k K(u)\,\mathrm{d}u$, $\nu_k = \int_{-1}^{1} u^k K^2(u)\,\mathrm{d}u$, $k \geq 0$;

(b) Bias term: $\boldsymbol{b}(\tau) = \frac{1}{2}\begin{pmatrix} \mu_2 \boldsymbol{\Upsilon}^{(2)}(\tau) \\ \mathbf{0} \end{pmatrix} + o_p(1)$, where $\boldsymbol{\Upsilon}(\tau) = (g(\tau), \boldsymbol{\beta}(\tau)')'$;

(c) Short/long-run covariance: $\boldsymbol{\Phi}(\tau) = \mathrm{diag}\left[\boldsymbol{\Phi}_1(\tau), \boldsymbol{\Phi}_2(\tau)\right]$, $\boldsymbol{\Phi}_{\nu\varepsilon}(\tau) = \mathrm{diag}\left[\boldsymbol{\Phi}_{\nu\varepsilon,1}(\tau), \boldsymbol{\Phi}_{\nu\varepsilon,2}(\tau)\right]$, where $\boldsymbol{\Phi}_1(\tau) = \bar{p}(\tau)\begin{pmatrix} 0 & \\ & \boldsymbol{\Sigma}_\nu \end{pmatrix} + \left[\bar{q}(\tau)\right]^{-1}\begin{pmatrix} 1 \\ \boldsymbol{\ell}(\tau) \end{pmatrix}\begin{pmatrix} 1 & \boldsymbol{\ell}(\tau)' \end{pmatrix}$, $\boldsymbol{\Phi}_2(\tau) = \mu_2 \bar{p}(\tau)\begin{pmatrix} 1 & \boldsymbol{\ell}(\tau)' \\ \boldsymbol{\ell}(\tau) & \boldsymbol{\Sigma}_\chi + \boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)' + \boldsymbol{\Sigma}_\nu \end{pmatrix}$, $\boldsymbol{\Phi}_{\nu\varepsilon,1}(\tau) = \nu_0\begin{pmatrix} 0 & \\ & \boldsymbol{\Lambda}_{\nu\varepsilon}(\tau) \end{pmatrix} + \nu_0 \Lambda_{\varepsilon,11}(\tau)\left[\bar{q}(\tau)\right]^{-2}\begin{pmatrix} 1 \\ \boldsymbol{\ell}(\tau) \end{pmatrix}\begin{pmatrix} 1 & \boldsymbol{\ell}(\tau)' \end{pmatrix}$, $\boldsymbol{\Phi}_{\nu\varepsilon,2}(\tau) = \nu_2\begin{pmatrix} 0 & \\ & \bar{\Lambda}_\varepsilon(\tau)\boldsymbol{\Sigma}_\chi + \boldsymbol{\Lambda}_{\nu\varepsilon}(\tau) \end{pmatrix} + \nu_2 \Lambda_{\varepsilon,22}(\tau)\begin{pmatrix} 1 \\ \boldsymbol{\ell}(\tau) \end{pmatrix}\begin{pmatrix} 1 & \boldsymbol{\ell}(\tau)' \end{pmatrix}$.

Theorem 2 (Pointwise asymptotic distribution) *Under Assumptions A1 - A8, for any fixed $\tau \in (0,1)$, as $(N,T) \to \infty$,*

$$\sqrt{NTh}\left(\widehat{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}(\tau) - h^2 \boldsymbol{b}(\tau)\right) \xrightarrow{d} \mathcal{N}\Big(\mathbf{0}, \boldsymbol{\Phi}(\tau)^{-1}\boldsymbol{\Phi}_{\nu\varepsilon}(\tau)\boldsymbol{\Phi}(\tau)^{-1}\Big). \tag{3.7}$$

Theorem 2 offers the first asymptotic result in the literature that gives an approximation in trending panels with missing observations and flexible dependence structures. Building on Proposition 1 and Theorem 2, we can establish the consistency of the estimated fixed effects.

Corollary 1 *Recall $\widehat{\alpha}_i = T^{-1}\sum_{t=1}^{T}\widehat{\alpha}_i^\dagger(\tau_t)$, $i = 1, \ldots, N$, where $\widehat{\alpha}_i^\dagger(\cdot)$ is defined above Eq. (2.11). Under Assumptions A1 - A8, as $(N,T) \to \infty$,*

$$\max_{1 \leq i \leq N} |\widehat{\alpha}_i - \alpha_i| = O_p\left(\max\left\{h^2, \sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}h}}, \sqrt{\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}}\right\}\right). \tag{3.8}$$

We now discuss some implications of these results. First, the second-order bias $\boldsymbol{b}(\tau)$ in Theorem 2 arises from the Taylor approximation (2.6), a standard result in the nonparametric literature. Second, the pointwise asymptotic distribution depends on various nuisance parameters such as $\boldsymbol{b}(\tau)$ and the long-run covariance matrix, which, in turn, depends on the local trends $\boldsymbol{\ell}(\cdot)$, the pattern of missing observations, and some second-order moment terms. Therefore, conducting inference based on the asymptotic distribution requires the estimation of these terms, which is far from straightforward in practice. While a jackknife procedure may be employed to eliminate bias, estimating long-run covariance matrices remains unavoidable and highly challenging. The estimation of these nuisance parameters demands careful selection of tuning parameters, substantially affecting the performance of asymptotic inference (Friedrich and Lin, 2022). For this reason, we propose in the next section a residual-based bootstrap procedure as an alternative method to construct confidence intervals and bands. The proposed bootstrap method automatically reproduces the nuisance parameters, without requiring a cumbersome selection of tuning parameters. Note that the uniform convergence results in Proposition 1 and Corollary 1 are key stepping stones for establishing the theoretical validity of our bootstrap method.

4 Bootstrap inference

Asymptotic inference based on Theorem 2 requires the estimation of various nuisance parameters. This challenge can be conveniently circumvented by employing a bootstrap method. However, a simple univariate, naive wild bootstrap, as commonly employed in empirical studies (for instance, Liddle et al., 2020), is not valid in the presence of cross-sectional and/or serial dependence. We face three main challenges: (i) serial dependence (and heteroskedasticity), (ii) cross-sectional dependence, and (iii) missing observations. To accommodate serial dependence and heteroskedasticity, a sieve wild bootstrap scheme (Smeekes and Taylor, 2012) can be implemented to obtain bootstrap samples separately for each unit in the panel. However, this neglects and eliminates potential cross-sectional dependence in the panel. To address this, we bootstrap the residuals $\widehat{e}_{it}$ jointly by stacking them in a vector over i, namely $\widehat{\boldsymbol{e}}_{\boldsymbol{\cdot} t} = \left(\widehat{e}_{1t}, \ldots, \widehat{e}_{Nt}\right)'$. Then, one can perform a multivariate sieve wild bootstrap using $\left\{\widehat{\boldsymbol{e}}_{\boldsymbol{\cdot} t}\right\}$ to replicate serial dependence and heteroskedasticity. However, the number of parameters will quickly increase as N increases. A potential remedy to reduce the number of parameters might entail imposing certain factor structures on error processes $\{\boldsymbol{e}_{\boldsymbol{\cdot} t}\}$ and subsequently applying the techniques outlined in, for instance, Trapani (2013) and Gonçalves and Perron (2020). However, extending this approach to accommodate missing observations is not straightforward.

In Section 4.1, we propose an Autoregressive Wild Bootstrap (AWB) procedure that correctly captures cross-sectional and serial dependence, and accommodates heteroskedasticity. Importantly, the procedure allows for missing observations. The AWB was originally introduced in Smeekes and Urbain (2014) for multivariate unit root testing and has demonstrated superior performance in modeling nonparametric trends in time series with missing observations, as highlighted in Friedrich et al. (2020). In Section 4.2, we discuss obtaining confidence intervals and provide a theoretical justification.

4.1 The autoregressive wild bootstrap

The core concept behind our AWB is to incorporate an autoregressive scalar series into the stacked residuals to capture heteroskedasticity and serial dependence in the error processes. By stacking residuals, we preserve the cross-sectional dependence pattern, while the inclusion of an autoregressive series accounts for the serial dependence and heteroskedasticity. The bootstrap algorithm involves five key steps:

Step 1 Let $\widetilde{\alpha}_i$, $\widetilde{g}(\cdot)$, and $\widetilde{\boldsymbol{\beta}}(\cdot)$, be the adapted LLDV estimates described in Section 2.1, but using a larger bandwidth $\tilde{h} > h$. Obtain residuals

$$\widetilde{e}_{it} = M_{it}\left(y_{it} - \widetilde{\alpha}_i - \widetilde{g}(t/T) - \boldsymbol{x}_{it}'\widetilde{\boldsymbol{\beta}}(t/T)\right), \qquad i = 1,\dots,N, \quad t = 1,\dots,T. \tag{4.1}$$

Step 2 For $\gamma \in (0,1)$, generate a scalar sequence $\nu_1^*,\dots,\nu_T^*$ as i.i.d. $\mathcal{N}(0,1-\gamma^2)$ and let $\xi_t^* = \gamma\xi_{t-1}^* + \nu_t^*$, $t = 2,\dots,T$, where $\xi_1^* \sim \mathcal{N}(0,1)$.

Step 3 For $i = 1,\dots,N$, $t = 1,\dots,T$, calculate the bootstrap errors $e_{it}^* = M_{it}\xi_t^*\widetilde{e}_{it}$, and generate the bootstrap observations by

$$y_{it}^* = M_{it}\left(\widetilde{\alpha}_i + \widetilde{g}(t/T) + \boldsymbol{x}_{it}'\widetilde{\boldsymbol{\beta}}(t/T) + e_{it}^*\right), \tag{4.2}$$

where $\widetilde{\alpha}_i$, $\widetilde{g}(\cdot)$, and $\widetilde{\boldsymbol{\beta}}(\cdot)$ are the same estimates given in Step 1.

STEP 4 Using $\{(y_{it}^*, \boldsymbol{x}_{it}),\, i = 1, \ldots, N,\, t = 1, \ldots, T\}$, construct the bootstrap adapted LLDV estimates $(\widehat{\alpha}_i^*, \widehat{g}^*(\cdot), \widehat{\boldsymbol{\beta}}^*(\cdot))$ with the same bandwidth h as used for the original estimates.

STEP 5 Repeat STEP 2 to STEP 4 B times, and let

$$\widehat{q}_{j,\alpha}(\tau) = \inf\left\{u \in \mathbb{R} : \mathbb{P}^*\left(\widehat{\beta}_j^*(\tau) - \widetilde{\beta}_j(\tau) \le u\right) \ge \alpha\right\}, \qquad j = 1, \ldots, d, \tag{4.3}$$

denote the 100αth percentile of the B centered bootstrap statistics $\widehat{\beta}_j^*(\tau) - \widetilde{\beta}_j(\tau)$, similarly for $\widehat{g}^*(\tau) - \widetilde{g}(\tau)$, $\tau \in (0, 1)$. These bootstrap quantiles shall be used to construct confidence intervals/bands.

In STEP 1, a larger bandwidth $\tilde{h}$ is used to produce an oversmoothed estimate. This approach is similar to the time series setting considered in Friedrich and Lin (2022), and ensures that the asymptotic bias is consistently estimated by the bootstrap; see Remark 2 below. Bootstrap performance is not sensitive to the choice of $\tilde{h}$. The parameter γ in STEP 2 accounts for both serial dependence and heteroskedasticity. Its interpretation is akin to the block length in block bootstrap methods, representing a tradeoff between capturing more dependence and allowing for more variation in the bootstrap samples (Smeekes and Urbain, 2014). In theory (Assumption B1), we assume $\gamma = \theta^{1/\ell}$, where $\theta \in (0, 1)$ and $\ell = \ell_{NT} > 0$ is some sequence that diverges to ∞ as the sample sizes grow. Further discussion on the role of γ is provided in Remark 1. It is worth noting that the normality of ν_t^* (and ξ_1^*) in STEP 2 is not necessary for obtaining bootstrap consistency. We only require ν_t^* to be i.i.d. conditionally on the original sample, with $\mathbb{E}^*(\nu_t^*) = 0$, $\mathbb{E}^*(\nu_t^*)^2 = 1 - \gamma^2$, and $\mathbb{E}^*(\nu_t^*)^4 < \infty$. Finally, the bootstrap observations corresponding to $M_{it} = 0$ are artificially set to zero in our procedure, as they do not play a role due to the construction of our loss function in (2.5). The missing pattern is automatically preserved without any modifications.

Remark 1 *We briefly discuss the role of the parameter γ in mimicking the asymptotic long-run variance. To illustrate, consider a simple time series regression model without missing observations: $Y_t = X_t\beta + e_t$, $t = 1, \ldots, T$, where $\hat{\beta}$ is the OLS estimator of the parameter β, and $\hat{\beta}^*$ is the bootstrap counterpart using the AWB. To establish bootstrap consistency, we shall establish the distribution* *of $\sqrt{T}(\hat{\beta}^* - \hat{\beta}) = \left(T^{-1}\sum_{t=1}^T X_t^2\right)^{-1}\left(T^{-1/2}\sum_{t=1}^T X_t\widehat{e}_t\xi_t^*\right)$, conditionally on the original sample. As a result, $T^{-1/2}\sum_{t=1}^T X_t\widehat{e}_t\xi_t^*$ must consistently estimate the limiting variance of $T^{-1/2}\sum_{t=1}^T X_t e_t$. Let $\gamma = \theta^{1/\ell}$ for some $\theta \in (0,1)$, where $\ell = \ell_T \to \infty$, as $T \to \infty$, is some positive sequence. We define the kernel function $k(\cdot)$ in de Jong and Davidson (2000, Theorem 2.1) as $k(x) = \theta^{|x|}$. Under some regularity conditions, it is straightforward to obtain that:*

$$\begin{aligned}
&\mathbb{E}^*\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T X_t\widehat{e}_t\xi_t^*\right)^2 = \frac{1}{T}\sum_{t=1}^T\sum_{s=1}^T X_tX_s\widehat{e}_t\widehat{e}_s\mathbb{E}^*\left(\xi_t^*\xi_s^*\right) = \frac{1}{T}\sum_{t=1}^T\sum_{s=1}^T X_tX_s\widehat{e}_t\widehat{e}_s\gamma^{|t-s|}\\
&\approx \frac{1}{T}\sum_{t=1}^T\sum_{s=1}^T X_tX_se_te_s\gamma^{|t-s|} = \frac{1}{T}\sum_{t=1}^T\sum_{s=1}^T X_tX_se_te_sk\left(\frac{t-s}{\ell}\right) \approx \frac{1}{T}\sum_{t=1}^T\sum_{s=1}^T \mathbb{E}\left(X_tX_se_te_s\right).
\end{aligned}$$

The term $T^{-1}\sum_{t=1}^T\sum_{s=1}^T \mathbb{E}\left(X_tX_se_te_s\right)$ will asymptotically converge to the desired long-run variance. This example highlights the role of γ as a kernel in estimating long-run variances.

4.2 Confidence intervals

For any $\tau \in (0,1)$, asymptotic pointwise confidence intervals $I_{j,N,T,\alpha}(\tau)$ for $\beta_j(\tau)$ are constructed to satisfy $\liminf_{T\to\infty,N\to\infty}\mathbb{P}\Big(\beta_j(\tau) \in I_{j,N,T,\alpha}(\tau)\Big) \geq 1-\alpha$. In other words, $I_{j,N,T,\alpha}(\tau)$ is statistically valid for a fixed time point $\tau \in (0,1)$. By utilizing $\widehat{q}_{j,\alpha}(\tau)$ in (4.3), $j = 1\ldots,d$, one can immediately obtain pointwise (equal-tailed) bootstrap confidence intervals with the level of $1-\alpha$ as follows:

$$I_{j,N,T,\alpha}^{P*}(\tau) = \left[\widehat{\beta}_j(\tau) - \widehat{q}_{j,1-\alpha/2}(\tau),\ \widehat{\beta}_j(\tau) - \widehat{q}_{j,\alpha/2}(\tau)\right], \qquad \tau \in (0,1). \tag{4.4}$$

The construction for $g(\cdot)$ is similar.

Now we justify the AWB-based confidence intervals by demonstrating that the bootstrap estimator asymptotically replicates the distribution of the proposed LLDV estimator in Section 2.1. This requires the following additional assumptions.

Assumptions: *B1 Suppose $\gamma = \theta^{1/\ell}$ for some $\theta \in (0,1)$, where $\ell = \ell_{NT} > 0$ satisfying*

$$\max\left\{N\ell^2\tilde{h}^4, \frac{\ell^2}{N^{\varpi(p_0/2)/2}}\frac{N^{\varpi(p_0/2)}\ln(NT)}{T^{\phi(p_0/2)}\tilde{h}}, \frac{N\ell^2}{(TN^2)^{\phi(p_0)-\phi(p_0/2)}}\frac{N^{\varpi(p_0/2)}\ln(NT)}{T^{\phi(p_0/2)}\tilde{h}}, \frac{1}{\ell}, \frac{\ell}{Th}\right\} \to 0,$$

as $(N,T)\to\infty$, where $\tilde{h}$ is the oversmoothing parameter used in the AWB (Step 1).

B2 Suppose $\sum_{i,j,m=1}^{N}\left(\sum_{k=1}^{T}|\mathbb{E}\left(\varepsilon_{i1}\varepsilon_{jk}\right)|\right)\left(\sum_{k=1}^{T}|\mathbb{E}\left(\varepsilon_{i1}\varepsilon_{mk}\right)|\right)=o(N^2)$*. Moreover,*

$$\sup_{t\geq 1}\mathbb{E}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sigma_i(\tau_t)M_{it}\varepsilon_{it}\left(\boldsymbol{\nu}_{it}-\boldsymbol{\nu}_{it}^{(m)}\right)\right\|^2=O\left(m^{-\varsigma}\right),\qquad \varsigma>1. \tag{4.5}$$

B3 Choose $\tilde{h}=\tilde{h}(N,T)$ *such that* $\max\left\{\tilde{h},NTh^7,NTh\tilde{h}^4,(NT)^{\varpi(p_0/2)/2}\ln(NT)h/\tilde{h}\right\}\to 0$*, as* $(N,T)\to\infty$*.*

We first comment on Assumption B1. By the identities $1-\phi(p_0/2)=\varpi(p_0/2)/2$ and $\phi(p_0)-\phi(p_0/2)=[\varpi(p_0/2)-\varpi(p_0)]/2$, Assumption B1 implies that $N^{1/2}\ell\big\{\max_{1\leq i\leq N}|\widetilde{\alpha}_i-\alpha_i|+\sup_{\tau\in[0,1]}\big\|\widetilde{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}(\tau)\big\|\big\}=o_p(1)$. Moreover, note that $\tilde{h}>h$ by assumption. Under Assumption A8, we have $N^{\varpi(p_0/2)}\ln(NT)/\big(T^{\phi(p_0/2)}\tilde{h}\big)\to 0$. Therefore, a key requirement to maintain the rates in Assumption B1 is to ensure that ℓ does not grow too rapidly. It indicates that when T is not sufficiently large, ℓ cannot be too large. Otherwise, estimation errors such as those from the fixed effects estimates in the cross-sectional dimension could accumulate quickly, leading to bootstrap inconsistency. Specifically, the rapid growth of ℓ hinders the consistent estimation of the long-run variance. For this reason, we recommend choosing a relatively small value for γ in finite samples. Assumption B2 aligns with the spirit of Assumptions A7*(a)* and A7*(c)*. Finally, Assumption B3 establishes conditions on the rate of the oversmoothing parameter such that the AWB correctly captures the second-order bias (Remark 2).

The following theorem indicates that the AWB consistently mimics the asymptotic distribution presented in Theorem 2.

Theorem 3 (Pointwise bootstrap validity) *Let* $\widehat{\boldsymbol{\theta}}^*(\cdot)$ *be the bootstrap counterpart of* $\widehat{\boldsymbol{\theta}}(\cdot)$ *and* $\widetilde{\boldsymbol{\theta}}(\tau)$ *be the oversmoothed estimate of* $\boldsymbol{\theta}(\cdot)$ *in Section 4.1. Under Assumptions A1 - A8, B1 - B3, for any fixed* $\tau\in(0,1)$*, as* $(N,T)\to\infty$*,*

$$\sqrt{NTh}\left(\widehat{\boldsymbol{\theta}}^*(\tau)-\widetilde{\boldsymbol{\theta}}(\tau)-h^2\boldsymbol{b}(\tau)\right)\xrightarrow{d^*}_p\mathcal{N}\Big(\mathbf{0},\boldsymbol{\Phi}(\tau)^{-1}\boldsymbol{\Phi}_{\nu\varepsilon}(\tau)\boldsymbol{\Phi}(\tau)^{-1}\Big). \tag{4.6}$$

Remark 2 *We can gain insight into the conditions related to the oversmoothing parameter* $\tilde{h}$ *in*

STEP 1 *of the AWB by examining the expression:* $\sqrt{NTh}\big(\widehat{\boldsymbol{\theta}}^*(\tau)-\widetilde{\boldsymbol{\theta}}(\tau)-h^2\boldsymbol{b}(\tau)\big)=\sqrt{NTh}\big[\widehat{\boldsymbol{\theta}}^*(\tau)-\mathbb{E}^*\big(\widehat{\boldsymbol{\theta}}^*(\tau)\big)\big]+\sqrt{NTh}\big[\mathbb{E}^*\big(\widehat{\boldsymbol{\theta}}^*(\tau)\big)-\widetilde{\boldsymbol{\theta}}(\tau)-h^2\boldsymbol{b}(\tau)\big]$. *The first term* $\sqrt{NTh}\big[\widehat{\boldsymbol{\theta}}^*(\tau)-\mathbb{E}^*\big(\widehat{\boldsymbol{\theta}}^*(\tau)\big)\big]$ *appears to mimic the stochastic variation and converges to the distribution in Eq.* (4.6)*. To achieve bootstrap consistency, we require the second term* $\sqrt{NTh}\big[\mathbb{E}^*\big(\widehat{\boldsymbol{\theta}}^*(\tau)\big)-\widetilde{\boldsymbol{\theta}}(\tau)-h^2\boldsymbol{b}(\tau)\big]$ *to be asymptotically negligible. We find that* $\mathbb{E}^*\big(\widehat{\boldsymbol{\theta}}^*(\tau)\big)-\widetilde{\boldsymbol{\theta}}(\tau)-h^2\boldsymbol{b}(\tau)=O_p(h^3)+O_p\Big(\tilde{h}^2+\sqrt{\ln(NT)/\big((NT)^{\phi(p_0/2)}\tilde{h}\big)}\Big)$, *uniformly in* $\tau\in[0,1]$*. Assumption B3 is therefore needed. Despite these technical details, the asymptotic negligibility of the term means that the bootstrap estimators can consistently estimate the second-order bias.*

Moreover, the overall variation of coefficient curves over a time period may be of interest in practice. Simultaneous confidence bands serve this purpose. However, there is currently no available asymptotic construction of simultaneous bands for our models, irrespective of the presence of missing observations. Even if a construction were available, it might still face challenges with slow convergence speeds, as recognized in the time series literature (Zhou and Wu, 2010). We elaborate on obtaining simultaneous bands using AWB-based bootstrap correction in Appendix F.

5 Simulation study

We examine the consistency of the proposed LLDV estimator, as well as the empirical coverage and length of the pointwise intervals and simultaneous bands from the AWB, in the presence of missing observations. Throughout, we employ the Epanechnikov kernel $K(x)=3/4(1-x^2)\mathbb{1}_{\{|x|\leq 1\}}$. To examine the impact of the bandwidth, we explore fixed values of $h\in\{0.09,0.12,0.15\}$. Additionally, we introduce a data-driven bandwidth selection (denoted as $\hat{h}_{\text{PLMCV}}$) in Appendix E using the panel local modified cross-validation (PLMCV). Since the results are relatively robust to the oversmoothing bandwidth $\tilde{h}$ in STEP 1 of the AWB, we adopt the suggestion in Friedrich and Lin (2022) and set $\tilde{h}=Ch^{5/9}$ with $C=2$. Moreover, there is currently no available method for selecting an "optimal" γ in STEP 2 of the AWB. As our theory suggests that the value of γ

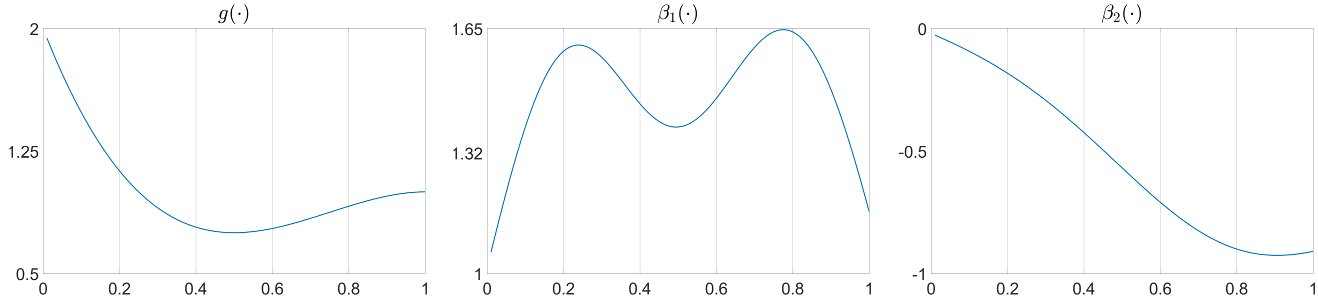


Figure 1: Plots of the global trend and the slope coefficients used in the simulation study.

should be relatively small when T is not very large, we consider $\gamma \in \{0.15, 0.2, 0.25, 0.3, 0.35, 0.4\}$ in simulations. As shown later, we find that a value around 0.25 consistently produces satisfactory results. All confidence intervals have a nominal level of 95%. The results are based on 10^3 Monte Carlo replications and $B = 999$ bootstrap samples. To evaluate the accuracy, we report the average mean squared errors (AMSE). Moreover, we employ the methodology described in Friedrich and Lin (2022, p.12) to evaluate empirical coverage and length of confidence intervals and bands (details are available in Appendix G.2). Additional discussions and the full set of results are provided in the supplementary materials, Appendices E - G.

5.1 The data generating process

We consider a model with a smooth global trend that mimics an asymmetric V-shape as observed in one of our applications in ethane (Figure 5). Additionally, we adopt two time-varying slope coefficients from Friedrich and Lin (2022); one positive slope coefficient with a two-peak shape, and one negative, smoothly decreasing coefficient. Since the model allows for heteroskedasticity, time-varying volatilities are also considered in our simulations. More specifically, we consider $y_{it} = \alpha_i + g\left(t/T\right) + \beta_1(t/T)x_{it,1} + \beta_2(t/T)x_{it,2} + \sigma_i(t/T)\varepsilon_{it}$, where $g(\tau) = -4\tau^3 + 9\tau^2 - 6\tau + 2$, and

$$\beta_1(\tau) = 1.5\exp(-10(\tau - 0.2)^2) + 1.6\exp(-8(\tau - 0.8)^2), \tag{5.1}$$

$$\beta_2(\tau) = -0.5\tau - 0.5\exp(-5(\tau - 0.8)^2). \tag{5.2}$$

Figure 1 depicts the curves of the global trend and the slope coefficients. For the regressors $\{\boldsymbol{x}_{it} = (x_{it,1}, x_{it,2})'\}$, we shall specify their local trends, individual specific effects, and the innovation processes. First, we take the following local trends $\boldsymbol{\ell}(\tau) = \big(\sin(\tau), (\tau-0.5)^2\big)'$. Second, the individual-specific effects $\boldsymbol{\chi}_i = (\chi_{i,1}, \chi_{i,2})'$ are i.i.d. U$[-1,1]$. Third, we generate the innovation processes from a VAR(1) process, where we allow for cross-sectional dependence among individuals i through the parameters $\rho_{u,1}$ and $\rho_{u,2}$ (given below) for the regressors, respectively. Additionally, we allow for dependence between regressors for each i through $\rho_{u,3}$. In summary, we generate the innovations of the regressors $\{\boldsymbol{\nu}_{\bullet t,j}\}$ as follows: $\boldsymbol{\nu}_{\bullet t,j} = (\nu_{1t,j}, \ldots, \nu_{Nt,j})' = \boldsymbol{A}_j \boldsymbol{\nu}_{\bullet t-1,j} + \boldsymbol{u}_{\bullet t,j}$, $\boldsymbol{A}_j = a_j \boldsymbol{I}_N$, $j = 1, 2$, where $\boldsymbol{u}_{\bullet t} = \big(\boldsymbol{u}'_{\bullet t,1}, \boldsymbol{u}'_{\bullet t,2}\big)' \overset{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_u)$ with $\boldsymbol{\Sigma}_u = \begin{pmatrix} \boldsymbol{\Sigma}_{u,1} & \boldsymbol{\Sigma}_{u,3} \\ \boldsymbol{\Sigma}_{u,3} & \boldsymbol{\Sigma}_{u,2} \end{pmatrix}$, $\boldsymbol{\Sigma}_{u,j} = \Big(\rho_{u,j}^{|l-i|},\ 1 \le i, l \le N\Big)$ for $j = 1, 2$, $\boldsymbol{\Sigma}_{u,3} = \rho_{u,3} \boldsymbol{I}_N$, and $a_j = 0.1$. We take $(\rho_{u,1}, \rho_{u,2}, \rho_{u,3}) = (0.3, 0.1, 0.3)$. Moreover, following Li et al. (2011), we take the fixed effects as $\alpha_i = \rho_\alpha \left(T^{-1} \sum_{t=1}^T x_{it,1}\right) + u_i^\alpha$, $u_i^\alpha \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, $i = 1, \ldots, N-1$, and $\alpha_N = -\sum_{i=1}^{N-1} \alpha_i$. We fix $\rho_\alpha = 1$. We consider heteroskedastic errors. Five patterns are adopted. In particular, the volatility process $\sigma_i(\tau)$ for individual i can be: (i) constant, i.e., $\sigma_i(\tau) = 1$; (ii) smoothly increasing, decreasing, i.e., $\sigma_i(\tau) = 1 + \kappa\tau$, $\kappa = 0.5, -0.5$; (iii) smoothly fluctuating, i.e., $\sigma_i(\tau) = 1 + a\cos(2\pi\kappa\tau)$, $a = 0.5$ and $\kappa = 4$; or (iv) smoothly increasing and fluctuating, i.e., $\sigma_i(\tau) = 1 + \tau + a\cos(2\pi\kappa\tau)$, $a = 0.5$ and $\kappa = 4$. We randomly assign one of the patterns to the individuals $i = 1, \ldots, N$. For the error process, we consider a VAR(1) process, where we produce cross-sectional dependence through the covariance matrix of the innovations. In particular, we generate the error term $\{\boldsymbol{\varepsilon}_{\bullet t} = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'\}$ according to $\boldsymbol{\varepsilon}_{\bullet t} = \boldsymbol{\Xi}\boldsymbol{\varepsilon}_{\bullet t-1} + \boldsymbol{\eta}_t$, $\boldsymbol{\eta}_t \overset{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_\eta)$, where $\boldsymbol{\Xi} = \rho_\varepsilon \boldsymbol{I}_N$, and $\boldsymbol{\Sigma}_\eta = \Big(\rho_\eta^{|j-i|},\ 1 \le i, j \le N\Big)$ with $\rho_\eta = 0.1$. We vary the strength of serial dependence and consider $\rho_\varepsilon \in \{0.1, 0.3\}$. The final step is to generate the missing patterns. We simulate a strictly stationary Markov chain of missing values independently for each unit, with the transition matrix given by

$$
\begin{array}{cc}
 & \begin{array}{cc} M_{it}=0 & M_{it}=1 \end{array} \\
\begin{array}{c} M_{i(t-1)}=0 \\ M_{i(t-1)}=1 \end{array} & \begin{pmatrix} 0.3 & 0.7 \\ 0.1 & 0.9 \end{pmatrix}
\end{array}
$$

where we let $M_{i1} \sim \text{Bern}(7/8)$. This will yield approximately 12.5% of the sample to be missing.

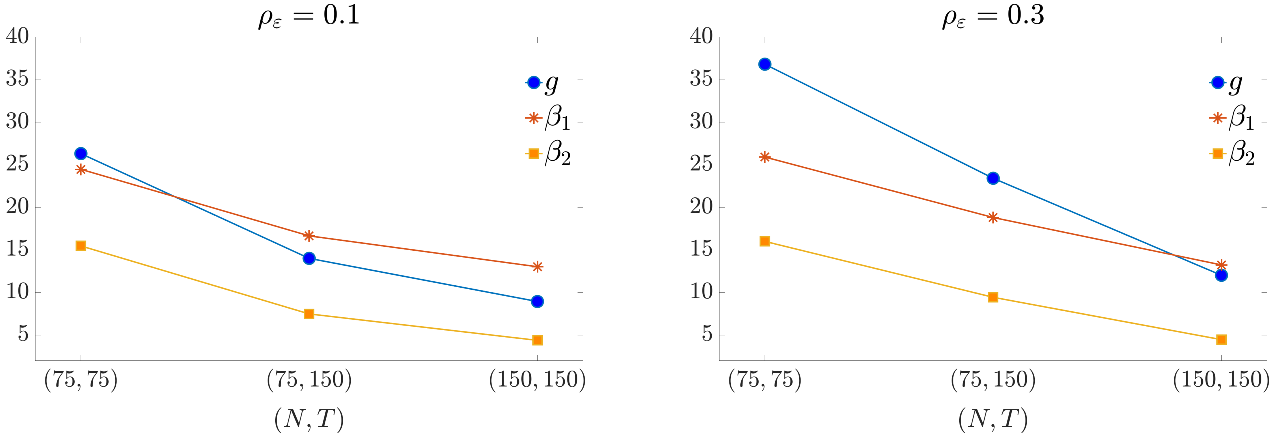


Figure 2: Average AMSE $\times 10^4$ under heteroskedasticity, and $h = 0.15$.

Table 1: Pointwise empirical coverage of confidence bands for various sample sizes and heteroskedastic errors, $\gamma = 0.2$.

		$\rho_\varepsilon = 0.1$			$\rho_\varepsilon = 0.3$		
	h	$N = 75$, $T = 75$	$N = 75$, $T = 150$	$N = 150$, $T = 150$	$N = 75$, $T = 75$	$N = 75$, $T = 150$	$N = 150$, $T = 150$
g	0.09	0.945	0.957	0.968	0.913	0.928	0.943
	0.12	0.954	0.960	0.971	0.929	0.932	0.950
	0.15	0.959	0.964	0.970	0.929	0.942	0.948
	$\hat{h}_{\text{PLMCV}}$	0.958	0.960	0.969	0.928	0.936	0.944
β_1	0.09	0.953	0.965	0.979	0.957	0.965	0.977
	0.12	0.945	0.948	0.957	0.954	0.947	0.956
	0.15	0.925	0.901	0.888	0.924	0.897	0.885
	$\hat{h}_{\text{PLMCV}}$	0.926	0.947	0.968	0.927	0.940	0.970
β_2	0.09	0.939	0.953	0.950	0.942	0.954	0.949
	0.12	0.943	0.952	0.956	0.945	0.949	0.954
	0.15	0.947	0.947	0.954	0.943	0.951	0.954
	$\hat{h}_{\text{PLMCV}}$	0.945	0.950	0.955	0.944	0.946	0.954

5.2 Simulation findings

We draw the following four main conclusions from our simulation results.

(i) Figure 2 shows the empirical accuracy increases when N or T increases. These results provide evidence of the uniform consistency of our LLDV estimator under missing values (Proposition 1).

(ii) Pointwise intervals show overall accurate empirical coverage in Table 1, even when N and T are small. They show robustness to the degree of serial correlation. For $\rho_\varepsilon = 0.3$, empirical coverage is slightly lower than for $\rho_\varepsilon = 0.1$ when $N = T = 75$. But they are nevertheless

Table 2: Pointwise empirical coverage (Cov.) and length (Lgth.) for $(\rho_\varepsilon, N, T) = (0.3, 75, 75)$ with the bandwidth $\hat{h}_{\text{PLMCV}}$ selected by PLMCV (Appendix E) and heteroskedastic errors.

	γ	0.15	0.2	0.25	0.3	0.35	0.4
	g	0.926	0.928	0.933	0.931	0.938	0.937
Cov.	β_1	0.924	0.927	0.931	0.933	0.932	0.936
	β_2	0.945	0.944	0.943	0.942	0.943	0.942
	g	0.221	0.226	0.230	0.233	0.243	0.250
Lgth.	β_1	0.190	0.193	0.195	0.199	0.206	0.211
	β_2	0.147	0.147	0.146	0.145	0.148	0.148

close to the nominal level (95%). Similar conclusions are found by Friedrich and Lin (2022) for time-varying time series models.

(iii) The pointwise results are relatively robust to the choice of bandwidth. The empirical coverage of all coefficient curves is close to the nominal level of 95%, irrespective of fixed or data-driven bandwidth ($\hat{h}_{\text{PLMCV}}$), except for $h = 0.15$ in the case of β_1. This is expected, considering that β_1 has two peaks, making it more challenging to cover compared to g and β_2. Empirical simultaneous coverage, on the other hand, is more sensitive to the choice of bandwidth but generally exhibits accurate coverage or mild undercoverage (Appendix F.1).

(iv) The pointwise results demonstrate overall robustness to the choice of γ, as shown in Table 2. A slight tradeoff between empirical coverage and length is observable.

Table G.2 (Appendix G.3) shows the empirical length decreases as either N or T increases.

6 Empirical application

We examine two empirical examples. In Section 6.1, we investigate the relationship between surface particulate matter air pollution ($PM_{2.5}$) and mortality. Subsequently, in Section 6.2, we perform a common trend analysis of atmospheric ethane emissions in the Northern Hemisphere. Both examples only involve missing observations in the dependent variables and thus fit into our framework. The bandwidths are determined through PLMCV (Appendix E). For both analyses, we present results for $\gamma = 0.2$ and $B = 1,500$. Additional outputs for different values of γ can be

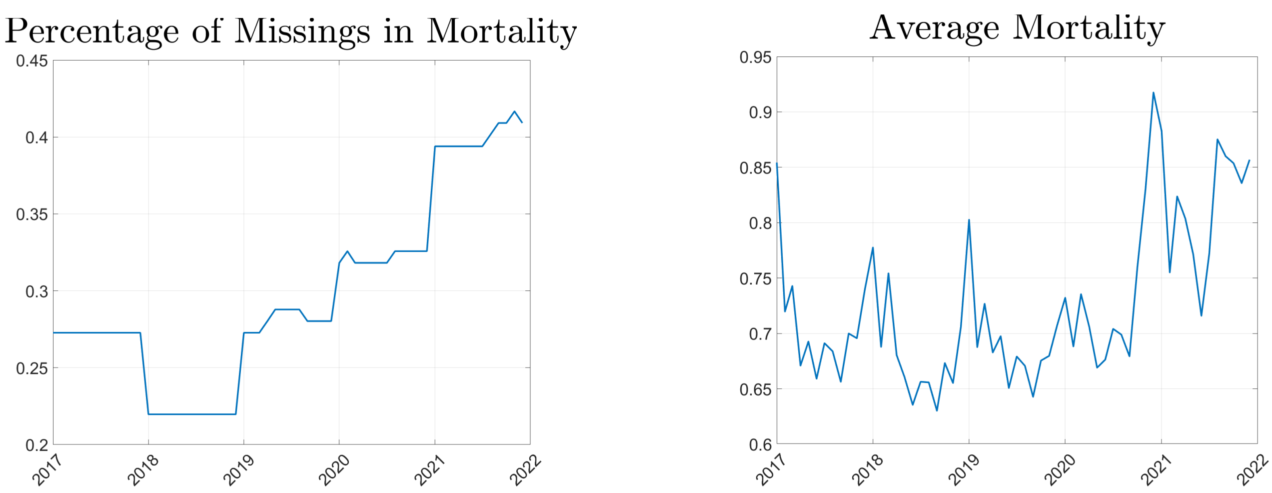


Figure 3: Proportion of missing mortality data (left) and cross-sectional average for mortality (right).

found in Appendix H.1, where the results remain similar. Further details on data acquisition can be found in Appendix H.3.

6.1 Mortality and $PM_{2.5}$

The effect of $PM_{2.5}$ on mortality has been studied by, for example, Pope III et al. (2009), Samet et al. (2000), Lelieveld et al. (2015), and Zhang et al. (2017). Existing research identifies a range of causes of death attributable to $PM_{2.5}$ exposure (Burnett et al., 2018; Landrigan et al., 2018). Current studies characterize the relationship between $PM_{2.5}$ concentrations and mortality as time-invariant. In our empirical analysis, we revisit this relationship, examining the trending patterns and a potentially time-varying effect. We collect population-weighted surface $PM_{2.5}$ data in microgram per cubic metre (μg/m^3) from the Atmospheric Composition Analysis Group. For the mortality, we take the total death cases per month from UNdata and divide it by the total population per month. The monthly total population data is part of the Atmospheric Composition Analysis Group dataset. Our sample, in total, includes 132 countries all over the globe and spans the months from January 2017 to December 2021, resulting in $T = 60$.

Due to registration and accounting limitations, mortality data are not transferred to the UN for each month for every country. This results in an average number of missing observations per month of around 30%. In Figure 3 (left), we display the average of missing data across all

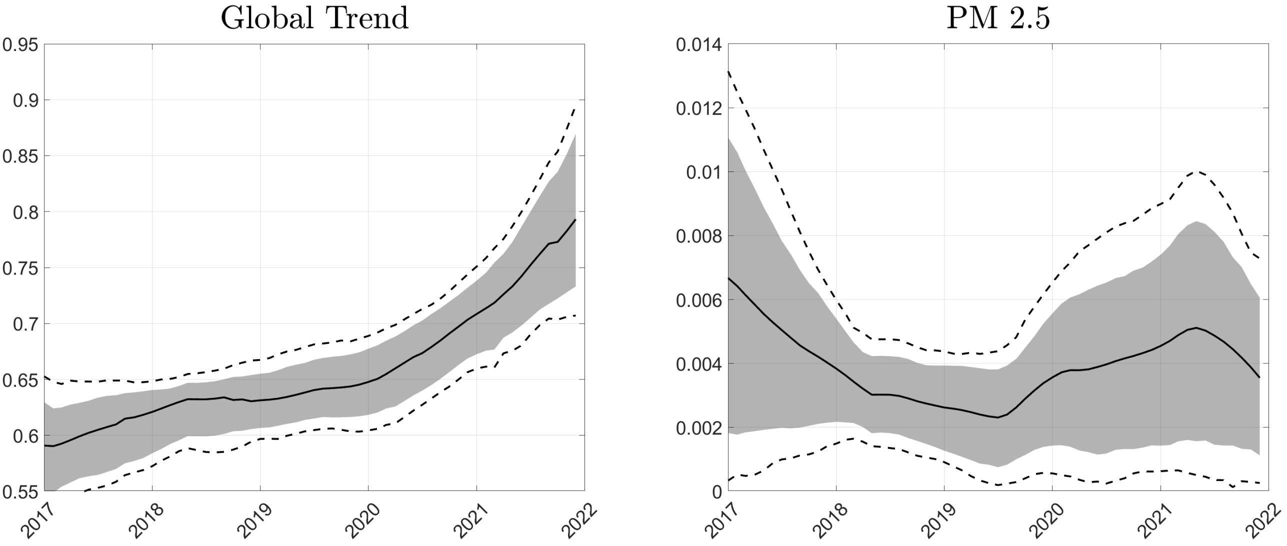


Figure 4: Estimated coefficient curves with 95%-level pointwise intervals (shaded area) and simultaneous bands (dashed lines, see Appendix F). The simultaneity is over the full sample period.

countries over time. The data availability varies throughout the sample period and appears to follow a Markovian pattern on average, aligning with our assumptions on the missing pattern.

Since we are interested in the effect of $PM_{2.5}$ on mortality, we denote y_{it} as the average population-weighted mortality for country i in month t, and x_{it} as the corresponding level of $PM_{2.5}$. In Figure 3 (right), we observe a gradual upward trend in y_{it}, while Figure H.1 (Appendix H.1) illustrates a slight downward trend in x_{it}. Both of these trending patterns align with our theoretical framework. We thus estimate our model including a global trend and x_{it} with a data-driven bandwidth $\hat{h}_{\mathrm{PLMCV}} = 0.2638$.

Figure 4 displays the estimated global trend in the left panel as well as the estimated coefficient curve for $PM_{2.5}$ in the right panel (black lines). In both panels, the pointwise confidence intervals (shaded area) and the simultaneous confidence bands (dashed lines) are added. The pointwise intervals show us the significance of the trend or the effect of $PM_{2.5}$ on mortality at any given point in time, while we consult the simultaneous confidence bands for inference on the development over time. The global trend in mortality (left panel) shows an overall upward-sloping behavior over the whole sample. According to the 95% simultaneous confidence bands, this trending behavior is significant. In addition, we find a significant and positive effect of $PM_{2.5}$ concentrations on mortality over the whole sample (right panel). The estimated coefficient curve of $PM_{2.5}$ shows a relatively flat downward-upward-downward pattern. As highlighted by previous studies (Li et al., 2023), a positive effect provides strong motivation for effective policy on air quality to reduce mortality. The pointwise confidence intervals for the $PM_{2.5}$ indicate that the effect is significantly positive at each time point in our sample. Additionally, the width of the confidence intervals seems to increase as the estimated effect of $PM_{2.5}$ increases, for example, the width around the peak in 2021 is larger than the width around the trough in 2018. Exploring the reasons behind the time-varying width may be an interesting avenue for further research.

6.2 Ethane emissions

Ethane is the most abundant non-methane hydrocarbon gas and its main sources are the oil and gas industry, where it is co-emitted with methane. It contributes to the formation of ground-level ozone which is a major pollutant. Friedrich et al. (2020) analyze trends in ethane emissions from four different measurement stations on a series-by-series basis. We extend their analysis by extracting a common trend among $N = 11$ series from the Northern Hemisphere. The data are obtained from the Network for the Detection of Atmospheric Composition Change (NDACC) using the Fourier Transform InfraRed (FTIR) remote-sensing technique, spanning from 1986 to 2022 with daily observations ($T = 13,394$). The dry air mole fraction of ethane is given in parts per billion (ppb). We deseasonalize the data on a station-by-station basis by subtracting their station-specific mean and regressing the series on one Fourier term. Since measurements can only be taken under clear-sky conditions, the average across stations of missing observations in our sample is around 90%. We apply our LLDV to estimate model (2.1) with $d = 0$, as the model in Robinson (2012), and obtain only the global trend estimate $\widehat{g}$ with the bootstrapped confidence intervals/bands. The data-driven bandwidth is $\hat{h}_{\mathrm{PLMCV}} = 0.1298$; see details in Appendix H.2.

The global trend and corresponding confidence intervals/bands are shown in Figure 5. The left panel plots the data in colored circles as well as the estimated global trend (black line) and the 95% simultaneous confidence intervals (dashed lines). The right panel zooms in on the trend

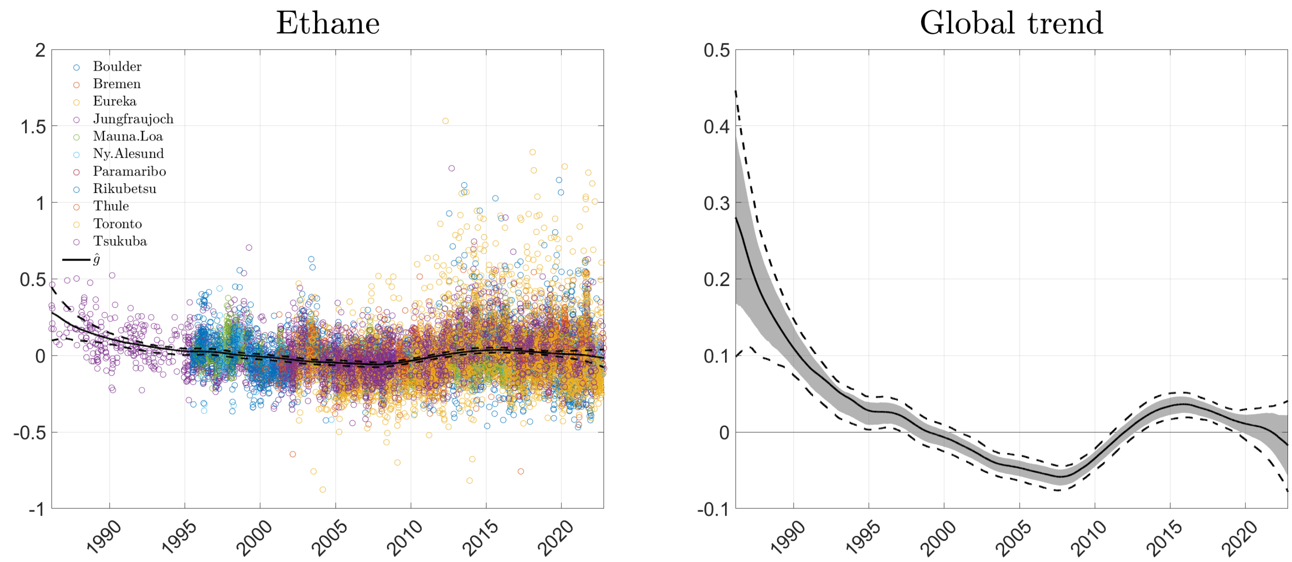


Figure 5: Estimated global trend for ethane emissions; refer to Figure 4 for additional details.

and confidence bands. We observe an overall downward-upward-downward trending pattern which confirms the main findings of the univariate analysis in Friedrich et al. (2020), but here it is common to many stations. The first trend reversal around 2007 is often attributed to the oil and gas boom in North America while the second reversal could be explained by a drastic drop in oil prices making it less profitable to exploit shale gas wells. The local peak in 1997/1998 can be attributed to boreal forest fires happening in Russia (Friedrich et al., 2020).

7 Conclusion

We explored a class of trending time-varying panel models that allow for missing observations, cross-sectional and serial dependence, and heteroskedasticity. We introduced a local linear dummy variable estimation method to handle missing observations in the dependent variable without relying on imputation. We obtained the limiting distribution of the parameter estimators based on a new uniform convergence result. This result accommodates cross-sectional dependence and near-epoch dependence over time. The limiting distribution contains various nuisance parameters. Estimating these nuisance parameters is challenging; hence, we propose an autoregressive wild bootstrap (AWB) procedure to construct confidence intervals and bands. The AWB automatically captures these nuisance parameters and demonstrates good performance in finite samples. We illustrated the proposed methods through two applications. First, we investigated the relationship between environmental quality and mortality in 132 countries, uncovering a significant upward trend. Second, we examined the global trend in atmospheric ethane, revealing an overall downward-upward-downward trending pattern. For future research, developing results of strong approximations will be crucial to justify the simultaneous bands. Additionally, the impact of missing observations in covariates on the asymptotic distribution remains an open question.

8 Acknowledgments

The paper has benefited from many inspiring discussions with Eric Beutner, Andre Lucas, Simon Reese, and Stephan Smeekes. We especially thank Ignace De Vos and Hanno Reuvers for their careful reading and valuable suggestions during the drafting of the initial versions of this paper. We thank Mingxuan Song for his excellent research assistantship during the initiation of the project. The paper has been presented at department seminars at Lund University, Vrije Universiteit Amsterdam, and conferences NESG 2023, IPDC 2023, We thank all the participants. Special gratitude is expressed to Aaron van Donkelaar for providing and preprocessing the $PM_{2.5}$ data, and to Emmanuel Mahieu for supporting the acquisition of ethane emissions data. We gratefully acknowledge J. Notholt (University of Bremen, Bremen, Germany), J.W. Hannigan (National Center for Atmospheric Research, NCAR, Boulder, CO, USA), E. Mahieu (University of Liège, Belgium), K. Strong (University of Toronto, Toronto, ON, Canada), T. Nagahama (ISEE Nagoya University), I. Morino (National Institute for Environmental Studies, Tsukuba, Japan) for providing FTIR data for the Bremen, Bolder, Eureka, Jungfraujoch, Ny Alesund, Mauna Loa, Paramaribo, Rikubetsu, Thule, Toronto and Tsukuba stations. Network for the Detection of Atmospheric Composition Change data are publicly available at http://www.ndacc.org.

References

Atak, A., T. T. Yang, Y. Zhang, and Q. Zhou (2023). Specification tests for time-varying coefficient panel data models. *Econometric Theory*, 1–48.

Awaworyi Churchill, S., K. Ivanovski, and M. E. Munyanyi (2021). Income inequality and renewable energy consumption: Time-varying non-parametric evidence. *Journal of Cleaner Production 296*, 126306.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica 77*(4), 1229–1279.

Bard, A., B. Renard, M. Lang, I. Giuntoli, J. Korck, G. Koboltschnig, M. Janža, M. d'Amico, and D. Volken (2015). Trends in the hydrologic regime of Alpine rivers. *Journal of Hydrology 529*, 1823–1837.

Bashir, F. and H.-L. Wei (2018). Handling missing data in multivariate time series using a vector autoregressive model-imputation (VAR-IM) algorithm. *Neurocomputing 276*, 23–30.

Bradley, R. C. (2005). Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions. *Probability Surveys 2*, 107 – 144.

Burnett, R., H. Chen, M. Szyszkowicz, N. Fann, B. Hubbell, C. A. Pope III, et al. (2018). Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proceedings of the National Academy of Sciences 115*(38), 9592–9597.

Cahan, E., J. Bai, and S. Ng (2023). Factor-based imputation of missing values and covariances in panel data of large dimensions. *Journal of Econometrics 233*(1), 113–131.

Cai, B., T. Cheng, and C. Yan (2018). Time-varying skills (versus luck) in U.S. active mutual funds and hedge funds. *Journal of Empirical Finance 49*, 81–106.

Cavaliere, G., A. Rahbek, and A. R. Taylor (2010). Testing for co-integration in vector autoregressions with non-stationary volatility. *Journal of Econometrics 158*(1), 7–24.

Chen, B. and L. Huang (2018). Nonparametric testing for smooth structural changes in panel data models. *Journal of Econometrics 202*(2), 245–267.

Chen, J., J. Gao, and D. Li (2012). Semiparametric trending panel data models with cross-sectional dependence. *Journal of Econometrics 171*(1), 71–85.

Corradi, V. and N. R. Swanson (2014). Testing for structural stability of factor augmented forecasting models. *Journal of Econometrics 182*(1), 100–118. Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions.

Davidson, J. (2002). Establishing conditions for the functional central limit theorem in nonlinear and semiparametric time series processes. *Journal of Econometrics 106*(2), 243–269.

de Jong, R. M. and J. Davidson (2000). Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices. *Econometrica 68*(2), 407–423.

de Jong, R. M. and T. Woutersen (2011). Dynamic time series binary choice. *Econometric Theory 27*(4), 673–702.

Dong, C., J. Gao, and B. Peng (2015). Semiparametric single-index panel data models with cross-sectional dependence. *Journal of Econometrics 188*(1), 301–312.

Friedrich, M., E. Beutner, H. Reuvers, S. Smeekes, J.-P. Urbain, W. Bader, et al. (2020). A statistical analysis of time trends in atmospheric ethane. *Climatic Change 162*, 105–125.

Friedrich, M. and Y. Lin (2022). Sieve bootstrap inference for linear time-varying coefficient models. *Journal of Econometrics (forthcoming)*.

Friedrich, M., Y. Lin, P. Ramdara, S. Telg, and B. van der Sluis (2023). Time-varying effects of housing attributes and economic environment on housing prices. Tinbergen Institute Discussion Paper, 2023-039/III.

Friedrich, M., S. Smeekes, and J.-P. Urbain (2020). Autoregressive wild bootstrap inference for nonparametric trends. *Journal of Econometrics 214*(1), 81–109. Annals Issue: Econometric Models of Climate Change.

Gao, J., B. Peng, and Y. Yan (2023). Higher-order expansions and inference for panel data models. *Journal of the American Statistical Association (forthcoming)*.

Gao, J., K. Xia, and H. Zhu (2020). Heterogeneous panel data models with cross-sectional dependence. *Journal of Econometrics 219*(2), 329–353. Annals Issue: Econometric Estimation and Testing: Essays in Honour of Maxwell King.

Gonçalves, S. and B. Perron (2014). Bootstrapping factor-augmented regression models. *Journal of Econometrics 182*(1), 156–173. Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions.

Gonçalves, S. and B. Perron (2020). Bootstrapping factor models with cross sectional dependence. *Journal of Econometrics 218*(2), 476–495.

Hailemariam, A., R. Smyth, and X. Zhang (2019). Oil prices and economic policy uncertainty: Evidence from a nonparametric panel data model. *Energy Economics 83*, 40–51.

Hansen, B. E. (2008). Uniform convergence rates for kernel estimation with dependent data. *Econometric Theory 24*, 726–748.

Jiang, W. (2009). On uniform deviations of general empirical risks with unboundedness, dependence, and high dimensionality. *Journal of Machine Learning Research 10*(36), 977–996.

Keef, C., J. Tawn, and C. Svensson (2009). Spatial risk assessment for extreme river flows. *Journal of the Royal Statistical Society Series C: Applied Statistics 58*(5), 601–618.

Kristensen, D. (2009). Uniform convergence rates of kernel estimators with heterogeneous dependent data. *Econometric Theory 25*, 1433–1445.

Landrigan, P. J., R. Fuller, N. J. Acosta, O. Adeyi, R. Arnold, A. B. Baldé, et al. (2018). The Lancet commission on pollution and health. *The lancet 391*(10119), 462–512.

Lelieveld, J., J. S. Evans, M. Fnais, D. Giannadaki, and A. Pozzer (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature 525*(7569), 367–371.

Li, D., J. Chen, and J. Gao (2011). Non-parametric time-varying coefficient panel data models with fixed effects. *The Econometrics Journal 14*(3), 387–408.

Li, D., Z. Lu, and O. Linton (2012). Local linear fitting under near epoch dependence: Uniform consistency with convergence rates. *Econometric Theory 28*(5), 935–958.

Li, L., J. Zhang, Y. Wang, and B. Ran (2018). Missing value imputation for traffic-related time series data based on a multi-view learning method. *IEEE Transactions on Intelligent Transportation Systems 20*(8), 2933–2943.

Li, W., A. Tian, Y. Shi, B. Chen, R. Ji, J. Ge, et al. (2023). Associations of long-term fine particulate matter exposure with all-cause and cause-specific mortality: results from the ChinaHEART project. *The Lancet Regional Health–Western Pacific 41*.

Liddle, B., R. Smyth, and X. Zhang (2020). Time-varying income and price elasticities for energy demand: Evidence from a middle-income panel. *Energy Economics 86*, 104681.

Lin, Y. and H. Reuvers (2022). Cointegrating polynomial regressions with power law trends: Environmental Kuznets curve or omitted time effects? Tinbergen Institute Discussion Paper, 2022-092/III.

Lu, Z. and O. Linton (2007). Local linear fitting under near epoch dependence. *Econometric Theory 23*(1), 37–70.

Pope III, C. A., M. Ezzati, and D. W. Dockery (2009). Fine-particulate air pollution and life expectancy in the United States. *New England Journal of Medicine 360*(4), 376–386.

Ren, X., Z. Tong, X. Sun, and C. Yan (2022). Dynamic impacts of energy consumption on economic growth in China: Evidence from a non-parametric panel data model. *Energy Economics 107*, 105855.

Robinson, P. M. (2012). Nonparametric trending regression with cross-sectional dependence. *Journal of Econometrics 169*(1), 4–14. Recent Advances in Panel Data, Nonlinear and Nonparametric Models: A Festschrift in Honor of Peter C.B. Phillips.

Samet, J. M., F. Dominici, F. C. Curriero, I. Coursac, and S. L. Zeger (2000). Fine particulate air pollution and mortality in 20 US cities, 1987–1994. *New England Journal of Medicine 343*(24), 1742–1749.

Shao, Q.-M. and H. Yu (1996). Weak convergence for weighted empirical processes of dependent sequences. *The Annals of Probability 24*(4), 2098 – 2127.

Silvapulle, P., R. Smyth, X. Zhang, and J.-P. Fenech (2017). Nonparametric panel data model for crude oil and stock market prices in net oil importing countries. *Energy Economics 67*, 255–267.

Smeekes, S. and A. R. Taylor (2012). Bootstrap union tests for unit roots in the presence of nonstationary volatility. *Econometric Theory 28*(2), 422–456.

Smeekes, S. and J.-P. Urbain (2014). A multivariate invariance principle for modified wild bootstrap methods with an application to unit root testing. *Maastricht University Research Papers, RM/14/008*.

Sun, X., S. Xiao, X. Ren, and B. Xu (2023). Time-varying impact of information and communication technology on carbon emissions. *Energy Economics 118*, 106492.

Sun, Y., R. J. Carroll, and D. Li (2009). Semiparametric estimation of fixed-effects panel data varying coefficient models. In *Nonparametric econometric methods*. Emerald Group Publishing Limited.

Trapani, L. (2013). On bootstrapping panel factor series. *Journal of Econometrics 172*(1), 127–141.

Uddin, M. M., V. Mishra, and R. Smyth (2020). Income inequality and CO2 emissions in the G7, 1870-2014: Evidence from non-parametric modelling. *Energy Economics 88*, 104780.

Yao, L. and W. Jiang (2012). On extensions of Hoeffding's inequality for panel data. *Statistics & Probability Letters 82*(3), 446–454.

Zhang, Q., X. Jiang, D. Tong, S. J. Davis, H. Zhao, G. Geng, et al. (2017). Transboundary health impacts of transported global air pollution and international trade. *Nature 543*(7647), 705–709.

Zhou, Z. and W. B. Wu (2010). Simultaneous inference of linear models with time varying coefficients. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 72*, 513–531.

Supplemental Appendix to:

Bootstrapping trending time-varying coefficient panel models with missing observations

Yicong Lin[1,2], *Bernhard van der Sluis*[3] *and Marina Friedrich*[1,2]

[1] *Vrije Universiteit Amsterdam*

[2] *Tinbergen Institute*

[3] *Erasmus University Rotterdam*

Contents

A Notation and definition of near-epoch processes

We collect the notation used throughout this paper in this section. Let $\mathbb{1}\{\cdot\}$ be an indicator function. For any vector $\boldsymbol{x} = (x_j) \in \mathbb{R}^n$, its p-norm is denoted by $\|\boldsymbol{x}\|_p = (\sum_{j=1}^n |x_j|^p)^{1/p}$. The induced p-norm for a matrix $\boldsymbol{A}$ is $\|\boldsymbol{A}\|_p = \sup_{\boldsymbol{x}\neq\boldsymbol{0}} \|\boldsymbol{A}\boldsymbol{x}\|_p/\|\boldsymbol{x}\|_p$. We omit the subscripts whenever $p = 2$. Let $\operatorname{diag}(\boldsymbol{A}, \boldsymbol{B}) = \begin{pmatrix} \boldsymbol{A} & \\ & \boldsymbol{B} \end{pmatrix}$ for matrices A and B. For any vector $\boldsymbol{x} \in \mathbb{R}^n$ and diagonal matrix $\boldsymbol{D} = \operatorname{diag}(d_1, \ldots, d_n) \in \mathbb{R}^{n\times n}$, $\boldsymbol{x}^k = \left(x_j^k\right)$ and $\boldsymbol{D}^k = \operatorname{diag}\left(d_1^k, \ldots, d_n^k\right)$ take the power k element-wise. The Kronecker product is denoted by "$\otimes$". The symbols "$\xrightarrow{p}$" and "$\xrightarrow{d}$" denote convergence in probability and in distribution, respectively. For two sequences of positive numbers $\{a_{NT}\}$ and $\{b_{NT}\}$, we write $a_{N,T} \sim b_{N,T}$ if there exists some constant $\varepsilon > 0$ such that $\varepsilon^{-1} \le a_{NT}/b_{NT} \le \varepsilon$ for all large N and T. For $a, b \in \mathbb{R}$, let $a \wedge b = \min\{a, b\}$ and $a \vee b = \max\{a, b\}$. Moreover, we denote $\sum_{i_1,i_2,\ldots,i_k=m}^n = \sum_{i_1=m}^n \sum_{i_2=m}^n \cdots \sum_{i_k=m}^n$, where $k, m, n \in \mathbb{Z}^+$ and $m \le n$. Bootstrap quantities are given a superscript *, expressing that it is conditional on the original sample as in, e.g., Boswijk et al. (2021). For instance, "$\xrightarrow{d^*}_p$" denotes bootstrap weak convergence in probability (Gine and Zinn, 1990). Let $\mathcal{C}^i\mathcal{I}$, $i \in \mathbb{N}$, be the collection of functions that have i_{th}-order continuous derivatives on the interval $\mathcal{I} \subset \mathbb{R}$, and $f^{(i)}(x) = \mathrm{d}^i f(x)/\mathrm{d}x^i$ represent the i-th-order derivative of f with respect to x. The generic constants $C, C_1, C_2, \ldots$ can change from line to line.

The following definition of near-epoch process is an extension from time series to panel models.

Definition A.1 (Near-Epoch Dependence) *Let $\{\boldsymbol{X}_{it},\, t \in \mathbb{Z}\}$ be an $\mathbb{R}^d$-valued process, defined based on a possibly vector-valued process $\{\boldsymbol{\varepsilon}_{it},\, t \in \mathbb{Z}\}$ by*

$$\boldsymbol{X}_{it} = (X_{it,1}, \ldots, X_{it,d})' = \Psi_{\boldsymbol{X}_i}\left(\boldsymbol{\varepsilon}_{it}, \boldsymbol{\varepsilon}_{i(t-1)}, \ldots\right), \qquad i = 1, \ldots, N,$$

where $\Psi_{\boldsymbol{X}_i} : \mathbb{R}^\infty \to \mathbb{R}^d$ are Borel measurable functions. For $i = 1, \ldots, N$, the process $\{\boldsymbol{X}_{it},\, t \in \mathbb{Z}\}$ is said to be (strictly stationary) near-epoch dependent in L_p-norm (NED in L_p) with respect to $\{\boldsymbol{\varepsilon}_{it},\, t \in \mathbb{Z}\}$, if $\{\boldsymbol{\varepsilon}_{it},\, t \in \mathbb{Z}\}$ is (strictly stationary) α-mixing process, and

$$\psi_{i,p}(m) = \sup_{t\in\mathbb{Z}} \mathbb{E}\left\|\boldsymbol{X}_{it} - \boldsymbol{X}_{it}^{(m)}\right\|^p \to 0, \qquad p > 0, \qquad \text{as} \quad m \to \infty, \tag{A.1}$$

where $\boldsymbol{X}_{it}^{(m)} = \left(X_{it,1}^{(m)}, \ldots, X_{it,d}^{(m)}\right)' = \Psi_{\boldsymbol{X}_i,m}(\boldsymbol{\varepsilon}_{it}, \ldots, \boldsymbol{\varepsilon}_{i(t-m+1)})$, and $\Psi_{\boldsymbol{X}_i,m}$ are $\mathbb{R}^d$-valued Borel measurable functions with m arguments. The term $\psi_{i,p}(m)$ is said to be the stability coefficients of order p of the process $\{\boldsymbol{X}_{it},\, t \in \mathbb{Z}\}$.

The process $\{\boldsymbol{X}_{it},\, t \in \mathbb{Z}\}$ is not required to be strictly stationary. If it is strictly stationary, then $\sup_{t\in\mathbb{Z}} \mathbb{E}\|\boldsymbol{X}_{it} - \boldsymbol{X}_{it}^{(m)}\|^p = \mathbb{E}\|\boldsymbol{X}_{it} - \boldsymbol{X}_{it}^{(m)}\|^p$ in Eq. (A.1).

B Proofs of auxiliary results

This section provides some intermediate results that are useful in the proofs of the main results. For convenience, define $w_t^k(\tau) = \left(\frac{\tau_t-\tau}{h}\right)^k K\left(\frac{\tau_t-\tau}{h}\right)$, $k \geq 0$. $\forall \tau \in [0,1]$, let $\mathcal{I}_{h,s}(\tau) = \{t : |\tau_{t+s} - \tau| \leq h\}$, $s \geq 0$. Without confusion, we suppress the dependence on s whenever $s = 0$, i.e., $\mathcal{I}_h(\tau) = \mathcal{I}_{h,0}(\tau)$. We first establish the general theorem of uniform convergence which proves to be useful in our proofs.

Proof of Theorem 1 For $m \geq 0$, define $Q_{N,t} = N^{-1}\sum_{i=1}^N Y_{it}$ and $Q_{N,t}^{(m)} = N^{-1}\sum_{i=1}^N Y_{it}^{(m)}$. Then the LHS of (3.5) can be bounded by

$$\sup_{\tau\in[0,1]}\left|\frac{1}{Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)} w_t^k(\tau)\left[Q_{N,t}^{(m)} - \mathbb{E}\left(Q_{N,t}^{(m)}\right)\right]\right| + \sup_{\tau\in[0,1]}\left|\frac{1}{Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)} w_t^k(\tau)\mathbb{E}\left(Q_{N,t} - Q_{N,t}^{(m)}\right)\right|$$
$$+ \sup_{\tau\in[0,1]}\left|\frac{1}{Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)} w_t^k(\tau)\left(Q_{N,t} - Q_{N,t}^{(m)}\right)\right| =: K_{NT,1} + K_{NT,2} + K_{NT,3}, \quad \text{(B.1)}$$

where $\mathcal{I}_{h,u_0}(\tau) = \{t : |\tau_t - \tau| \leq u_0 h\}$. Consider the term $K_{NT,1}$ first. We cover the interval $[0,1]$ by a finite number of subintervals B_l, $l = 1,\ldots,L_{NT}$. These subintervals are centered at b_l with a radius of $h\,r_{NT}\,[\delta_N(m)]^{-1}$, where $r_{NT} = \sqrt{\frac{\ln(NT)}{(NT)^{\phi(q)}h}}$ and $\delta_N(m) = N^{-1}\sum_{i=1}^N [\psi_{i,p}(m)]^{1/p} + N^{-1/2}$. Then, by the triangle inequality, $K_{NT,1}$ is bounded by

$$\max_{1\leq l\leq L_{NT}}\sup_{\tau\in B_l}\left|\frac{1}{Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)}\left(w_t^k(\tau) - w_t^k(b_l)\right)\left[Q_{N,t}^{(m)} - \mathbb{E}\left(Q_{N,t}^{(m)}\right)\right]\right|$$
$$+ \max_{1\leq l\leq L_{NT}}\left|\frac{1}{Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)} w_t^k(b_l)\left[Q_{N,t}^{(m)} - \mathbb{E}\left(Q_{N,t}^{(m)}\right)\right]\right|. \quad \text{(B.2)}$$

By Minkowski's and Lyapunov's inequalities, (3.2) and (3.3), we obtain

$$\sup_{t\geq 1}\left(\mathbb{E}\left|Q_{N,t}^{(m)} - \mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|^q\right)^{1/q} \leq \sup_{t\geq 1}\left\{2\left(\mathbb{E}\left|Q_{N,t} - Q_{N,t}^{(m)}\right|^q\right)^{1/q} + \left(\mathbb{E}\left|Q_{N,t} - \mathbb{E}(Q_{N,t})\right|^q\right)^{1/q}\right\}$$
$$= N^{-1}\sup_{t\geq 1}\left\{2\left(\mathbb{E}\left|\sum_{i=1}^N\left[Y_{it} - Y_{it}^{(m)}\right]\right|^q\right)^{1/q} + \left(\mathbb{E}\left|\sum_{i=1}^N\left[Y_{it} - \mathbb{E}\left(Y_{it}\right)\right]\right|^q\right)^{1/q}\right\} \leq C\delta_N(m). \quad \text{(B.3)}$$

Note that, for any $\tau \in B_l$, $l = 1,\ldots,L_{NT}$, $\left|w_t^k(\tau) - w_t^k(b_l)\right| \leq C r_{NT}\,[\delta_N(m)]^{-1}$ by the Lipschitz continuity of $x \mapsto x^k K(x)$, where C does not depend on (k,l,τ). By (B.3), the first component in (B.2) can then be bounded by

$$\frac{1}{Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)}\max_{1\le l\le L_{NT}}\sup_{\tau\in B_l}\left|w_t^k(\tau)-w_t^k(b_l)\right|\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|$$

$$\le Cr_{NT}\frac{1}{Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)}\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|=O_p(r_{NT}). \quad \text{(B.4)}$$

Next, we consider the second component in (B.2). Let $\mathcal{F}_{\eta,t_1}^{t_2}=\sigma\left(\boldsymbol{\eta}_t,\, t_1\le t\le t_2\right)$. For any $\epsilon>0$, by Bonferroni's inequality and the Triplex inequality (Jiang, 2009, Theorem 1), we obtain

$$\begin{aligned}
&\mathbb{P}\left(\max_{1\le l\le L_{NT}}\left|\frac{1}{Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)}w_t^k(b_l)\left[Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right]\right|>2u_0\,r_{NT}\,\epsilon\right)\\
&\le\sum_{l=1}^{L_{NT}}\mathbb{P}\left(\left|\frac{1}{Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)}w_t^k(b_l)\left[Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right]\right|>2u_0\,r_{NT}\,\epsilon\right)\\
&\le\sum_{l=1}^{L_{NT}}\left\{4m\exp\left(-\frac{2u_0Th\,r_{NT}^2}{4m^2\kappa_{NT}^2}\frac{\epsilon^2}{288}\right)+\frac{6}{\epsilon}\frac{1}{r_{NT}}\frac{1}{2u_0Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)}w_t^k(b_l)\mathbb{E}\left|\mathbb{E}\left(Q_{N,t}^{(m)}\,|\,\mathcal{F}_{\eta,-\infty}^{t-2m}\right)-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|\right.\\
&\quad\left.+\frac{15}{\epsilon}\frac{1}{r_{NT}}\frac{1}{2u_0Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)}w_t^k(b_l)\mathbb{E}\Big(\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|\mathbb{1}\left\{w_t^k(b_l)\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|>\kappa_{NT}\right\}\Big)\right\},
\end{aligned}$$

(B.5)

where $\kappa_{NT}>0$ can be any sequence that depends on (N,T). Following the proof of Theorem 14.1 in Davidson (1994), we observe that $\left\{Q_{N,t}^{(m)},\, t\in\mathbb{Z}\right\}$ is an α-mixing process with mixing coefficients $\alpha^{(m)}(j)$ bounded by

$$\alpha^{(m)}(j)\le\begin{cases}\alpha(j-m), & j\ge m+1,\\ 1/4, & j\le m.\end{cases} \quad \text{(B.6)}$$

By Theorem 14.2 of Davidson (1994) and (B.3), the "dependence term" in (B.5), i.e., the second term (without multiplicative constants), can be bounded as

$$\frac{1}{r_{NT}}\frac{1}{2u_0Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)}w_t^k(b_l)\mathbb{E}\left|\mathbb{E}\left(Q_{N,t}^{(m)}\,|\,\mathcal{F}_{\eta,-\infty}^{t-2m}\right)-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|\le Cr_{NT}^{-1}\left[\alpha(m)\right]^{1-1/q}\delta_N(m). \quad \text{(B.7)}$$

Moreover, by Hölder's and Markov's inequalities, the "tail term" in (B.5), i.e., the last term (without multiplicative constants), can be bounded as

$$
\begin{aligned}
&\frac{1}{r_{NT}}\frac{1}{2u_0Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)} w_t^k(b_l)\mathbb{E}\Big(\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|\mathbb{1}\Big\{w_t^k(b_l)\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|>\kappa_{NT}\Big\}\Big)\\
&\le \frac{1}{r_{NT}}\frac{1}{2u_0Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)} w_t^k(b_l)\left(\mathbb{E}\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|^q\right)^{1/q}\left[\mathbb{P}\left(w_t^k(b_l)\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|>\kappa_{NT}\right)\right]^{1-1/q}\\
&\le \frac{1}{r_{NT}}\frac{1}{2u_0Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)} w_t^k(b_l)\left(\mathbb{E}\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|^q\right)^{1/q}\left[\frac{\left[w_t^k(b_l)\right]^q\mathbb{E}\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|^q}{\kappa_{NT}^q}\right]^{1-1/q}\\
&= \frac{1}{r_{NT}\kappa_{NT}^{1-q}}\frac{1}{2u_0Th}\sum_{t\in\mathcal{I}_{h,u_0}(\tau)}\left[w_t^k(b_l)\right]^q\mathbb{E}\left|Q_{N,t}^{(m)}-\mathbb{E}\left(Q_{N,t}^{(m)}\right)\right|^q\\
&\le Cr_{NT}^{-1}\kappa_{NT}^{1-q}\left[\delta_N(m)\right]^q. \qquad\qquad (B.8)
\end{aligned}
$$

By Eqs. (B.2), (B.4), (B.5), (B.7), and (B.8), and given that $L_{NT}=O\left(h^{-1}r_{NT}^{-1}\delta_N(m)\right)$, to establish $K_{NT,1}=O_p(r_{NT})$, it suffices to show there exist a positive sequence $m=m_{NT}$ such that, as $(N,T)\to\infty$,

(a) $h^{-1}r_{NT}^{-1}\delta_N(m_{NT})m_{NT}\exp\left(-\frac{Thr_{NT}^2}{m_{NT}^2\kappa_{NT}^2}C\right)\to 0$ for any $C>0$;

(b) $h^{-1}r_{NT}^{-2}\left[\alpha(m_{NT})\right]^{1-1/q}\left[\delta_N(m_{NT})\right]^2\to 0$;

(c) $h^{-1}r_{NT}^{-2}\kappa_{NT}^{1-q}\left[\delta_N(m_{NT})\right]^{q+1}\to 0$.

Observe that, from Eq. (3.3), we have

$$
\delta_N(m_{NT})=m_{NT}^{-\lambda/p}N^{-1/2}\left(N^{-1/2}\sum_{i=1}^N d_i\right)+N^{-1/2}=O(N^{-1/2}). \qquad (B.9)
$$

We then take

$$
\begin{aligned}
m_{NT}&=\left[\left(\frac{\ln(NT)}{h\,r_{NT}^2}\frac{1}{N}\right)^{1/[\beta(1-1/q)]}\right]\sim\left[(NT)^{\phi_\beta(q)}N^{-1}\right]^{1/[\beta(1-1/q)]},\\
\kappa_{NT}&=\left(\frac{\ln(NT)}{h\,r_{NT}^2}\frac{1}{N^{(q+1)/2}}\right)^{1/(q-1)}\sim\left[(NT)^{\phi_\beta(q)}N^{-(q+1)/2}\right]^{1/(q-1)}.
\end{aligned}
$$

Given (B.9), we obtain

(a') $\ln\left(h^{-1}r_{NT}^{-1}\delta_N(m_{NT})m_{NT}\right)\sim\ln(NT)$ and $Thr_{NT}^2/(m_{NT}^2\kappa_{NT}^2)\sim N^{\ell_0}\ln(NT)$ for some $\ell_0=\ell_0(q,\beta)>0$, and therefore,

$$
h^{-1}r_{NT}^{-1}\delta_N(m_{NT})m_{NT}\exp\left(-\frac{Thr_{NT}^2}{m_{NT}^2\kappa_{NT}^2}C\right)=\exp\left\{\ln\left(h^{-1}r_{NT}^{-1}\delta_N(m_{NT})m_{NT}\right)-\frac{Thr_{NT}^2}{m_{NT}^2\kappa_{NT}^2}C\right\}\to 0,
$$

for any $C>0$.

(b') $h^{-1}r_{NT}^{-2}\left[\alpha(m_{NT})\right]^{1-1/q}\left[\delta_N(m_{NT})\right]^2\le C\left[\ln(NT)\right]^{-1}\to 0$;

(c') $h^{-1}r_{NT}^{-2}\kappa_{NT}^{1-q}\left[\delta_N(m_{NT})\right]^{q+1}\le C\left[\ln(NT)\right]^{-1}\to 0$.

Therefore, $K_{NT,1} = O_p(r_{NT})$. For the term $K_{NT,2}$ in (B.1), we have

$$K_{NT,2} \le \left\{ \sup_{\tau\in[0,1]} \frac{1}{Th} \sum_{t\in\mathcal{I}_{h,u_0}(\tau)} w_t^k(\tau) \right\} \left\{ \sup_{t\ge 1} \mathbb{E} \left| Q_{N,t} - Q_{N,t}^{(m)} \right| \right\} = O\left(m_{NT}^{-\lambda/p} N^{-1/2} \right) = O\left(r_{NT}\right), \quad \text{(B.10)}$$

where the last step follows from Eq. (3.4) and

$$m_{NT}^{-\lambda/p} N^{-1/2} r_{NT}^{-1} \le C \left[(NT)^{\phi_\beta(q)} N^{-1} \right]^{-\lambda/[p\beta(1-1/q)]} (NT)^{\phi_\beta(q)/2} N^{-1/2} \sqrt{h} \frac{1}{\sqrt{\ln(NT)}} \le C.$$

Similarly, by applying Markov's inequality, we can deduce that $K_{NT,3} = O_p(r_{NT})$, which is the last term in (B.1). Summing up, we have (3.5). ■

Lemma B.1 *Recall* $\phi(x) = \dfrac{(1-1/x)\varphi_\alpha}{2+(1+1/x)\varphi_\alpha}$ *in Assumption A8, where* $x > 1$. *Under Assumptions A1, A3, A5(c), A7(a), A8, for any nonnegative integers* $k, k_1, k_2 \ge 0$,

$$\sup_{\tau\in[0,1]} \left\| \frac{1}{NTh} \sum_{i=1}^N \sum_{t=1}^T w_t^{k_1}(\tau) \boldsymbol{\ell}^{k_2}(\tau_t) M_{it} - \frac{1}{NTh} \sum_{i=1}^N \sum_{t=1}^T w_t^{k_1}(\tau) \boldsymbol{\ell}^{k_2}(\tau_t) \mathbb{E}\left(M_{it}\right) \right\| = O_p\left(R_{1,NT}\right), \quad \text{(B.11)}$$

$$\sup_{\tau\in[0,1]} \left\| \frac{1}{NTh} \sum_{i=1}^N \sum_{t=1}^T w_t^{k_1}(\tau) \boldsymbol{\ell}^{k_2}(\tau_t) M_{it} \boldsymbol{\nu}_{it}' - \frac{1}{NTh} \sum_{i=1}^N \sum_{t=1}^T w_t^{k_1}(\tau) \boldsymbol{\ell}^{k_2}(\tau_t) \mathbb{E}\left(M_{it} \boldsymbol{\nu}_{it}'\right) \right\| = O_p\left(R_{2,NT}\right), \quad \text{(B.12)}$$

$$\sup_{\tau\in[0,1]} \left\| \frac{1}{NTh} \sum_{i=1}^N \sum_{t=1}^T w_t^{k}(\tau) M_{it} \boldsymbol{\nu}_{it} \boldsymbol{\nu}_{it}' - \frac{1}{NTh} \sum_{i=1}^N \sum_{t=1}^T w_t^{k}(\tau) \mathbb{E}\left(M_{it} \boldsymbol{\nu}_{it} \boldsymbol{\nu}_{it}'\right) \right\| = O_p\left(R_{3,NT}\right), \quad \text{(B.13)}$$

where $R_{i,NT} = \sqrt{\dfrac{\ln(NT)}{(NT)^{\eta_i} h}}$, *with* $\eta_1 = \phi(m_0)$, $\eta_2 = \phi(q_0)$, *and* $\eta_3 = \phi(r_0)$.

Proof of Lemma B.1 By inspecting each element in the vector/matrix, a straightforward application of Theorem 1 establishes the lemma. It is worth noting that for the process $\{M_{it},\, t \ge 1\}$, one can simply set $d_i \equiv 0$, for $i = 1, \ldots, N$, in Eq. (3.2). The details of these steps are omitted. ■

Lemma B.2 *Under Assumptions A1, A3, A5 - A8, for any nonnegative integer* $k \ge 0$,

$$\sup_{\tau\in[0,1]} \left\| \frac{1}{NTh} \sum_{i=1}^N \sum_{t=1}^T w_t^{k}(\tau) M_{it} \boldsymbol{x}_{it} - \frac{1}{NTh} \sum_{i=1}^N \sum_{t=1}^T w_t^{k}(\tau) \mathbb{E}\left(M_{it} \boldsymbol{x}_{it}\right) \right\| = O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{\ln(NT)}{(NT)^{\phi(m_0\wedge q_0)} h}} \right), \quad \text{(B.14)}$$

$$\sup_{\tau\in[0,1]} \left\| \frac{1}{NTh} \sum_{i=1}^N \sum_{t=1}^T w_t^{k}(\tau) M_{it} \boldsymbol{x}_{it} \boldsymbol{x}_{it}' - \frac{1}{NTh} \sum_{i=1}^N \sum_{t=1}^T w_t^{k}(\tau) \mathbb{E}\left(M_{it} \boldsymbol{x}_{it} \boldsymbol{x}_{it}'\right) \right\| = O_p\left(\frac{1}{\sqrt{N}} + \sqrt{\frac{\ln(NT)}{(NT)^{\eta_{min}} h}} \right), \quad \text{(B.15)}$$

where $\phi(\cdot)$ *and* η_{min} *are defined in Lemma B.1.*

Proof of Lemma B.2 By definition, the LHS of (B.14) is bounded by $\sum_{j=1}^3 \boldsymbol{\Xi}_{NT,j}(M, \boldsymbol{x})$, where, by

Lemma B.1, $\varXi_{NT,1}(M,\boldsymbol{x}) = \sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)\boldsymbol{\ell}(\tau_t)\left[M_{it}-\mathbb{E}\left(M_{it}\right)\right]\right\| = O_p\left(R_{1,NT}\right)$, $\varXi_{NT,2}(M,\boldsymbol{x}) = \sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)\left[M_{it}\boldsymbol{\nu}_{it}-\mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}\right)\right]\right\| = O_p\left(R_{2,NT}\right)$. Moreover, by Assumption A7(b),

$$\begin{aligned}\varXi_{NT,3}(M,\boldsymbol{x}) &= \sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)\left[M_{it}\boldsymbol{\chi}_i-\mathbb{E}\left(M_{it}\boldsymbol{\chi}_i\right)\right]\right\| \\ &= \sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)M_{it}\boldsymbol{\chi}_i\right\| \le \sup_{\tau\in[0,1]}\frac{1}{Th}\sum_{t=1}^{T}\left|w_t^k(\tau)\right|\left\|\frac{1}{N}\sum_{i=1}^{N}M_{it}\boldsymbol{\chi}_i\right\|.\end{aligned}$$

By Assumptions A5(b) and A7(b),

$$\begin{aligned}\mathbb{E}\left\|N^{-1}\sum_{i=1}^{N}M_{it}\boldsymbol{\chi}_i\right\|^2 &= N^{-2}\sum_{j=1}^{d}\mathbb{E}\left(\sum_{i=1}^{N}M_{it}\chi_{i,j}\right)^2 \\ &= N^{-2}\sum_{j=1}^{d}\mathbb{E}\left(\sum_{i=1}^{N}M_{it}\chi_{i,j}^2+\sum_{i\neq\ell}M_{it}M_{\ell t}\chi_{i,j}\chi_{\ell,j}\right) \le CN^{-1}.\end{aligned}$$

By Markov's inequality, we have $\left\|N^{-1}\sum_{i=1}^{N}M_{it}\boldsymbol{\chi}_i\right\| = O_p\left(N^{-1/2}\right)$ uniformly in τ. Then, $\varXi_{NT,3}(M,\boldsymbol{x}) = O_p\left(N^{-1/2}\right)$, and thus (B.14) is obtained by noting $O_p\left(\sum_{i=1}^{2}R_{i,NT}\right) = O_p\left(\sqrt{\ln(NT)/[(NT)^{\phi(m_0\wedge q_0)}h]}\right)$.

Similarly, the LHS of (B.15) is bounded by $\sum_{j=1}^{6}\varXi_{NT,j}(M,\boldsymbol{x}\boldsymbol{x}')$, where

$$\begin{aligned}\varXi_{NT,1}(M,\boldsymbol{x}\boldsymbol{x}') &= \sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)\boldsymbol{\ell}(\tau_t)\boldsymbol{\ell}(\tau_t)'\left[M_{it}-\mathbb{E}\left(M_{it}\right)\right]\right\| = O_p\left(R_{1,NT}\right), \\ \varXi_{NT,2}(M,\boldsymbol{x}\boldsymbol{x}') &= 2\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)\boldsymbol{\ell}(\tau_t)\left[M_{it}\boldsymbol{\nu}_{it}'-\mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}'\right)\right]\right\| = O_p\left(R_{2,NT}\right), \\ \varXi_{NT,3}(M,\boldsymbol{x}\boldsymbol{x}') &= \sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)\left[M_{it}\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}'-\mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}'\right)\right]\right\| = O_p\left(R_{3,NT}\right),\end{aligned}$$

using Lemma B.1. Note that, by Assumptions A5 and A7(b), we have $\mathbb{E}\left\|N^{-1}\sum_{i=1}^{N}M_{it}\left(\boldsymbol{\chi}_i\boldsymbol{\chi}_i'-\mathbb{E}\left(\boldsymbol{\chi}_i\boldsymbol{\chi}_i'\right)\right)\right\|^2 \le CN^{-1}$ and $\mathbb{E}\left\|N^{-1}\sum_{i=1}^{N}\left[M_{it}\boldsymbol{\chi}_i\boldsymbol{\nu}_{it}'-\mathbb{E}\left(M_{it}\boldsymbol{\chi}_i\boldsymbol{\nu}_{it}'\right)\right]\right\|^2 = \mathbb{E}\left\|N^{-1}\sum_{i=1}^{N}M_{it}\boldsymbol{\chi}_i\boldsymbol{\nu}_{it}'\right\|^2 \le CN^{-1}$. Using similar arguments for $\varXi_{NT,3}(M,\boldsymbol{x})$ above, we obtain

$$\begin{aligned}\varXi_{NT,4}(M,\boldsymbol{x}\boldsymbol{x}') &= \sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)\left[M_{it}\boldsymbol{\chi}_i\boldsymbol{\chi}_i'-\mathbb{E}\left(M_{it}\boldsymbol{\chi}_i\boldsymbol{\chi}_i'\right)\right]\right\| = O_p\left(N^{-1/2}+R_{1,NT}\right), \\ \varXi_{NT,5}(M,\boldsymbol{x}\boldsymbol{x}') &= 2\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)\boldsymbol{\ell}(\tau_t)\left[M_{it}\boldsymbol{\chi}_i'-\mathbb{E}\left(M_{it}\boldsymbol{\chi}_i'\right)\right]\right\| = O_p\left(N^{-1/2}\right), \\ \varXi_{NT,6}(M,\boldsymbol{x}\boldsymbol{x}') &= 2\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)\left[M_{it}\boldsymbol{\chi}_i\boldsymbol{\nu}_{it}'-\mathbb{E}\left(M_{it}\boldsymbol{\chi}_i\boldsymbol{\nu}_{it}'\right)\right]\right\| = O_p\left(N^{-1/2}\right).\end{aligned}$$

Eq. (B.15) is then obtained by combining these results. ■

Lemma B.3 *Under Assumptions A1, A5(a), and A6, for any nonnegative integer $k \geq 0$,*

(a) $\sup_{\tau\in[0,1]} \left|(NTh)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} w_t^k(\tau) p_i(\tau_t) - \mu_k \bar{p}(\tau)\right| = O\left(h^2 + (Th^2)^{-1} + \phi_{p,N}\right)$,

(b) $\sup_{\tau\in[0,1]} \left\|(NTh)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} w_t^k(\tau) p_i(\tau_t) \boldsymbol{\ell}(\tau_t) - \mu_k \bar{p}(\tau)\boldsymbol{\ell}(\tau)\right\| = O\left(h^2 + (Th^2)^{-1} + \phi_{p,N}\right)$,

(c) $\sup_{\tau\in[0,1]} \left\|(NTh)^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} w_t^k(\tau) p_i(\tau_t) \boldsymbol{\ell}(\tau_t)\boldsymbol{\ell}(\tau_t)' - \mu_k \bar{p}(\tau)\boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)'\right\| = O\left(h^2 + (Th^2)^{-1} + \phi_{p,N}\right)$,

where $\phi_{p,N}$ is given in Assumption A6.

Proof of Lemma B.3 The following result is used repeatedly:

$$\left|T^{-1}\sum_{t=1}^{T} g\left(t/T\right) - \int_0^1 g(z)\,\mathrm{d}z\right| \leq \sup_{|x-y|\leq T^{-1}} |g(x)-g(y)|\,, \tag{B.16}$$

where $g(\cdot)$ is continuous and Riemann-integrable (Bühlmann, 1998, Eq. (6.5)). We only show Part (c) as the others are similar. By (B.16), Assumption A6, and that $\bar{p}(\cdot)\boldsymbol{\ell}(\cdot)\boldsymbol{\ell}(\cdot)' \in \mathcal{C}^2[0,1]$,

$$\begin{aligned}
\frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)\frac{1}{N}\sum_{i=1}^{N} p_i(\tau_t)\boldsymbol{\ell}(\tau_t)\boldsymbol{\ell}(\tau_t)' &= \frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)\bar{p}(\tau_t)\boldsymbol{\ell}(\tau_t)\boldsymbol{\ell}(\tau_t)' + O(\phi_{p,N}) \\
&= \int_0^1 h^{-1}\left(\frac{z-\tau}{h}\right)^k K\left(\frac{z-\tau}{h}\right)\bar{p}(z)\boldsymbol{\ell}(z)\boldsymbol{\ell}(z)'\,\mathrm{d}z + O\left(\frac{1}{Th^2}\right) + O(\phi_{p,N}) \\
&= \int_{-1}^1 u^k K(u)\bar{p}(\tau+uh)\boldsymbol{\ell}(\tau+uh)\boldsymbol{\ell}(\tau+uh)'\,\mathrm{d}u + O\left(\frac{1}{Th^2}\right) + O(\phi_{p,N}) \\
&= \mu_k\bar{p}(\tau)\boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)' + O(h^2) + O\left(\frac{1}{Th^2}\right) + O(\phi_{p,N}),
\end{aligned}$$

where the final step follows from a Taylor expansion of each element of $\bar{p}(\tau+uh)\boldsymbol{\ell}(\tau+uh)\boldsymbol{\ell}(\tau+uh)'$ around the corresponding element of $\bar{p}(\tau)\boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)'$, and $\mu_1 = \int_{-1}^1 uK(u)\,\mathrm{d}u = 0$. Note that the asymptotic terms $O(\cdot)$ and $o(\cdot)$ are uniform in τ, and thus Part (c) follows. ■

Recall the definition of $\nu_{\tau,i}$ Eq. (2.8). When T is sufficiently large, Assumption A6 ensures the presence of at least one observation in the h-neighborhood $\tau \in [0,1]$, we have $\nu_{\tau,i} \neq 0$. Without loss of generality, we will proceed with the assumption $\nu_{\tau,i} = \left(\sum_{t=1}^{T} w_t^0(\tau) M_{it}\right)^{-1}$.

Lemma B.4 *Recall $\phi(x) = \dfrac{(1-1/x)\varphi_\alpha}{2+(1+1/x)\varphi_\alpha}$ and $\varpi(x) = \dfrac{4(1+\varphi_\alpha/x)}{2+(1+1/x)\varphi_\alpha}$ in Assumption A8. Note that $\phi(\cdot)$ and $\varpi(\cdot)$ are strictly increasing and decreasing functions, respectively. Suppose Assumptions A1, A3, and A5(c), are imposed for all the parts below. Let $k, k_1, k_2 \geq 0$ be any nonnegative integers (if not specified).*

(a) $\max\limits_{1\leq i\leq N}\sup\limits_{\tau\in[0,1]} \left|(Th)^{-1}\sum_{t=1}^{T} w_t^k(\tau)\left[M_{it} - \mathbb{E}(M_{it})\right]\right| = O_p\left(\sqrt{\dfrac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}\right)$, $\eta_{max} = \lim\limits_{x\to+\infty}\phi(x) = \varphi_\alpha/(2+\varphi_\alpha)$. *One has* $\dfrac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h} \leq \dfrac{N^{\varpi(p_0/2)}\ln(NT)}{T^{\phi(p_0/2)}h} \to 0$, *as $(N,T)\to\infty$, by Assumption A8.*

(b) $\max_{1\le i\le N}\sup_{\tau\in[0,1]}(Th\nu_{\tau,i}) = O_p(1)$.

(c) $\max_{1\le i\le N}\sup_{\tau\in[0,1]}\left|(Th)^{-1}\sum_{t=1}^{T} w_t^k(\tau)M_{it}\right| = O_p\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}} + h + \frac{1}{Th^2}\right)$ *if k is odd, and is $O_p(1)$ if k is even.*

(d) $\max_{1\le i\le N}\sup_{\tau\in[0,1]}\left\|(Th)^{-1}\sum_{t=1}^{T} w_t^k(\tau)\left[M_{it}\boldsymbol{\nu}_{it} - \mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}\right)\right]\right\| = O_p\left(\sqrt{\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}}\right)$. *By Assumption A8, one has* $\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h} < \frac{N^{\varpi(p_0/2)}\ln(NT)}{T^{\phi(p_0/2)}h} \to 0$, *as* $(N,T)\to\infty$.

(e) $\max_{1\le i\le N}\sup_{\tau\in[0,1]}\left\|(Th)^{-1}\sum_{t=1}^{T} w_t^{k_1}(\tau)w_t^{k_2}(\tau)\left[M_{it}\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}' - \mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}'\right)\right]\right\| = O_p\left(\sqrt{\frac{N^{\varpi(p_0/2)}\ln(NT)}{T^{\phi(p_0/2)}h}}\right)$.

(f) $\max_{1\le i\le N}\sup_{\tau\in[0,1]}\left\|(Th)^{-1}\sum_{t=1}^{T} w_t^k(\tau)M_{it}\boldsymbol{x}_{it}\right\| = O_p\left(\sqrt{\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}} + h + \frac{1}{Th^2}\right) = o_p(1)$ *for an odd integer k and is $O_p(1)$ for an even k, if Assumptions A5(b), A7(b), and A8 hold.*

(g) $\sup_{\tau\in[0,1]}\left\|N^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^k(\tau)M_{it}\boldsymbol{\nu}_{it}\right\| = O_p\left(\sqrt{\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}}\right)$.

(h) $\sup_{\tau\in[0,1]}\left|ThN^{-1}\omega_\tau - \bar{q}(\tau)\right| = O_p\left(\phi_{q,N} + h + \frac{1}{Th^2} + \sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}\right)$ *if Assumption A6 holds.*

(i) $\sup_{\tau\in[0,1]}\left|N^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^k(\tau)M_{it} - \mu_k\right| = O_p\left(h + \frac{1}{Th^2} + \sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}\right)$.

Proof of Lemma B.4 We start by showing Part (a). The steps are similar to Proof of Theorem 1. That is, we cover the interval $[0,1]$ by a finite number of subintervals $\bar{B}_l$, $l = 1,\ldots,\bar{L}_{NT}$, which are centered at $\bar{b}_l$ with radius $h\varrho_{NT}$, where $\varrho_{NT} = \left[T^{-\eta_{max}}h^{-1}N^{1-\eta_{max}}\ln(NT)\right]^{1/2}$. It is easy to obtain

$$\max_{1\le i\le N}\sup_{\tau\in[0,1]}\left|\frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)\left[M_{it}-\mathbb{E}(M_{it})\right]\right| \le O_p(\varrho_{NT}) + \max_{1\le i\le N}\max_{1\le l\le L_{NT}}\left|\frac{1}{Th}\sum_{t\in\mathcal{I}_h(\tau)} w_t^k(b_l)\left[M_{it}-\mathbb{E}(M_{it})\right]\right|.$$

Let $\mathcal{F}_{\xi,t_1}^{t_2} = \sigma\left(\boldsymbol{\xi}_{\cdot t},\, t_1\le t\le t_2\right)$. By Bonferroni's inequality and the Triplex inequality (Jiang 2009, Theorem 1), for any $M_\epsilon, \bar{\delta}_{NT} > 0$ and positive integer $\bar{q}_{NT}$,

$$\begin{aligned}
&\mathbb{P}\left(\max_{1\le i\le N}\max_{1\le l\le L_{NT}}\left|\frac{1}{Th}\sum_{t\in\mathcal{I}_h(\tau)} w_t^k(b_l)\left[M_{it}-\mathbb{E}(M_{it})\right]\right| \ge 2M_\epsilon\varrho_{NT}\right)\\
&\le \sum_{i=1}^{N}\sum_{l=1}^{L_{NT}}\Bigg\{2\bar{q}_{NT}\exp\left(-\frac{2Th\varrho_{NT}^2}{\bar{q}_{NT}^2\bar{\delta}_{NT}^2}\frac{M_\epsilon^2}{288}\right) + \frac{6}{M_\epsilon}\frac{1}{\varrho_{NT}}\frac{1}{2Th}\sum_{t\in\mathcal{I}_h(\tau)} w_t^k(b_l)\mathbb{E}\left|\mathbb{E}\left(M_{it}\,|\,\mathcal{F}_{\xi,-\infty}^{t-\bar{q}_{NT}}\right) - \mathbb{E}\left(M_{it}\right)\right|\\
&\quad + \frac{15}{M_\epsilon}\frac{1}{\varrho_{NT}}\frac{1}{2Th}\sum_{t\in\mathcal{I}_h(\tau)} w_t^k(b_l)\mathbb{E}\left(M_{it}\mathbb{1}\left\{\left|w_t^k(b_l)M_{it}\right| > \bar{\delta}_{NT}\right\}\right)\Bigg\}.
\end{aligned}\tag{B.17}$$

Putting $p=1$ and $r=\infty$ in Theorem 14.2 by Davidson (1994), we have

$$\frac{1}{\varrho_{NT}}\frac{1}{2Th}\sum_{t\in\mathcal{I}_h(\tau)} w_t^k(b_l)\mathbb{E}\left|\mathbb{E}\left(M_{it}\,|\,\mathcal{F}_{\xi,-\infty}^{t-\bar{q}_{NT}}\right) - \mathbb{E}\left(M_{it}\right)\right| \le C\varrho_{NT}^{-1}\alpha\left(\bar{q}_{NT}\right).$$

Moreover, by Markov's inequality, for any $\bar{r} > 0$,

$$\frac{1}{\varrho_{NT}} \frac{1}{2Th} \sum_{t \in \mathcal{I}_h(\tau)} w_t^k(b_l) \mathbb{E}\left(M_{it} \mathbb{1}\left\{\left|w_t^k(b_l) M_{it}\right| > \bar{\delta}_{NT}\right\}\right) \leq C \bar{\delta}_{NT}^{-\bar{r}} \varrho_{NT}^{-1}.$$

By (B.17), one obtains $\max_{1 \leq i \leq N} \max_{1 \leq l \leq L_{NT}} \left|(Th)^{-1} \sum_{t \in \mathcal{I}_h(\tau)} w_t^k(b_l) \left[M_{it} - \mathbb{E}(M_{it})\right]\right| = O_p(\varrho_{NT})$ if for any $\epsilon > 0$, there exists $M_\epsilon > 0$ such that

$$2NL_{NT}\bar{q}_{NT} \exp\left(-\frac{Th\varrho_{NT}^2}{\bar{q}_{NT}^2 \bar{\delta}_{NT}^2} \frac{M_\epsilon^2}{144}\right) + \tilde{C}_1 NL_{NT} \varrho_{NT}^{-1} \alpha\left(\bar{q}_{NT}\right) + \tilde{C}_2 NL_{NT} \bar{\delta}_{NT}^{-\bar{r}} \varrho_{NT}^{-1} \leq \epsilon, \tag{B.18}$$

where $\tilde{C}_1, \tilde{C}_2 > 0$ are some constants.

Let $\bar{q}_{NT} = \left\lceil \left(\frac{N \ln(NT)}{h \varrho_{NT}^2}\right)^{1/\varphi_\alpha} \right\rceil$ and $\bar{\delta}_{NT} = \left(\frac{N \ln(NT)}{h \varrho_{NT}^2}\right)^{1/\bar{r}}$. Note that $L_{NT} \leq K_0 h^{-1} \varrho_{NT}^{-1}$ for some $K_0 > 0$. Simple linear algebra leads to

$$\ln\left(2NL_{NT}\bar{q}_{NT}\right) < \ln(N) + \ln(NT) < 2\ln(NT), \qquad \text{as} \quad (N,T) \to \infty.$$

Moreover, we have $\frac{Th\varrho_{NT}^2}{\bar{q}_{NT}^2 \bar{\delta}_{NT}^2} \sim (NT)^{-2\eta_{max}/\bar{r}} \ln(NT)$. Since $\bar{r} > 0$ is arbitrary, we can take $\bar{r} = \bar{r}_{NT} = NT \to \infty$, as $(N,T) \to \infty$, leading to $(NT)^{-2\eta_{max}/\bar{r}} \to 1$. Therefore, for any $\epsilon > 0$, one can always find an $M_\epsilon > 0$ such that

$$2NL_{NT}\bar{q}_{NT} \exp\left(-\frac{Th\varrho_{NT}^2}{\bar{q}_{NT}^2 \bar{\delta}_{NT}^2} \frac{M_\epsilon^2}{144}\right) < \exp\left(2\ln(NT) - \frac{Th\varrho_{NT}^2}{\bar{q}_{NT}^2 \bar{\delta}_{NT}^2} \frac{M_\epsilon^2}{144}\right) \leq \epsilon/3.$$

Finally, it is easy to obtain $\tilde{C}_1 NL_{NT} \varrho_{NT}^{-1} \alpha\left(\bar{q}_{NT}\right) \leq C \left[\ln(NT)\right]^{-1} \leq \epsilon/3$ and $\tilde{C}_2 NL_{NT} \bar{\delta}_{NT}^{-\bar{r}} \varrho_{NT}^{-1} \leq C \left[\ln(NT)\right]^{-1} \leq \epsilon/3$, as $(N,T) \to \infty$, for any $\tilde{C}_1, \tilde{C}_2 > 0$. Combining these results, we obtain (B.18).

Next, we show Part *(b)*. By Part *(a)* and Assumption A8, we have

$$\frac{1}{Th} \sum_{t=1}^{T} w_t^0(\tau) M_{it} = \frac{1}{Th} \sum_{t=1}^{T} w_t^0(\tau) p_i(\tau_t) + o_p(1) \geq p_L \mu_0 + o_p(1), \tag{B.19}$$

for all i, where the o_p-terms are uniform in τ and i. That is, $(Th)^{-1} \sum_{t=1}^{T} w_t^0(\tau) M_{it}$ is bounded below away from zero with arbitrarily large probability as $(N,T) \to \infty$, uniformly in τ and i. Therefore, we have $\max_{1 \leq i \leq N} \sup_{\tau \in [0,1]} (Th\nu_{\tau,i}) \leq C$ with arbitrarily large probability.

Part *(c)* is obtained using Part *(a)* and the Riemann sum approximation (B.16).

The proof of Part *(d)* is similar to Part *(a)* and Theorem 1; we only sketch the steps. We first construct open subintervals centered at b_l with the radius $hN^{-1}\tilde{\varrho}_{NT}$, where $\tilde{\varrho}_{NT} = \sqrt{N^{\varpi(p_0)} \ln(NT)/(T^{\phi(p_0)} h)}$.

Using $\mathbb{E}\|M_{it}\boldsymbol{\nu}_{it}\| \leq \mathbb{E}\|\boldsymbol{\nu}_{it}\| \leq C$, it is easy to obtain

$$\begin{aligned}
&\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left\|\frac{1}{Th}\sum_{t=1}^{T}w_t^k(\tau)\left[M_{it}\boldsymbol{\nu}_{it}-\mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}\right)\right]\right\| \\
&\leq O_p(\tilde{\varrho}_{NT})+\max_{1\leq i\leq N}\max_{1\leq l\leq L_{NT}}\left\|\frac{1}{Th}\sum_{t\in\mathcal{I}_h(\tau)}w_t^k(b_l)\left[M_{it}\boldsymbol{\nu}_{it}-\mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}\right)\right]\right\| \\
&\leq O_p(\tilde{\varrho}_{NT})+O_p\left(\sum_{i=1}^{N}\left[\psi_{i,p_0}(m)\right]^{1/p_0}\right)+\max_{1\leq i\leq N}\max_{1\leq l\leq L_{NT}}\left\|\frac{1}{Th}\sum_{t\in\mathcal{I}_h(\tau)}w_t^k(b_l)\left[M_{it}\boldsymbol{\nu}_{it}^{(m)}-\mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}^{(m)}\right)\right]\right\|,
\end{aligned}$$

where $m=m_{NT}>0$. By the Triplex inequality again, there exist some constants $C_1, C_2>0$ such that

$$\begin{aligned}
&\mathbb{P}\left(\max_{1\leq i\leq N}\max_{1\leq l\leq L_{NT}}\left\|\frac{1}{Th}\sum_{t\in\mathcal{I}_h(\tau)}w_t^k(b_l)\left[M_{it}\boldsymbol{\nu}_{it}^{(m)}-\mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}^{(m)}\right)\right]\right\|\geq 2M_\epsilon\tilde{\varrho}_{NT}\right) \\
&\leq d\left\{4NL_{NT}m_{NT}\exp\left(-\frac{2Th\tilde{\varrho}_{NT}^2}{4m_{NT}^2\kappa_{NT}^2}\frac{M_\epsilon^2}{288d}\right)+C_1NL_{NT}\tilde{\varrho}_{NT}^{-1}\left[\alpha(m_{NT})\right]^{1-1/p_0}+C_2NL_{NT}\tilde{\varrho}_{NT}^{-1}\kappa_{NT}^{1-p_0}\right\}.
\end{aligned}$$

Let $m_{NT}=\left\lceil\left(N^{2-\varpi(p_0)}T^{\phi(p_0)}\right)^{1/[(1-1/p_0)\varphi_\alpha]}\right\rceil$ and $\kappa_{NT}=\left(N^{2-\varpi(p_0)}T^{\phi(p_0)}\right)^{1/(p_0-1)}$. For any $\epsilon>0$, it is not hard to show the second and third terms in the brackets are bounded by $C\left[\ln(NT)\right]^{-1}<\epsilon/3$. By somewhat cumbersome linear algebra, we find that the construction of $\varpi(\cdot)$ ensures that the first term

$$4NL_{NT}m_{NT}\exp\left(-\frac{2Th\tilde{\varrho}_{NT}^2}{4m_{NT}^2\kappa_{NT}^2}\frac{M_\epsilon^2}{288d}\right)<\exp\Big(K_1\ln(NT)-K_2M_\epsilon^2\ln(NT)\Big), \tag{B.20}$$

where $K_1, K_2>0$ are some constants. For any $\epsilon>0$, one can choose a sufficiently large $M_\epsilon>0$ such that (B.20) is bounded by $\epsilon/3$. Summing up, we obtain

$$\max_{1\leq i\leq N}\max_{1\leq l\leq L_{NT}}\left\|\frac{1}{Th}\sum_{t\in\mathcal{I}_h(\tau)}w_t^k(b_l)\left[M_{it}\boldsymbol{\nu}_{it}^{(m)}-\mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}^{(m)}\right)\right]\right\|=O_p\left(\tilde{\varrho}_{NT}\right).$$

Given the chosen m_{NT} above, simple linear algebra gives

$$\sum_{i=1}^{N}\left[\psi_{i,p_0}(m_{NT})\right]^{1/p_0}\tilde{\varrho}_{NT}^{-1}\leq Cm_{NT}^{-\varphi_\nu/p_0}N^{1/2}\tilde{\varrho}_{NT}^{-1}\leq C\sqrt{h/\ln(NT)}\left(N\sqrt{T}\right)^{\eta_0}N^{-1/2}\to 0,$$

where $\eta_0=\dfrac{(p_0-1)\varphi_\alpha-2\varphi_\nu}{2p_0+(p_0+1)\varphi_\alpha}<0$ by the condition $2\varphi_\nu\geq p_0\varphi_\alpha$ as required in Assumption A8. We then conclude that $O_p\left(\sum_{i=1}^{N}\left[\psi_{i,p_0}(m)\right]^{1/p_0}\right)=o_p(\tilde{\varrho}_{NT})$. Combining these results, we obtain Part (d).

Note that $x\mapsto x^\ell K^2(x)$ is Lipschitz continuous on $[-1,1]$ for any nonnegative integer ℓ, and

$\mathbb{E}\left\|M_{it}\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}'\right\| \leq \mathbb{E}\left\|\boldsymbol{\nu}_{it}\right\|^2 \leq C$. Moreover,

$$\begin{aligned}
&\left\{\mathbb{E}\left\|M_{it}\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}' - M_{it}\boldsymbol{\nu}_{it}^{(m)}\boldsymbol{\nu}_{it}^{(m)\prime}\right\|^{p_0/2}\right\}^{2/p_0} \\
&\leq C\left\{\mathbb{E}\left\|\boldsymbol{\nu}_{it} - \boldsymbol{\nu}_{it}^{(m)}\right\|^{p_0} + 2\left(\mathbb{E}\left\|\boldsymbol{\nu}_{it} - \boldsymbol{\nu}_{it}^{(m)}\right\|^{p_0}\right)^{1/2}\left(\mathbb{E}\left\|\boldsymbol{\nu}_{it}\right\|^{p_0}\right)^{1/2}\right\}^{2/p_0} \leq C\left[\psi_{i,p_0}(m)\right]^{1/p_0}. \quad \text{(B.21)}
\end{aligned}$$

By (B.21), Part (e) follows from similar arguments above.

Next, we note that Part (f) can be bounded by

$$\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left\|\frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)\mathbb{E}\left(M_{it}\boldsymbol{x}_{it}\right)\right\| + \max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left\|\frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)\left[M_{it}\boldsymbol{x}_{it} - \mathbb{E}\left(M_{it}\boldsymbol{x}_{it}\right)\right]\right\|. \quad \text{(B.22)}$$

By the moment conditions in Assumptions A5(b) - (c), and A7(b), using the Riemann sum approximation (B.16), $\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left\|(Th)^{-1}\sum_{t=1}^{T} w_t^k(\tau)\mathbb{E}\left(M_{it}\boldsymbol{x}_{it}\right)\right\| = O\left(1/(Th^2)\right) + O(h)$ if k is odd, and is $O(1)$ if k is even. Moreover, the second term in (B.22) can be bounded by

$$\begin{aligned}
&\left(\max_{1\leq i\leq N}\left\|\boldsymbol{\chi}_i\right\|\right)\left(\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left\|\frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)M_{it}\right\|\right) + \max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left\|\frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)\boldsymbol{\ell}(\tau_t)\left[M_{it} - \mathbb{E}\left(M_{it}\right)\right]\right\| \\
&\quad + \max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left\|\frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)\left[M_{it}\boldsymbol{\nu}_{it} - \mathbb{E}\left(M_{it}\boldsymbol{\nu}_{it}\right)\right]\right\| \\
&= \left(\max_{1\leq i\leq N}\left\|\boldsymbol{\chi}_i\right\|\right)\left(\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left\|\frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)M_{it}\right\|\right) + O_p\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}\right) \\
&\quad + O_p\left(\sqrt{\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}}\right),
\end{aligned}$$

where the final equality follows from a straightforward modification (to include the components of $\boldsymbol{\ell}(\cdot)$ as in (B.11)) of Part (a), and Part (d) above. By $\max_{1\leq i\leq N}\left\|\boldsymbol{\chi}_i\right\| = O_p(1)$ in Assumption A5(b) and Part (c), we have

$$\left(\max_{1\leq i\leq N}\left\|\boldsymbol{\chi}_i\right\|\right)\left(\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left\|\frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)M_{it}\right\|\right) = O_p\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}} + h + \frac{1}{Th^2}\right)$$

if k is odd, and is $O_p(1)$ if k is even. Note that $\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}} \leq \sqrt{\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}}$. We obtain Part (f) by combining these results.

For Part (g), note that

$$\sup_{\tau\in[0,1]}\left\|\frac{1}{N}\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^k(\tau)M_{it}\boldsymbol{\nu}_{it}\right\| \leq \left\{\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left(Th\nu_{\tau,i}\right)\right\}\left\{\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left\|\frac{1}{Th}\sum_{t=1}^{T} w_t^k(\tau)M_{it}\boldsymbol{\nu}_{it}\right\|\right\}.$$

By Parts (b), (d), and Assumption A7(b), Part (g) is obtained.

Consider Part *(h)*. Note that $\mathbb{P}(\omega_\tau = 1) = \mathbb{P}\left(\sum_{i=1}^N \nu_{\tau,i} = 0\right) = \mathbb{P}\left(M_{it} = 0,\, \forall i = 1, \ldots, N,\, \forall t \in \mathcal{I}_h(\tau)\right) = o(1)$ by Assumption A6, as $(N,T) \to \infty$. It is only necessary to consider the case $\omega_\tau \neq 1$. Since $|p_i(\tau_t) - p_i(\tau)| \leq Ch$ whenever $|\tau_t - \tau| \leq h$ by Assumption A6, and that $K(\cdot)$ has the support $[-1,1]$, by Lemma B.4*(a)*,

$$\begin{aligned}
\frac{Th}{N}\omega_\tau &= \frac{1}{N}\sum_{i=1}^N \left\{\frac{1}{Th}\sum_{t\in\mathcal{I}_h(\tau)} K\left(\frac{\tau_t-\tau}{h}\right) p_i(\tau_t) + O_p\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}\right)\right\}^{-1} \\
&= \left(\frac{1}{Th}\sum_{t:|\tau_t-\tau|\leq h} K\left(\frac{\tau_t-\tau}{h}\right)\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^N p_i^{-1}(\tau)\right) + O(h) + O_p\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}\right) \\
&= \bar{q}(\tau) + O(\phi_{q,N}) + O\left(\frac{1}{Th^2}\right) + O(h) + O_p\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}\right),
\end{aligned}$$

where the asymptotic terms $O(\cdot)$ and $o(\cdot)$ are uniform in τ.

Finally, for Part *(i)*, note that

$$\sup_{\tau\in[0,1]}\left|\frac{1}{N}\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^k(\tau)M_{it} - \mu_k\right| \leq \frac{1}{N}\sum_{i=1}^N \sup_{\tau\in[0,1]}\left|\nu_{\tau,i}\sum_{t=1}^T w_t^k(\tau)M_{it} - \mu_k\right|.$$

By $|p_i(\tau_t) - p_i(\tau)| \leq Ch$ if $|\tau_t - \tau| \leq h$, Part *(a)* above, the Riemann sum approximation (B.16), and Eq. (B.19),

$$\begin{aligned}
\sup_{\tau\in[0,1]}\left|\nu_{\tau,i}\sum_{t=1}^T w_t^k(\tau)M_{it} - \mu_k\right| &= \sup_{\tau\in[0,1]}\left|\frac{(Th)^{-1}\sum_{t=1}^T w_t^k(\tau)p_i(\tau_t)}{(Th)^{-1}\sum_{t=1}^T w_t^0(\tau)p_i(\tau_t)} - \mu_k\right| + O_p\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}\right) \\
&= \sup_{\tau\in[0,1]}\left|\frac{\mu_k p_i(\tau)}{\mu_0 p_i(\tau)} - \mu_k\right| + O\left(\frac{1}{Th^2}\right) + O(h) + O_p\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}\right) \\
&= O\left(\frac{1}{Th^2}\right) + O(h) + O_p\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}\right), \qquad \text{(B.23)}
\end{aligned}$$

where the $O(\cdot)/O_p(\cdot)$-terms are uniform in i. We conclude Part *(i)*. ■

Lemma B.5 *Define*

$$\boldsymbol{A}_{NT}(\tau) = \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\boldsymbol{x}_{it}' - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)',$$

$$\boldsymbol{B}_{NT}(\tau) = \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it} - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T} w_t^1(\tau)M_{it}\right),$$

$$\boldsymbol{C}_{NT}(\tau) = \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\boldsymbol{x}_{it}' - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)',$$

$$\boldsymbol{D}_{NT}(\tau) = \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^2(\tau)M_{it} - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^1(\tau)M_{it}\right)^2,$$

$$\boldsymbol{E}_{NT}(\tau) = \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^2(\tau)M_{it}\boldsymbol{x}_{it}' - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^1(\tau)M_{it}\right)\left(\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)',$$

$$\boldsymbol{F}_{NT}(\tau) = \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^2(\tau)M_{it}\boldsymbol{x}_{it}\boldsymbol{x}_{it}' - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)',$$

$$\boldsymbol{G}_{NT}^k(\tau) = \sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^k(\tau)M_{it}\boldsymbol{x}_{it}.$$

Under the assumptions in Lemmas B.1 - B.4, we have

(*a*) $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\boldsymbol{A}_{NT}(\tau) - \bar{p}(\tau)\boldsymbol{\Sigma}_\nu\right\| = o_p(1)$,

(*b*) $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\boldsymbol{B}_{NT}(\tau)\right\| = o_p(1)$,

(*c*) $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\boldsymbol{C}_{NT}(\tau)\right\| = o_p(1)$,

(*d*) $\sup_{\tau\in[0,1]}\left|(NTh)^{-1}\boldsymbol{D}_{NT}(\tau) - \mu_2\bar{p}(\tau)\right| = o_p(1)$,

(*e*) $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\boldsymbol{E}_{NT}(\tau) - \mu_2\bar{p}(\tau)\boldsymbol{\ell}(\tau)'\right\| = o_p(1)$,

(*f*) $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\boldsymbol{F}_{NT}(\tau) - \mu_2\bar{p}(\tau)\left[\boldsymbol{\Sigma}_\chi + \boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)' + \boldsymbol{\Sigma}_\nu\right]\right\| = o_p(1)$.

(*g*) $\sup_{\tau\in[0,1]}\left\|N^{-1}\boldsymbol{G}_{NT}^k(\tau) - \mu_k\boldsymbol{\ell}(\tau)\right\| = o_p(1)$, $k = 0, 1, 2$.

Proof of Lemma B.5 We start with Part (a). Recall that $\boldsymbol{x}_{it} = \boldsymbol{\chi}_i + \boldsymbol{\ell}(t/T) + \boldsymbol{\nu}_{it}$. One can decompose $\boldsymbol{A}_{NT}(\tau)$ into $\boldsymbol{A}_{NT}(\tau) = \boldsymbol{A}_{NT,1}(\tau) + \boldsymbol{A}_{NT,2}(\tau) + \boldsymbol{A}_{NT,2}(\tau)' + \boldsymbol{A}_{NT,3}(\tau)$, where

$$\boldsymbol{A}_{NT,1}(\tau) = \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\ell}(\tau_t)\boldsymbol{\ell}(\tau_t)' - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\ell}(\tau_t)\right)\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\ell}(\tau_t)\right)',$$

$$\boldsymbol{A}_{NT,2}(\tau) = \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\ell}(\tau_t)\boldsymbol{\nu}_{it}' - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\ell}(\tau_t)\right)\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\right)',$$

$$\boldsymbol{A}_{NT,3}(\tau) = \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}' - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\right)\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\right)'.$$

For the first component of $(NTh)^{-1}\boldsymbol{A}_{NT,1}(\tau)$, we have $(NTh)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\ell}(\tau_t)\boldsymbol{\ell}(\tau_t)' = \bar{p}(\tau)\boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)' + o_p(1)$ by Lemma B.3(c), where the o_p-term is uniform in τ. Moreover, since $K(\cdot)$

has the support $[-1,1]$,

$$\frac{1}{Th}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\ell}(\tau_t) = \frac{1}{Th}\sum_{t\in\mathcal{I}_h(\tau)} w_t^0(\tau)M_{it}\boldsymbol{\ell}(\tau_t) = \frac{1}{Th}\nu_{\tau,i}^{-1}\boldsymbol{\ell}(\tau) + O_p(h), \tag{B.24}$$

by a Taylor expansion of $\boldsymbol{\ell}(\tau_t)$ around $\boldsymbol{\ell}(\tau)$, where the O_p-term is uniform in τ. Using (B.24), we have

$$\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\ell}(\tau_t)\right)\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\ell}(\tau_t)\right)' - \bar{p}(\tau)\boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)'\right\| = o_p(1),$$

and finally, $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\boldsymbol{A}_{NT,1}(\tau)\right\| = o_p(1)$. By Assumptions A5, A7$(b)$, Eq. (B.12) in Lemma B.1, and (b), (d) in Lemma B.4, we have $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\boldsymbol{A}_{NT,2}(\tau)\right\| = o_p(1)$. Before continuing, note that, by Assumption A5(c),

$$\frac{1}{NTh}\left\|\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^k(\tau)p_i(\tau_t)\left[\mathbb{E}\left(\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}'\right) - \boldsymbol{\Sigma}_\nu\right]\right\| \le C\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbb{E}\left(\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}'\right) - \boldsymbol{\Sigma}_\nu\right\| = o(1), \quad k\ge 0. \tag{B.25}$$

By Assumption A7(b) and Eq. (B.13) in Lemma B.1, the first component of $(NTh)^{-1}\boldsymbol{A}_{NT,3}(\tau)$ can be written as

$$\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}' = \frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)p_i(\tau_t)\boldsymbol{\Sigma}_\nu + o_p(1) = \bar{p}(\tau)\boldsymbol{\Sigma}_\nu + o_p(1),$$

where the o_p-terms are uniform in τ. Note that $\mathbb{E}(\boldsymbol{\nu}_{it}) = \mathbf{0}$, by (b), (d) in Lemma B.4, Assumption A7(b),

$$\begin{aligned}\frac{1}{NTh}\sup_{\tau\in[0,1]}&\left\|\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\right)\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\right)'\right\| \\ &\le \sup_{\tau\in[0,1]}\frac{1}{N}\sum_{i=1}^{N}(Th\nu_{\tau,i})\left\|\frac{1}{Th}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\right\|^2 = O_p\left(\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}\right) = o_p(1).\end{aligned} \tag{B.26}$$

Therefore, $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\boldsymbol{A}_{NT,3}(\tau) - \bar{p}(\tau)\boldsymbol{\Sigma}_\nu\right\| = o_p(1)$. Combining these results leads to Part (a). By Assumptions A1, A5, A7(b), Lemma B.2, and (b), (c), (f) in Lemma B.4, it is simple to obtain Part (b):

$$\begin{aligned}\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\boldsymbol{B}_{NT}(\tau)\right\| \le& \sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right\| \\ &+ \sup_{\tau\in[0,1]}\frac{1}{N}\sum_{i=1}^{N}(Th\nu_{\tau,i})\left\|\frac{1}{Th}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right\|\left|\frac{1}{Th}\sum_{t=1}^{T} w_t^1(\tau)M_{it}\right| = o_p(1).\end{aligned}$$

Similarly, Part (c) can be obtained using Lemmas B.2, B.3, and (b), (f) in Lemma B.4.

Now we consider Part (d). By Lemmas B.1, B.3(a), and B.4(b) - (c),

$$\sup_{\tau\in[0,1]}\left|\frac{1}{NTh}\boldsymbol{D}_{NT}(\tau)-\mu_2\bar{p}(\tau)\right|\leq\left|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^2(\tau)M_{it}-\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^2(\tau)\mathbb{E}(M_{it})\right|$$
$$+\left|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^2(\tau)\mathbb{E}(M_{it})-\mu_2\bar{p}(\tau)\right|+\left|\frac{1}{N}\sum_{i=1}^{N}(Th\nu_{\tau,i})\left(\frac{1}{Th}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\right)^2\right|=o_p(1).$$

Part (e) is similar to Part (f) below, and thus omitted.

Next, by Lemmas B.2 - B.3, Assumptions A5, A7(b), and the identity $\mathbb{E}\left(\boldsymbol{x}_{it}\boldsymbol{x}_{it}'\right)=\boldsymbol{\Sigma}_\chi+\boldsymbol{\ell}(\tau_t)\boldsymbol{\ell}(\tau_t)'+\mathbb{E}\left(\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{it}'\right)$, the first component of $(NTh)^{-1}\boldsymbol{F}_{NT}(\tau)$ can be written as

$$\begin{aligned}\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^2(\tau)M_{it}\boldsymbol{x}_{it}\boldsymbol{x}_{it}'&=\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^2(\tau)p_i(\tau_t)\mathbb{E}\left(\boldsymbol{x}_{it}\boldsymbol{x}_{it}'\right)+o_p(1)\\&=\mu_2\bar{p}(\tau)\left[\boldsymbol{\Sigma}_\chi+\boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)'+\boldsymbol{\Sigma}_\nu\right]+o_p(1),\end{aligned}\tag{B.27}$$

where the o_p-terms are uniform in τ. Note that by (b), (f) in Lemma B.4,

$$\begin{aligned}&\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)'\right\|\\&\qquad\leq\sup_{\tau\in[0,1]}\frac{1}{N}\sum_{i=1}^{N}(Th\nu_{\tau,i})\left\|\frac{1}{Th}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right\|^2=o_p(1).\end{aligned}\tag{B.28}$$

Part (f) is obtained by combining (B.27) and (B.28).

Finally, consider $N^{-1}\boldsymbol{G}_{NT}^k(\tau)$. Note that

$$\begin{aligned}\boldsymbol{G}_{NT}^k(\tau)&=\sum_{i=1}^{N}\boldsymbol{\chi}_i\nu_{\tau,i}\sum_{t=1}^{T}w_t^k(\tau)M_{it}+\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}w_t^k(\tau)M_{it}\boldsymbol{\ell}(\tau_t)+\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}w_t^k(\tau)M_{it}\boldsymbol{\nu}_{it}\\&=:\boldsymbol{G}_{NT,1}^k(\tau)+\boldsymbol{G}_{NT,2}^k(\tau)+\boldsymbol{G}_{NT,3}^k(\tau).\end{aligned}$$

By Assumption A5(b) and Eq. (B.23), when N and T pass to infinity jointly,

$$\frac{1}{N}\boldsymbol{G}_{NT,1}^k(\tau)=\mu_k\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\chi}_i+\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\chi}_i\left(\nu_{\tau,i}\sum_{t=1}^{T}w_t^k(\tau)M_{it}-\mu_k\right)=\mu_k\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\chi}_i+o_p(1)=o_p(1),$$

uniformly in τ. Using similar argument as (B.24) and Lemma B.4(i), we have $N^{-1}\boldsymbol{G}_{NT,2}^k(\tau)=\mu_k\boldsymbol{\ell}(\tau)+O_p(h)\xrightarrow{p}\mu_k\boldsymbol{\ell}(\tau)$ uniformly in τ. Finally, Lemma B.4(g) implies that $N^{-1}\boldsymbol{G}_{NT,3}^k(\tau)=o_p(1)$ uniformly in τ. These results jointly lead to Part (g). ■

C Proofs of main asymptotic results

Proposition C.1 *Under the assumptions in Lemma B.5, we have*

$$\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^{N}\widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{Z}}_i^M(\tau)-\boldsymbol{\Phi}(\tau)\right\|=o_p(1),\qquad as\quad (N,T)\to\infty. \tag{C.1}$$

Proof of Proposition C.1 Recall the construction of $\widetilde{\boldsymbol{Z}}_i^M(\tau)$ in Eq. (2.7). We have

$$\begin{aligned}
&\sum_{i=1}^{N}\widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{Z}}_i^M(\tau)\\
&=\sum_{i=1}^{N}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{K}_h(\tau)\boldsymbol{Z}_i^M(\tau)-\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\boldsymbol{k}_h(\tau)'\left(\boldsymbol{Z}_i^M(\tau)-\bar{\boldsymbol{Z}}^M(\tau)\right)\\
&=\left\{\sum_{i=1}^{N}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{K}_h(\tau)\boldsymbol{Z}_i^M(\tau)-\sum_{i=1}^{N}\nu_{\tau,i}\left(\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\right)\left(\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\right)'\right\}\\
&\quad+\omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\right)'\\
&=\left\{\sum_{i=1}^{N}\sum_{t=1}^{T}K\left(\frac{\tau_t-\tau}{h}\right)M_{it}\boldsymbol{z}_{it}(\tau)\boldsymbol{z}_{it}(\tau)'-\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}K\left(\frac{\tau_t-\tau}{h}\right)M_{it}\boldsymbol{z}_{it}(\tau)\right)\left(\sum_{t=1}^{T}K\left(\frac{\tau_t-\tau}{h}\right)M_{it}\boldsymbol{z}_{it}(\tau)\right)'\right\}\\
&\quad+\omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}K\left(\frac{\tau_t-\tau}{h}\right)M_{it}\boldsymbol{z}_{it}(\tau)\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}K\left(\frac{\tau_t-\tau}{h}\right)M_{it}\boldsymbol{z}_{it}(\tau)\right)'=:\boldsymbol{\Pi}_1+\boldsymbol{\Pi}_2.
\end{aligned}\tag{C.2}$$

Consider $\boldsymbol{\Pi}_1$ first. By the definition of $\boldsymbol{z}_{it}(\tau)$, one can separate $\boldsymbol{\Pi}_1$ into blocks $\boldsymbol{\Pi}_1=(\boldsymbol{\Pi}_{1,ij},\ 1\le i,j\le 4)$, where $\boldsymbol{\Pi}_{1,1j}=\boldsymbol{\Pi}_{1,j1}'=\boldsymbol{0}'$, $\boldsymbol{\Pi}_{1,22}=\boldsymbol{A}_{NT}(\tau)$, $\boldsymbol{\Pi}_{1,23}=\boldsymbol{B}_{NT}(\tau)$, $\boldsymbol{\Pi}_{1,24}=\boldsymbol{C}_{NT}(\tau)$, $\boldsymbol{\Pi}_{1,33}=\boldsymbol{D}_{NT}(\tau)$, $\boldsymbol{\Pi}_{1,34}=\boldsymbol{E}_{NT}(\tau)$, $\boldsymbol{\Pi}_{1,44}=\boldsymbol{F}_{NT}(\tau)$, and $\boldsymbol{\Pi}_{1,ij}=\boldsymbol{\Pi}_{1,ji}'$, $1\le i,j\le 4$, where $\boldsymbol{A}_{NT}(\tau)$ to $\boldsymbol{F}_{NT}(\tau)$ are defined in Lemma B.5. Using Lemma B.5, we obtain

$$\frac{1}{NTh}\boldsymbol{\Pi}_1=\begin{pmatrix}0&&\\&\bar{p}(\tau)\boldsymbol{\Sigma}_\nu&\\&&\boldsymbol{\Phi}_2(\tau)\end{pmatrix}+o_p(1), \tag{C.3}$$

where the o_p-term is uniform in τ. Similarly, one can write $\boldsymbol{\Pi}_2=(\boldsymbol{\Pi}_{2,ij},\ 1\le i,j\le 4)$, where $\boldsymbol{\Pi}_{2,11}=$

$\omega_\tau^{-1}N^2$, $\boldsymbol{\Pi}_{2,12} = \omega_\tau^{-1}N\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^0(\tau)M_{it}\boldsymbol{x}_{it}'$,

$$\boldsymbol{\Pi}_{2,13} = \omega_\tau^{-1}N\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^1(\tau)M_{it}, \qquad \boldsymbol{\Pi}_{2,14} = \omega_\tau^{-1}N\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^1(\tau)M_{it}\boldsymbol{x}_{it}',$$
$$\boldsymbol{\Pi}_{2,22} = \omega_\tau^{-1}\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)',$$
$$\boldsymbol{\Pi}_{2,23} = \omega_\tau^{-1}\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^1(\tau)M_{it}\right),$$
$$\boldsymbol{\Pi}_{2,24} = \omega_\tau^{-1}\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)',$$
$$\boldsymbol{\Pi}_{2,33} = \omega_\tau^{-1}\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^1(\tau)M_{it}\right)^2, \quad \boldsymbol{\Pi}_{2,34} = \omega_\tau^{-1}\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^1(\tau)M_{it}\right)\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)',$$

and $\boldsymbol{\Pi}_{2,44} = \omega_\tau^{-1}\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{i=1}^N \nu_{\tau,i}\sum_{t=1}^T w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)'$, with $\boldsymbol{\Pi}_{2,ij} = \boldsymbol{\Pi}_{2,ji}'$, $1 \leq i,j \leq 4$. By *(h)* and *(i)* in Lemma B.4, Lemma B.5*(g)*, we obtain

$$\frac{1}{NTh}\boldsymbol{\Pi}_2 = [\bar{q}(\tau)]^{-1}\begin{pmatrix} 1 & \boldsymbol{\ell}(\tau)' & \\ \boldsymbol{\ell}(\tau) & \boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)' & \\ & & \mathbf{O}\end{pmatrix} + o_p(1), \tag{C.4}$$

where the o_p-term is uniform in τ. Combining (C.3) and (C.4), we have (C.1). ■

Proposition C.2 *Recall* $\boldsymbol{b}(\tau) = \frac{1}{2}\begin{pmatrix}\mu_2\boldsymbol{\Upsilon}^{(2)}(\tau)\\ \mathbf{0}\end{pmatrix} + o_p(1)$, *where* $\boldsymbol{\Upsilon}(\tau) = (g(\tau), \boldsymbol{\beta}(\tau)')'$. *Let* $\boldsymbol{\Delta}_i^M(\tau) = \mathrm{diag}(\boldsymbol{m}_i)\boldsymbol{b}_i - \boldsymbol{Z}_i^M(\tau)\boldsymbol{\theta}(\tau)$, *where* $\boldsymbol{b}_i = (g_1,\ldots,g_T)' + (\boldsymbol{x}_{i1}'\boldsymbol{\beta}_1,\ldots,\boldsymbol{x}_{iT}'\boldsymbol{\beta}_T)'$. *Define*

$$\boldsymbol{\mathcal{B}}_{NT}(\tau) = \sum_{i=1}^N \boldsymbol{Z}_i^M(\tau)'\boldsymbol{K}_h(\tau)\boldsymbol{\Delta}_i^M(\tau) - \sum_{i=1}^N \nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\boldsymbol{k}_h(\tau)'\left(\boldsymbol{\Delta}_i^M(\tau) - \omega_\tau^{-1}\sum_{i=1}^N \nu_{\tau,i}\boldsymbol{\Delta}_i^M(\tau)\right). \tag{C.5}$$

Under the assumptions in Lemma B.5, we have

$$\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\boldsymbol{\mathcal{B}}_{NT}(\tau) - \mathrm{diag}\left[\boldsymbol{\Phi}_1(\tau), \mathbf{O}\right]h^2\boldsymbol{b}(\tau)\right\| = O_p(h^3). \tag{C.6}$$

Proof of Proposition C.2 By a Taylor expansion of $\boldsymbol{\Upsilon}(\tau_t)$ around $\boldsymbol{\Upsilon}(\tau)$ for $|\tau_t - \tau| \leq h$,

$$\boldsymbol{\Delta}_i^M(\tau) = \frac{h^2}{2}\,\mathrm{diag}\left[\left(\frac{\tau_1-\tau}{h}\right)^2,\ldots,\left(\frac{\tau_T-\tau}{h}\right)^2\right]\boldsymbol{Z}_i^M(\tau)\begin{pmatrix}\boldsymbol{\Upsilon}^{(2)}(\tau)\\ \mathbf{0}_{d+1}\end{pmatrix} + O_p(h^3), \tag{C.7}$$

where the $O_p(h^3)$-term is uniform in τ. In a similar way to (C.2), using (C.7), $(NTh)^{-1}\boldsymbol{\mathcal{B}}_{NT}(\tau)$ can be written as

$$\frac{1}{NTh}\left[\sum_{i=1}^{N}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{K}_h(\tau)\boldsymbol{\varDelta}_i^M(\tau)-\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\boldsymbol{k}_h(\tau)'\boldsymbol{\varDelta}_i^M(\tau)\right.$$
$$\left.+\omega_\tau^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{k}_h(\tau)'\boldsymbol{\varDelta}_i^M(\tau)\right]=\frac{h^2}{2}\left[\frac{1}{NTh}\widetilde{\boldsymbol{\varPi}}_1+\frac{1}{NTh}\widetilde{\boldsymbol{\varPi}}_2\right]\begin{pmatrix}\boldsymbol{\varUpsilon}^{(2)}(\tau)\\ \mathbf{0}_{d+1}\end{pmatrix}+O_p(h^3), \tag{C.8}$$

where the $O_p(h^3)$-term is uniform in τ, and

$$\widetilde{\boldsymbol{\varPi}}_1=\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^2(\tau)M_{it}\boldsymbol{z}_{it}(\tau)\boldsymbol{z}_{it}(\tau)'-\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^0(\tau)M_{it}\boldsymbol{z}_{it}(\tau)\right)\left(\sum_{t=1}^{T}w_t^2(\tau)M_{it}\boldsymbol{z}_{it}(\tau)\right)',$$
$$\widetilde{\boldsymbol{\varPi}}_2=\omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}w_t^0(\tau)M_{it}\boldsymbol{z}_{it}(\tau)\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}w_t^2(\tau)M_{it}\boldsymbol{z}_{it}(\tau)\right)',$$

which are $2(d+1)$-dimensional square matrices. As the block decomposition of $\boldsymbol{\varPi}_1$ in Proof of Proposition C.1, by the definition of $\boldsymbol{z}_{it}(\tau)$, one can separate $\widetilde{\boldsymbol{\varPi}}_1$ into blocks $\widetilde{\boldsymbol{\varPi}}_1=\big(\widetilde{\boldsymbol{\varPi}}_{1,ij},\ 1\le i,j\le 4\big)$, where $\widetilde{\boldsymbol{\varPi}}_{1,ij}$ have the same dimensions of $\boldsymbol{\varPi}_{1,ij}$. The matrix $\widetilde{\boldsymbol{\varPi}}_2$ is similar. By noting the zero vector $\mathbf{0}_{d+1}$ in Eq. (C.8), it is only necessary to consider the first $d+1$ columns of $\widetilde{\boldsymbol{\varPi}}_k$ for $k=1,2$, i.e., $\big(\widetilde{\boldsymbol{\varPi}}_{k,ij},\ 1\le i\le 4,\ 1\le j\le 2\big)$.

For $\widetilde{\boldsymbol{\varPi}}_1$, we obtain $\widetilde{\boldsymbol{\varPi}}_{1,11}=0$, $\widetilde{\boldsymbol{\varPi}}_{1,12}=\mathbf{0}'$,

$$\widetilde{\boldsymbol{\varPi}}_{1,21}=\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^2(\tau)M_{it}\boldsymbol{x}_{it}-\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T}w_t^2(\tau)M_{it}\right),$$
$$\widetilde{\boldsymbol{\varPi}}_{1,22}=\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^2(\tau)M_{it}\boldsymbol{x}_{it}\boldsymbol{x}_{it}'-\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T}w_t^2(\tau)M_{it}\boldsymbol{x}_{it}\right)',$$
$$\widetilde{\boldsymbol{\varPi}}_{1,31}=\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^3(\tau)M_{it}-\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^1(\tau)M_{it}\right)\left(\sum_{t=1}^{T}w_t^2(\tau)M_{it}\right),$$
$$\widetilde{\boldsymbol{\varPi}}_{1,32}=\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^3(\tau)M_{it}\boldsymbol{x}_{it}'-\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^1(\tau)M_{it}\right)\left(\sum_{t=1}^{T}w_t^2(\tau)M_{it}\boldsymbol{x}_{it}\right)',$$
$$\widetilde{\boldsymbol{\varPi}}_{1,41}=\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^3(\tau)M_{it}\boldsymbol{x}_{it}-\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T}w_t^2(\tau)M_{it}\right),$$

and $\widetilde{\boldsymbol{\varPi}}_{1,42}=\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^3(\tau)M_{it}\boldsymbol{x}_{it}\boldsymbol{x}_{it}'-\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T}w_t^2(\tau)M_{it}\boldsymbol{x}_{it}\right)'$. By a straightforward modification of Lemma B.5, using Lemma B.4, we find

$$\frac{1}{NTh}\begin{pmatrix}\widetilde{\boldsymbol{\varPi}}_{1,11} & \widetilde{\boldsymbol{\varPi}}_{1,12}\\ \widetilde{\boldsymbol{\varPi}}_{1,21} & \widetilde{\boldsymbol{\varPi}}_{1,22}\\ \widetilde{\boldsymbol{\varPi}}_{1,31} & \widetilde{\boldsymbol{\varPi}}_{1,32}\\ \widetilde{\boldsymbol{\varPi}}_{1,41} & \widetilde{\boldsymbol{\varPi}}_{1,42}\end{pmatrix}=\mu_2\bar{p}(\tau)\begin{pmatrix}0 & \mathbf{0}'\\ \mathbf{0} & \boldsymbol{\varSigma}_\nu\\ 0 & \mathbf{0}'\\ \mathbf{0} & \mathbf{O}\end{pmatrix}+o_p(1), \tag{C.9}$$

where the o_p-term is uniformly in τ. For $\widetilde{\boldsymbol{\Pi}}_2$, we obtain $\widetilde{\boldsymbol{\Pi}}_{2,11} = \omega_\tau^{-1} N \sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^2(\tau) M_{it}$, $\widetilde{\boldsymbol{\Pi}}_{2,12} = \omega_\tau^{-1} N \sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^2(\tau) M_{it} \boldsymbol{x}_{it}'$,

$$
\begin{aligned}
\widetilde{\boldsymbol{\Pi}}_{2,21} &= \omega_\tau^{-1} \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^0(\tau) M_{it} \boldsymbol{x}_{it} \right) \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^2(\tau) M_{it} \right), \\
\widetilde{\boldsymbol{\Pi}}_{2,22} &= \omega_\tau^{-1} \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^0(\tau) M_{it} \boldsymbol{x}_{it} \right) \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^2(\tau) M_{it} \boldsymbol{x}_{it} \right)', \\
\widetilde{\boldsymbol{\Pi}}_{2,31} &= \omega_\tau^{-1} \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^1(\tau) M_{it} \right) \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^2(\tau) M_{it} \right), \\
\widetilde{\boldsymbol{\Pi}}_{2,32} &= \omega_\tau^{-1} \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^1(\tau) M_{it} \right) \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^2(\tau) M_{it} \boldsymbol{x}_{it} \right)', \\
\widetilde{\boldsymbol{\Pi}}_{2,41} &= \omega_\tau^{-1} \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^1(\tau) M_{it} \boldsymbol{x}_{it} \right) \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^2(\tau) M_{it} \right), \\
\widetilde{\boldsymbol{\Pi}}_{2,42} &= \omega_\tau^{-1} \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^1(\tau) M_{it} \boldsymbol{x}_{it} \right) \left(\sum_{i=1}^N \nu_{\tau,i} \sum_{t=1}^T w_t^2(\tau) M_{it} \boldsymbol{x}_{it} \right)'.
\end{aligned}
$$

Similarly,

$$
\frac{1}{NTh} \begin{pmatrix} \widetilde{\boldsymbol{\Pi}}_{2,11} & \widetilde{\boldsymbol{\Pi}}_{2,12} \\ \widetilde{\boldsymbol{\Pi}}_{2,21} & \widetilde{\boldsymbol{\Pi}}_{2,22} \\ \widetilde{\boldsymbol{\Pi}}_{2,31} & \widetilde{\boldsymbol{\Pi}}_{2,32} \\ \widetilde{\boldsymbol{\Pi}}_{2,41} & \widetilde{\boldsymbol{\Pi}}_{2,42} \end{pmatrix} = \mu_2 \frac{1}{\bar{q}(\tau)} \begin{pmatrix} 1 & \boldsymbol{\ell}(\tau)' \\ \boldsymbol{\ell}(\tau) & \boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)' \\ 0 & \mathbf{0}' \\ \mathbf{0} & \mathbf{O} \end{pmatrix} + o_p(1), \tag{C.10}
$$

where the o_p-term is uniformly in τ. Therefore, by disregarding the last $d+1$ columns, the initial $d+1$ columns of $(NTh)^{-1}(\widetilde{\boldsymbol{\Pi}}_1 + \widetilde{\boldsymbol{\Pi}}_2)$ can be written as $\begin{pmatrix} \mu_2 \boldsymbol{\Phi}_1(\tau) \\ \mathbf{O} \end{pmatrix} + o_p(1)$ by combining (C.8) and (C.9). Plugging it into (C.10), we obtain Eq. (C.6). ■

Lemma C.1 *Under Assumptions A1, A3, A4, A5(c), A6, A7, and A8, for any fixed $\tau \in (0,1)$, $k = 0, 1, \ldots$,*

$$
\frac{1}{\sqrt{NTh}} \sum_{i=1}^N \sum_{t=1}^T w_t^k(\tau) M_{it} \begin{pmatrix} Th\nu_{\tau,i} \\ 1 \\ \boldsymbol{\chi}_i \\ \boldsymbol{\nu}_{it} \end{pmatrix} e_{it} \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \nu_{2k} \begin{pmatrix} \boldsymbol{\Lambda}_\varepsilon(\tau) & & \\ & \bar{\Lambda}_\varepsilon(\tau)\boldsymbol{\Sigma}_\chi & \\ & & \boldsymbol{\Lambda}_{\nu\varepsilon}(\tau) \end{pmatrix} \right), \tag{C.11}
$$

as $(N,T) \to \infty$, where $\boldsymbol{\Lambda}_\varepsilon(\tau)$, $\bar{\Lambda}_\varepsilon(\tau)$, and $\boldsymbol{\Lambda}_{\nu\varepsilon}(\tau)$, are defined in Assumption A7(d).

Proof of Lemma C.1 We use the Cramér-Wold device to establish the asymptotic joint distribution. For any unit vector $\boldsymbol{a} = (a_1, \boldsymbol{a}_2')' \in \mathbb{R}^{2(d+1)}$, it suffices to consider

$$\sum_{t=1}^{T} \mathcal{Y}_{NT,t}(\tau) + a_1 \frac{1}{\sqrt{NTh}} \sum_{i=1}^{N} \sum_{t=1}^{T} w_t^k(\tau) M_{it} \left\{ (Th\nu_{\tau,i}) - [p_i(\tau)]^{-1} \right\} e_{it}, \tag{C.12}$$

where $\mathcal{Y}_{NT,t}(\tau) = (NTh)^{-1/2} w_t^k(\tau) \sum_{i=1}^{N} M_{it} \boldsymbol{a}' [1/p_i(\tau), 1, \boldsymbol{\chi}_i', \boldsymbol{\nu}_{it}']' e_{it}$. We shall argue that, as $(N,T) \to \infty$,

(a) $\sum_{t=1}^{T} \mathcal{Y}_{NT,t}(\tau) \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \boldsymbol{a}'\nu_{2k} \operatorname{diag}\left[\boldsymbol{\Lambda}_\varepsilon(\tau), \bar{\Lambda}_\varepsilon(\tau)\boldsymbol{\Sigma}_\chi, \boldsymbol{\Lambda}_{\nu\varepsilon}(\tau)\right]\boldsymbol{a}\right)$ for any $\tau \in (0,1)$;

(b) $(NTh)^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} w_t^k(\tau) M_{it} \left\{ (Th\nu_{\tau,i}) - [p_i(\tau)]^{-1} \right\} e_{it} = o_p(1)$ for any $\tau \in (0,1)$.

We prove Part (a) in two steps: (i) derive the asymptotic variance; (ii) show the asymptotic normality. For the asymptotic variance, we have $\mathbb{E}\left(\sum_{t=1}^{T} \mathcal{Y}_{NT,t}(\tau)\right) = 0$ by Assumption A7(b). Moreover, by the identities $e_{it} = \sigma_i(\tau_t)\varepsilon_{it}$ and

$$\left(\sum_{i=1}^{N}\sum_{t=1}^{T} a_{it}\right)^2 = \sum_{i,j=1}^{N}\left[\sum_{t=1}^{T} a_{it}a_{jt} + \sum_{s=1}^{T-1}\sum_{t=1}^{T-s}\left(a_{it}a_{j(t+s)} + a_{i(t+s)}a_{jt}\right)\right], \tag{C.13}$$

we have $\mathbb{E}\left(\sum_{t=1}^{T} \mathcal{Y}_{NT,t}(\tau)\right)^2 = \boldsymbol{a}'\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT}(\tau)\right]\boldsymbol{a}$, where $\boldsymbol{\mathcal{E}}_{NT}(\tau) = \left(\boldsymbol{\mathcal{E}}_{NT,k_1k_2}(\tau),\ k_1, k_2 = 1,2,3,4\right)$ is a block matrix with the elements given by

$$\begin{aligned}
\boldsymbol{\mathcal{E}}_{NT,k_1k_2}(\tau) = \frac{1}{NTh}\sum_{i,j=1}^{N}\Bigg\{ & \sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2 \sigma_i(\tau_t)\sigma_j(\tau_t) M_{it}M_{jt}\boldsymbol{A}_{k_1k_2}(i,t;j,t)\varepsilon_{it}\varepsilon_{jt} \\
& + \sum_{s=1}^{T-1}\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\,\sigma_i(\tau_t)\sigma_j(\tau_{t+s})M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)\varepsilon_{it}\varepsilon_{j(t+s)} \\
& + \sum_{s=1}^{T-1}\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\,\sigma_i(\tau_{t+s})\sigma_j(\tau_t)M_{i(t+s)}M_{jt}\boldsymbol{A}_{k_1k_2}(i,t+s;j,t)\varepsilon_{i(t+s)}\varepsilon_{jt}\Bigg\},
\end{aligned}$$

with $\boldsymbol{A}_{k_1k_2}(i,t_i;j,t_j) = \boldsymbol{a}_{k_1}(i,t_i)\boldsymbol{a}_{k_2}(j,t_j)'$. Moreover, for $\ell = i$ or $\ell = j$, we define $\boldsymbol{a}_1(\ell,t_\ell) = 1/p_\ell(\tau)$, $\boldsymbol{a}_2(\ell,t_\ell) = 1$, $\boldsymbol{a}_3(\ell,t_\ell) = \boldsymbol{\chi}_\ell$, $\boldsymbol{a}_4(\ell,t_\ell) = \boldsymbol{\nu}_{\ell t_\ell}$.

Before deriving $\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,k_1k_2}(\tau)\right]$, $k_1, k_2 = 1,\dots,4$, some intermediate results are given as follows. Recall the definition $\mathcal{I}_{h,s}(\tau) = \{t : |\tau_{t+s} - \tau| \le h\}$, $s \ge 0$, and $\mathcal{I}_h(\tau) = \mathcal{I}_{h,s}(\tau)$ for $s = 0$.

(a.1) $\forall \tau \in (0,1)$, $w_t^k(\tau)\, w_{t+s}^k(\tau) \equiv 0$ if $t \notin \mathcal{I}_h(\tau) \bigcap \mathcal{I}_{h,s}(\tau) = \{\lceil T(\tau - h)\rceil \le t \le \lfloor T(\tau+h) - s\rfloor, 0 \le s \le 2Th\}$. Therefore, for $s > 2Th$, $(Th)^{-1}\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau) = 0$, and

$$\frac{1}{Th}\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau) = \frac{1}{Th}\sum_{t=\lceil T(\tau-h)\rceil}^{\lfloor T(\tau+h)-s\rfloor} w_t^k(\tau)\, w_{t+s}^k(\tau), \qquad 0 \le s \le 2Th. \tag{C.14}$$

Furthermore, let $\{H_T\} \subset (0, \lfloor 2Th \rfloor)$ be some sequence that satisfies $\dfrac{1}{H_T} + \dfrac{H_T}{Th} \to 0$ as $T \to \infty$, then for $s < H_T$,

$$\frac{1}{Th}\sum_{t=\lceil T(\tau-h)\rceil}^{\lfloor T(\tau+h)-s\rfloor} w_t^k(\tau)\, w_{t+s}^k(\tau) = \frac{1}{Th}\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2 + O\left(\frac{H_T}{Th}\right) \to \nu_{2k}, \qquad T\to\infty, \tag{C.15}$$

uniformly in τ, due to $(Th)^{-1}\sum_{t=\lfloor T(\tau+h)-s\rfloor+1}^{\lfloor T(\tau+h)\rfloor}\left[w_t^k(\tau)\right]^2 \leq CH_T/(Th)$, and

$$\left|\frac{1}{Th}\sum_{t=\lceil T(\tau-h)\rceil}^{\lfloor T(\tau+h)-s\rfloor} w_t^k(\tau)\left[w_{t+s}^k(\tau) - w_t^k(\tau)\right]\right| \leq C\frac{H_T}{Th},$$

by the Lipschitz continuity of $x\mapsto x^k K(x)$ over $[-1,1]$.

$(a.2)$ For $t\in\mathcal{I}_h(\tau)\bigcap\mathcal{I}_{h,s}(\tau)$, we have $|\sigma_i(\tau_{t+s})-\sigma_i(\tau)|\leq Ch$ and $|\mathcal{R}_{i,j}(\tau_t,\tau_{t+s})-\mathcal{R}_{i,j}(\tau,\tau)|\leq Ch$ by the Lipschitz continuity (Assumptions A4, A6(b)), where $\tau\in(0,1)$, $s\geq 0$, $i,j\in\{1\ldots,N\}$.

Now we derive $\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,k_1k_2}(\tau)\right]$, $k_1,k_2=1,\ldots,4$. Let H_T be the sequence as given in $(a.1)$ above. Note that $\{\varepsilon_{it}\varepsilon_{j(t+s)}\}$, $s\in\mathbb{Z}$, are strictly stationary by Assumption A3 and Theorem 3.35 in White (2001), i.e., $\mathbb{E}\left(\varepsilon_{it}\varepsilon_{j(t+s)}\right)=\mathbb{E}\left(\varepsilon_{i0}\varepsilon_{js}\right)$. Using $(a.1)$ - $(a.2)$, Assumptions A3, A4, A6(b), A7(b) - (c), we have

$$\begin{aligned}
&\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,11}(\tau)\right]\\
&=\frac{1}{N}\sum_{i,j=1}^{N}\left[p_i(\tau)p_j(\tau)\right]^{-1}\Bigg\{\mathbb{E}\left(\varepsilon_{it}\varepsilon_{jt}\right)\frac{1}{Th}\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2\sigma_i(\tau_t)\sigma_j(\tau_t)\mathcal{R}_{i,j}(\tau_t,\tau_t)\\
&\quad+\sum_{s=1}^{T-1}\mathbb{E}\left(\varepsilon_{it}\varepsilon_{j(t+s)}\right)\frac{1}{Th}\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\,\sigma_i(\tau_t)\sigma_j(\tau_{t+s})\mathcal{R}_{i,j}(\tau_t,\tau_{t+s})\\
&\quad+\sum_{s=1}^{T-1}\mathbb{E}\left(\varepsilon_{i(t+s)}\varepsilon_{jt}\right)\frac{1}{Th}\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\,\sigma_i(\tau_{t+s})\sigma_j(\tau_t)\mathcal{R}_{i,j}(\tau_{t+s},\tau_t)\Bigg\}\\
&=\frac{1}{N}\sum_{i,j=1}^{N}\frac{\mathcal{R}_{i,j}(\tau,\tau)}{p_i(\tau)p_j(\tau)}\sigma_i(\tau)\sigma_j(\tau)\Bigg\{\mathbb{E}\left(\varepsilon_{it}\varepsilon_{jt}\right)\frac{1}{Th}\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2\\
&\quad+\sum_{s=1}^{T-1}\mathbb{E}\left(\varepsilon_{it}\varepsilon_{j(t+s)}\right)\frac{1}{Th}\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)+\sum_{s=1}^{T-1}\mathbb{E}\left(\varepsilon_{i(t+s)}\varepsilon_{jt}\right)\frac{1}{Th}\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\Bigg\}+O(h)\\
&=\frac{1}{N}\sum_{i,j=1}^{N}\frac{\mathcal{R}_{i,j}(\tau,\tau)}{p_i(\tau)p_j(\tau)}\sigma_i(\tau)\sigma_j(\tau)\Bigg\{\mathbb{E}\left(\varepsilon_{it}\varepsilon_{jt}\right)\frac{1}{Th}\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2\\
&\quad+\sum_{s=1}^{H_T-1}\mathbb{E}\left(\varepsilon_{it}\varepsilon_{j(t+s)}\right)\frac{1}{Th}\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2+\sum_{s=H_T}^{\lfloor 2Th\rfloor}\mathbb{E}\left(\varepsilon_{it}\varepsilon_{j(t+s)}\right)\frac{1}{Th}\sum_{t=\lceil T(\tau-h)\rceil}^{\lfloor T(\tau+h)-s\rfloor}w_t^k(\tau)\,w_{t+s}^k(\tau)+\\
&\quad+\sum_{s=1}^{H_T-1}\mathbb{E}\left(\varepsilon_{i(t+s)}\varepsilon_{jt}\right)\frac{1}{Th}\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2+\sum_{s=H_T}^{\lfloor 2Th\rfloor}\mathbb{E}\left(\varepsilon_{i(t+s)}\varepsilon_{jt}\right)\frac{1}{Th}\sum_{t=\lceil T(\tau-h)\rceil}^{\lfloor T(\tau+h)-s\rfloor}w_t^k(\tau)\,w_{t+s}^k(\tau)\Bigg\}+O\left(\frac{H_T}{Th}\right)+O(h)\\
&=\left\{\frac{1}{N}\sum_{i,j=1}^{N}\frac{\mathcal{R}_{i,j}(\tau,\tau)}{p_i(\tau)p_j(\tau)}\sigma_i(\tau)\sigma_j(\tau)\sum_{s=-H_T+1}^{H_T-1}\mathbb{E}\left(\varepsilon_{it}\varepsilon_{j(t+s)}\right)\right\}\left\{\frac{1}{Th}\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2\right\}+O\left(\frac{H_T}{Th}\right)+o(1)+O(h),
\end{aligned}$$

where the $o(1)$-term in the final equality is obtained by using A7(c),

$$\left|\frac{1}{N}\sum_{i,j=1}^{N}\frac{\mathcal{R}_{i,j}(\tau,\tau)}{p_i(\tau)p_j(\tau)}\sigma_i(\tau)\sigma_j(\tau)\sum_{s=H_T}^{\lfloor 2Th\rfloor}\left\{\mathbb{E}\left(\varepsilon_{it}\varepsilon_{j(t+s)}\right)+\mathbb{E}\left(\varepsilon_{i(t+s)}\varepsilon_{jt}\right)\right\}\frac{1}{Th}\sum_{t=\lceil T(\tau-h)\rceil}^{\lfloor T(\tau+h)-s\rfloor}w_t^k(\tau)\,w_{t+s}^k(\tau)\right|$$
$$\leq C\frac{1}{N}\sum_{i,j=1}^{N}\sum_{s=H_T}^{T}\left\{\left|\mathbb{E}\left(\varepsilon_{it}\varepsilon_{j(t+s)}\right)\right|+\left|\mathbb{E}\left(\varepsilon_{i(t+s)}\varepsilon_{jt}\right)\right|\right\}=o(1),$$

as $(N,T)\to\infty$. By Assumption A7(d), we obtain

$$\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,11}(\tau)\right]\to\nu_{2k}\Lambda_{\varepsilon,11}(\tau),\qquad (N,T)\to\infty, \tag{C.16}$$

where $\Lambda_{\varepsilon,,k_1k_2}(\tau)$ is the $(k_1,k_2)_{th}$ element of $\boldsymbol{\Lambda}_\varepsilon(\tau)$. Similarly, we obtain $\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,k_1k_2}(\tau)\right]\to\nu_{2k}\Lambda_{\varepsilon,k_1k_2}(\tau)$, $1\leq k_1,k_2,\leq 2$. Moreover, by Assumption A7(b), one can obtain $\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,13}(\tau)'\right]=\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,31}(\tau)\right]=\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,14}(\tau)'\right]=\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,41}(\tau)\right]=\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,23}(\tau)'\right]=\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,32}(\tau)\right]=\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,24}(\tau)'\right]=\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,42}(\tau)\right]=\mathbf{0}$, and $\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,34}(\tau)\right]=\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,43}(\tau)\right]=\mathbf{O}$. Note that $\boldsymbol{\chi}_i$, $i=1,\ldots,N$, are independent, by steps similar to $\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,11}(\tau)\right]$, we have

$$\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,33}(\tau)\right]=\boldsymbol{\Sigma}_\chi\left\{\frac{1}{N}\sum_{i=1}^{N}\mathcal{R}_{i,i}(\tau,\tau)\sigma_i^2(\tau)\sum_{s=-H_T+1}^{H_T-1}\mathbb{E}\left(\varepsilon_{it}\varepsilon_{i(t+s)}\right)\right\}$$
$$\times\left\{\frac{1}{Th}\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2\right\}+O\left(\frac{H_T}{Th}\right)+o(1)+O(h)\to\nu_{2k}\bar{\Lambda}_\varepsilon(\tau)\boldsymbol{\Sigma}_\chi, \tag{C.17}$$

as $(N,T)\to\infty$, and

$$\mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT,44}(\tau)\right]=\left\{\frac{1}{N}\sum_{i,j=1}^{N}\mathcal{R}_{i,j}(\tau,\tau)\sigma_i(\tau)\sigma_j(\tau)\sum_{s=-H_T+1}^{H_T-1}\mathbb{E}\left(\boldsymbol{\nu}_{it}\boldsymbol{\nu}_{j(t+s)}'\varepsilon_{it}\varepsilon_{j(t+s)}\right)\right\}$$
$$\times\left\{\frac{1}{Th}\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2\right\}+O\left(\frac{H_T}{Th}\right)+o(1)+O(h)\to\nu_{2k}\boldsymbol{\Lambda}_{\nu\varepsilon}(\tau). \tag{C.18}$$

Combining the results above, we have

$$\mathbb{E}\left(\sum_{t=1}^{T}\mathcal{Y}_{NT,t}(\tau)\right)^2\to\boldsymbol{a}'\nu_{2k}\begin{pmatrix}\boldsymbol{\Lambda}_\varepsilon(\tau) & & \\ & \bar{\Lambda}_\varepsilon(\tau)\boldsymbol{\Sigma}_\chi & \\ & & \boldsymbol{\Lambda}_{\nu\varepsilon}(\tau)\end{pmatrix}\boldsymbol{a}. \tag{C.19}$$

For Part (a), it remains to establish the asymptotic normality. We employ the central limit theorem (CLT) for NED processes given in Corollary 24.7 of Davidson (1994), as in Gao et al. (2020, Proof of Eq. (B.25)). As such, we verify the conditions required in the corollary.

$(a.3)$ $\mathbb{E}\left(\mathcal{Y}_{NT,t}(\tau)\right)=0$, and by Eq. (C.19), $\mathbb{E}\left(\sum_{t=1}^{T}\mathcal{Y}_{NT,t}(\tau)\right)^2\to\boldsymbol{a}'\nu_{2k}\,\mathrm{diag}\left[\boldsymbol{\Lambda}_\varepsilon(\tau),\bar{\Lambda}_\varepsilon(\tau)\boldsymbol{\Sigma}_\chi,\boldsymbol{\Lambda}_{\nu\varepsilon}(\tau)\right]\boldsymbol{a}$.

$(a.4)$ Let $c_{T,t} = 1/\sqrt{Th}$, $t \geq 1$, and $r = p_0/2 = 2 + \delta > 2$. By Assumption A7$(a)$, we have

$$\begin{aligned}
\left\{\mathbb{E}\left|\mathcal{Y}_{NT,t}(\tau)/c_{T,t}\right|^r\right\}^{1/r} &= \frac{1}{\sqrt{N}} w_t^k(\tau) \left\{\mathbb{E}\left|a_1 \sum_{i=1}^N \frac{\sigma_i(\tau_t)}{p_i(\tau)} M_{it}\varepsilon_{it} + \boldsymbol{a}_2' \sum_{i=1}^N \sigma_i(\tau_t) M_{it}\boldsymbol{\nu}_{it}\varepsilon_{it}\right|^r\right\}^{1/r} \\
&\leq C\frac{1}{\sqrt{N}} \left\{\left(\mathbb{E}\left|\sum_{i=1}^N \frac{\sigma_i(\tau_t)}{p_i(\tau)} M_{it}\varepsilon_{it}\right|^r\right)^{1/r} + \left(\mathbb{E}\left\|\sum_{i=1}^N \sigma_i(\tau_t) M_{it}\boldsymbol{\nu}_{it}\varepsilon_{it}\right\|^r\right)^{1/r}\right\} \\
&\leq C.
\end{aligned}$$

$(a.5)$ For any $\tau \in (0,1)$, $\mathcal{Y}_{NT,t}(\tau)$ is L_2-NED of size -1 on $\{\boldsymbol{\xi}_{\bullet t}\}$, where $\{\boldsymbol{\xi}_{\bullet t}\}$ is α-mixing of size $-(2+\delta)/\delta$ by Assumption A3, in the terminology of Davidson (1994). More specifically, recall $\mathcal{F}_{\xi,t_1}^{t_2} = \sigma\left(\boldsymbol{\xi}_{\bullet t},\, t_1 \leq t \leq t_2\right)$. By Minkowski's and Cauchy-Schwarz inequalities, Assumptions A4, A5(c), we have

$$\begin{aligned}
&\left\{\mathbb{E}\left|\mathcal{Y}_{NT,t}(\tau) - \mathbb{E}\left(\mathcal{Y}_{NT,t}(\tau) \,\middle|\, \mathcal{F}_{\xi,t-m}^{t+m}\right)\right|^2\right\}^{1/2} \\
&= \frac{1}{\sqrt{NTh}} w_t^k(\tau) \left\{\mathbb{E}\left|\sum_{i=1}^N \sigma_i(\tau_t) M_{it}\boldsymbol{a}_2' \left[\boldsymbol{\nu}_{it} - \mathbb{E}\left(\boldsymbol{\nu}_{it} \,\middle|\, \mathcal{F}_{\xi,t-m}^{t+m}\right)\right]\varepsilon_{it}\right|^2\right\}^{1/2} \\
&\leq C\frac{1}{\sqrt{NTh}} \sum_{i=1}^N \left\{\mathbb{E}\left\|\boldsymbol{\nu}_{it} - \mathbb{E}\left(\boldsymbol{\nu}_{it} \,\middle|\, \mathcal{F}_{\xi,t-m}^{t+m}\right)\right\|^4\right\}^{1/4} \left[\mathbb{E}(\varepsilon_{it}^4)\right]^{1/4} \qquad (C.20) \\
&\leq C\frac{1}{\sqrt{NTh}} \sum_{i=1}^N \left\{\mathbb{E}\left\|\boldsymbol{\nu}_{it} - \boldsymbol{\nu}_{it}^{(m)}\right\|^{p_0}\right\}^{1/p_0} \leq C\frac{1}{\sqrt{NTh}} \sum_{i=1}^N \left[\psi_{i,p_0}(m)\right]^{1/p_0} \leq C\frac{1}{\sqrt{Th}} m^{-\varphi_\nu/p_0},
\end{aligned}$$

where (C.20) follows from Theorem 10.28 in Davidson (1994) and Lyapunov's inequality. Since $\varphi_\nu/p_0 > 1$, we can conclude that $\mathcal{Y}_{NT,t}(\tau)$ is L_2-NED of size -1.

$(a.6)$ Note that $\sum_{t=1}^T \mathcal{Y}_{NT,t}(\tau) = \sum_{t \in \mathcal{I}_h(\tau)} \mathcal{Y}_{NT,t}(\tau)$, where the cardinality of $\mathcal{I}_h(\tau)$ is bounded by CTh. Eq. (24.28) in Davidson (1994) can be replaced by $\sup_T (Th) \max_{t\geq 1} c_{T,t}^2 < \infty$.

The conditions in Corollary 24.7 of Davidson (1994) are fulfilled by $(a.3)$ - $(a.6)$. Therefore, Part (a) is obtained.

We consider Part (b) next. Note that by Assumptions A7(b) - (c) and A8,

$$
\begin{aligned}
&\mathbb{E}\left|\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^k(\tau)\,M_{it}\left\{(Th\nu_{\tau,i})-[p_i(\tau)]^{-1}\right\}e_{it}\right|^2\\
&=\frac{1}{NTh}\mathbb{E}\left\{\sum_{i,j=1}^{N}\left\{(Th\nu_{\tau,i})-[p_i(\tau)]^{-1}\right\}\left\{(Th\nu_{\tau,j})-[p_j(\tau)]^{-1}\right\}\sum_{s,t=1}^{T}w_t^k(\tau)\,w_s^k(\tau)\,\sigma_i(\tau_t)\sigma_j(\tau_s)M_{it}M_{js}\varepsilon_{it}\varepsilon_{js}\right\}\\
&=\frac{1}{NTh}\mathbb{E}\left\{\sum_{i,j=1}^{N}\left\{(Th\nu_{\tau,i})-[p_i(\tau)]^{-1}\right\}\left\{(Th\nu_{\tau,j})-[p_j(\tau)]^{-1}\right\}\sum_{s,t=1}^{T}w_t^k(\tau)\,w_s^k(\tau)\,\sigma_i(\tau_t)\sigma_j(\tau_s)M_{it}M_{js}\mathbb{E}\left(\varepsilon_{it}\varepsilon_{js}\right)\right\}\\
&\leq C\left(\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}+h^4+\frac{1}{T^2h^4}\right)\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s,t=1}^{T}w_t^k(\tau)\,w_s^k(\tau)\left|\mathbb{E}\left(\varepsilon_{it}\varepsilon_{js}\right)\right| \qquad (C.21)\\
&\leq C\left(\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}+h^4+\frac{1}{T^2h^4}\right)\frac{1}{N}\sum_{i,j=1}^{N}\left\{\left|\mathbb{E}\left(\varepsilon_{i0}\varepsilon_{j0}\right)\right|+\sum_{k=1}^{T-1}\left(\left|\mathbb{E}\left(\varepsilon_{i0}\varepsilon_{jk}\right)\right|+\left|\mathbb{E}\left(\varepsilon_{ik}\varepsilon_{j0}\right)\right|\right)\right\}\\
&=o(1),\qquad \text{uniformly in } \tau, \qquad (C.22)
\end{aligned}
$$

where (C.21) is obtained using Lemma B.4(a) - (b) as follows:

$$
\begin{aligned}
\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left|(Th\nu_{\tau,i})-[p_i(\tau)]^{-1}\right| &= \max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left|(Th\nu_{\tau,i})\,[p_i(\tau)]^{-1}\left\{\frac{1}{Th}\sum_{t=1}^{T}w_t^0(\tau)M_{it}-p_i(\tau)\right\}\right|\\
&\leq p_L\left\{\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}(Th\nu_{\tau,i})\right\}\left\{\max_{1\leq i\leq N}\sup_{\tau\in[0,1]}\left|\frac{1}{Th}\sum_{t=1}^{T}w_t^0(\tau)M_{it}-p_i(\tau)\right|\right\}\\
&\leq C\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}+h^2+\frac{1}{Th^2}\right).
\end{aligned}
$$

Then Part (b) is obtained by Markov's inequality. ■

Proposition C.3 *Define*

$$
\boldsymbol{\mathcal{D}}_{NT}(\tau)=\sum_{i=1}^{N}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{K}_h(\tau)\boldsymbol{e}_i^M-\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\boldsymbol{k}_h(\tau)'\left(\boldsymbol{e}_i^M-\omega_\tau^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{e}_i^M\right). \qquad (C.23)
$$

Under Assumptions A1, A3, A5 - A8, for any fixed $\tau\in(0,1)$,

$$
\frac{1}{\sqrt{NTh}}\boldsymbol{\mathcal{D}}_{NT}(\tau)\xrightarrow{d}\mathcal{N}\Big(\boldsymbol{0},\boldsymbol{\Phi}_{\nu\varepsilon}(\tau)\Big), \qquad (C.24)
$$

as $(N,T)\to\infty$, *where* $\boldsymbol{\Phi}_{\nu\varepsilon}(\tau)$ *is defined above Theorem 2 (p. 16).*

Proof of Proposition C.3 We first split $\boldsymbol{\mathcal{D}}_{NT}(\tau)$ into four main blocks of vectors:

$$\begin{aligned}\boldsymbol{\mathcal{D}}_{NT}(\tau) &= \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau) M_{it} \boldsymbol{z}_{it}(\tau) e_{it} - \sum_{i=1}^{N} \nu_{\tau,i} \left(\sum_{t=1}^{T} w_t^0(\tau) M_{it} \boldsymbol{z}_{it}(\tau)\right)\left(\sum_{t=1}^{T} w_t^0(\tau) M_{it} e_{it}\right) \\ &+ \omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau) M_{it}\boldsymbol{z}_{it}(\tau)\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau) M_{it} e_{it}\right) =: \begin{pmatrix}\boldsymbol{\Upsilon}_{NT,1}(\tau)\\ \vdots \\ \boldsymbol{\Upsilon}_{NT,4}(\tau)\end{pmatrix}, \quad \text{(C.25)}\end{aligned}$$

where,

$$\begin{aligned}\boldsymbol{\Upsilon}_{NT,1}(\tau) &= \omega_\tau^{-1} N\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau) M_{it} e_{it}\right),\\ \boldsymbol{\Upsilon}_{NT,2}(\tau) &= \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau) M_{it}\boldsymbol{x}_{it} e_{it} - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau) M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T} w_t^0(\tau) M_{it} e_{it}\right)\\ &\quad + \omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau) M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau) M_{it} e_{it}\right),\\ \boldsymbol{\Upsilon}_{NT,3}(\tau) &= \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau) M_{it} e_{it} - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^1(\tau) M_{it}\right)\left(\sum_{t=1}^{T} w_t^0(\tau) M_{it} e_{it}\right)\\ &\quad + \omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^1(\tau) M_{it}\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau) M_{it} e_{it}\right),\\ \boldsymbol{\Upsilon}_{NT,4}(\tau) &= \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau) M_{it}\boldsymbol{x}_{it} e_{it} - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^1(\tau) M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T} w_t^0(\tau) M_{it} e_{it}\right)\\ &\quad + \omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^1(\tau) M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau) M_{it} e_{it}\right).\end{aligned}$$

We first make the following claims.

(CL.1) $(NTh)^{-1/2}\Big(\boldsymbol{\Upsilon}_{NT,1}(\tau), \boldsymbol{\Upsilon}_{NT,2}(\tau)'\Big)' \xrightarrow{d} \mathcal{N}\Big(\mathbf{0}, \boldsymbol{\Phi}_{\nu\varepsilon,1}(\tau)\Big)$;

(CL.2) $(NTh)^{-1/2}\Big(\boldsymbol{\Upsilon}_{NT,3}(\tau), \boldsymbol{\Upsilon}_{NT,4}(\tau)'\Big)' \xrightarrow{d} \mathcal{N}\Big(\mathbf{0}, \boldsymbol{\Phi}_{\nu\varepsilon,2}(\tau)\Big)$;

(CL.3) $(NTh)^{-1/2}\Big(\boldsymbol{\Upsilon}_{NT,1}(\tau), \boldsymbol{\Upsilon}_{NT,2}(\tau)'\Big)'$ and $(NTh)^{-1/2}\Big(\boldsymbol{\Upsilon}_{NT,3}(\tau), \boldsymbol{\Upsilon}_{NT,4}(\tau)'\Big)'$ are asymptotically independent.

These claims jointly imply the limiting distribution (C.24). We next prove these claims.

I. *Proof of Claim (CL.1)*

We prove the claim in two steps. First, we establish the asymptotic marginal distributions of $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,1}(\tau)$ and $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,2}(\tau)$, respectively. When showing the asymptotic distribution of $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,2}(\tau)$, we argue that $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,2}(\tau)$ is asymptotically dominated by

$$\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}e_{it}+\left(\frac{Th}{N}\omega_\tau\right)^{-1}\boldsymbol{\ell}(\tau)\left(\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}\left(Th\nu_{\tau,i}\right)e_{it}\right). \tag{C.26}$$

If (C.26) holds, then the joint distributional convergence of $\left(\boldsymbol{\Upsilon}_{NT,1}(\tau),\boldsymbol{\Upsilon}_{NT,2}(\tau)'\right)'$ is immediate by utilizing Lemma C.1. Given the joint convergence, as the second step, we only need to establish the asymptotic covariance matrix of $(NTh)^{-1/2}\left(\boldsymbol{\Upsilon}_{NT,1}(\tau),\boldsymbol{\Upsilon}_{NT,2}(\tau)'\right)'$.

Consider $\boldsymbol{\Upsilon}_{NT,1}(\tau)$ first. By Lemma B.4(h) and Lemma C.1, we have

$$\frac{1}{\sqrt{NTh}}\boldsymbol{\Upsilon}_{NT,1}(\tau)=\left(\frac{Th}{N}\omega_\tau\right)^{-1}\left(\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}\left(Th\nu_{\tau,i}\right)e_{it}\right)\xrightarrow{d}\mathcal{N}\left(\mathbf{0},\nu_0\Lambda_{\varepsilon,11}(\tau)/\left[\bar{q}(\tau)\right]^2\right), \tag{C.27}$$

where $\Lambda_{\varepsilon,,k_1k_2}(\tau)$ is the $(k_1,k_2)_{th}$ element of $\boldsymbol{\Lambda}_\varepsilon(\tau)$. Namely, the (marginal) asymptotic variance of $\boldsymbol{\Upsilon}_{NT,1}(\tau)$ is $\nu_0\Lambda_{\varepsilon,11}(\tau)/\left[\bar{q}(\tau)\right]^2$.

Now we show that

$$\frac{1}{\sqrt{NTh}}\boldsymbol{\Upsilon}_{NT,2}(\tau)\xrightarrow{d}\mathcal{N}\left(\mathbf{0},\nu_0\left(\Lambda_{\varepsilon,11}(\tau)\boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)'/\left[\bar{q}(\tau)\right]^2+\boldsymbol{\Lambda}_{\nu\varepsilon}(\tau)\right)\right),\qquad (N,T)\to\infty. \tag{C.28}$$

More specifically, plugging $\boldsymbol{x}_{it}=\boldsymbol{\chi}_i+\boldsymbol{\ell}(\tau_t)+\boldsymbol{\nu}_{it}$ in $\boldsymbol{\Upsilon}_{NT,2}(\tau)$, we have

$$\begin{aligned}
\boldsymbol{\Upsilon}_{NT,2}(\tau)&=\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)\boldsymbol{\ell}(\tau_t)M_{it}e_{it}-\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^0(\tau)\boldsymbol{\ell}(\tau_t)M_{it}\right)\left(\sum_{t=1}^{T}w_t^0(\tau)M_{it}e_{it}\right)\\
&+\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}e_{it}-\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\right)\left(\sum_{t=1}^{T}w_t^0(\tau)M_{it}e_{it}\right)\\
&+\left(\frac{Th}{N}\omega_\tau\right)^{-1}\left\{\left(\frac{1}{N}\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}w_t^0(\tau)\boldsymbol{\ell}(\tau_t)M_{it}\right)+\left(\frac{1}{N}\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\right)+\left(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\chi}_i\right)\right\}\\
&\times\left(\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}\left(Th\nu_{\tau,i}\right)e_{it}\right)\\
&=:\boldsymbol{\Upsilon}_{NT,2,1}(\tau)-\boldsymbol{\Upsilon}_{NT,2,2}(\tau)+\boldsymbol{\Upsilon}_{NT,2,3}(\tau)-\boldsymbol{\Upsilon}_{NT,2,4}(\tau)+\boldsymbol{\Upsilon}_{NT,2,5}(\tau).
\end{aligned}$$

We consider $\boldsymbol{\Upsilon}_{NT,2,1}(\tau)-\boldsymbol{\Upsilon}_{NT,2,2}(\tau)$ as a whole. By a Taylor expansion of $\boldsymbol{\ell}(\tau_t)$ around $\boldsymbol{\ell}(\tau)$, we have

$$\begin{aligned}
\frac{1}{\sqrt{NTh}}\boldsymbol{\Upsilon}_{NT,2,1}(\tau)&=\boldsymbol{\ell}(\tau)\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}e_{it}+\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)\left[\boldsymbol{\ell}(\tau_t)-\boldsymbol{\ell}(\tau)\right]M_{it}e_{it}\\
&=\boldsymbol{\ell}(\tau)\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}e_{it}+O_p(h),
\end{aligned}$$

where the second equality is by a straightforward modification of Eq. (C.22) and Markov's inequality.

Similarly,

$$\frac{1}{\sqrt{NTh}}\boldsymbol{\Upsilon}_{NT,2,2}(\tau)=\boldsymbol{\ell}(\tau)\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}e_{it}+O_p(h).$$

Combining these results, we have $(NTh)^{-1/2}\left[\boldsymbol{\Upsilon}_{NT,2,1}(\tau)-\boldsymbol{\Upsilon}_{NT,2,2}(\tau)\right]=O_p(h)$. Note that $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,2,3}(\tau)$ is the first component in Eq. (C.26). Moreover, by Assumptions A5(c), A7(b), and (b), (d) in Lemma B.4, a modification of (C.22) together with Markov's inequality, we have

$$\frac{1}{\sqrt{NTh}}\boldsymbol{\Upsilon}_{NT,2,4}(\tau)=O_p\left(\sqrt{\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}}\right). \tag{C.29}$$

For $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,2,5}(\tau)$, by the Taylor expansion of $\boldsymbol{\ell}(\tau_t)$ around $\boldsymbol{\ell}(\tau)$, Assumptions A5(b) - (c), A7(b), and Lemma B.4(b) - (d), we obtain

$$\begin{aligned}\frac{1}{\sqrt{NTh}}\boldsymbol{\Upsilon}_{NT,2,5}(\tau)&=\left(\frac{Th}{N}\omega_\tau\right)^{-1}\left[\left(\boldsymbol{\ell}(\tau)+o_p(h)\right)+O_p\left(\sqrt{\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}}\right)+O_p\left(\frac{1}{\sqrt{N}}\right)\right]\\&\quad\times\left(\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}\left(Th\nu_{\tau,i}\right)e_{it}\right)\\&=\left(\frac{Th}{N}\omega_\tau\right)^{-1}\boldsymbol{\ell}(\tau)\left(\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}\left(Th\nu_{\tau,i}\right)e_{it}\right)+o_p(1),\end{aligned}$$

implying that $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,2,5}(\tau)$ is asymptotically dominated by the second component in Eq. (C.26). Combining the results above, by Eq. (C.26) and Lemma C.1, we have the marginal distribution of $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,2}(\tau)$ as given in (C.28).

For the asymptotic covariance between $\boldsymbol{\Upsilon}_{NT,1}(\tau)$ and $\boldsymbol{\Upsilon}_{NT,2}(\tau)$, it is only necessary to consider the asymptotic covariance between the asymptotically dominating terms, namely, between $\boldsymbol{\Upsilon}_{NT,1}(\tau)$ and $\boldsymbol{\Upsilon}_{NT,2,3}(\tau)+\boldsymbol{\Upsilon}_{NT,2,5}(\tau)$. Note that $\boldsymbol{\Upsilon}_{NT,1}(\tau)$ and $\boldsymbol{\Upsilon}_{NT,2,3}(\tau)$ are asymptotically independent by Lemma C.1. Hence, the asymptotic covariance between $\boldsymbol{\Upsilon}_{NT,1}(\tau)$ and $\boldsymbol{\Upsilon}_{NT,2}(\tau)$ is determined by $\boldsymbol{\Upsilon}_{NT,1}(\tau)$ and $\boldsymbol{\Upsilon}_{NT,2,5}(\tau)$, resulting in $\nu_0\Lambda_{\varepsilon,11}(\tau)/\left[\bar{q}(\tau)\right]^2\boldsymbol{\ell}(\tau)$. Putting these results together, Claim (CL.1) is obtained.

II. Proof of Claim (CL.2)

Consider $\boldsymbol{\Upsilon}_{NT,3}(\tau)$ first. By employing (b), (c), (i) in Lemma B.4, Markov's inequality, and an adaption of Eq. (C.22), we have

$$\begin{aligned}
&\frac{1}{\sqrt{NTh}}\boldsymbol{\Upsilon}_{NT,3}(\tau)\\
&=\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}e_{it}-\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}(Th\nu_{\tau,i})\left(\frac{1}{Th}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\right)\left(\sum_{t=1}^{T}w_t^0(\tau)M_{it}e_{it}\right)\\
&\quad+\left(\frac{Th}{N}\omega_\tau\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\right)\left(\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^0(\tau)M_{it}\left(Th\nu_{\tau,i}\right)e_{it}\right)\\
&=\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}e_{it}+O_p\left(\sqrt{\frac{N^{1-\eta_{max}}\ln(NT)}{T^{\eta_{max}}h}}+h+\frac{1}{Th^2}\right).
\end{aligned}$$

By Lemma C.1, we have

$$\frac{1}{\sqrt{NTh}}\boldsymbol{\Upsilon}_{NT,3}(\tau)\xrightarrow{d}\mathcal{N}\Big(0,\nu_2\Lambda_{\varepsilon,22}(\tau)\Big),\qquad (N,T)\to\infty. \tag{C.30}$$

Now consider $\boldsymbol{\Upsilon}_{NT,4}(\tau)$. By Lemma C.1, the first component in $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,4}(\tau)$ can be written as:

$$\begin{aligned}
&\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{x}_{it}e_{it}=\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{\chi}_i e_{it}+\boldsymbol{\ell}(\tau)\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}e_{it}\\
&\quad+\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{\nu}_{it}e_{it}+O_p(h)\xrightarrow{d}\mathcal{N}\Big(\mathbf{0},\nu_2\big(\bar{\Lambda}_\varepsilon(\tau)\boldsymbol{\Sigma}_\chi+\Lambda_{\varepsilon,22}(\tau)\boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)'+\boldsymbol{\Lambda}_{\nu\varepsilon}(\tau)\big)\Big).
\end{aligned}$$

Moreover, the last two components in $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,4}(\tau)$ are asymptotically negligible:

$$\begin{aligned}
\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T}w_t^0(\tau)M_{it}e_{it}\right)&=o_p(1),\\
\frac{1}{\sqrt{NTh}}\omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T}w_t^0(\tau)M_{it}e_{it}\right)&=o_p(1),
\end{aligned}$$

using similar arguments for (C.22), Assumptions A5(c) and A7(b), Lemma B.4(b) - (d), and Lemma B.5(g). Combining these results, we have

$$\frac{1}{\sqrt{NTh}}\boldsymbol{\Upsilon}_{NT,4}(\tau)\xrightarrow{d}\mathcal{N}\Big(\mathbf{0},\nu_2\big(\bar{\Lambda}_\varepsilon(\tau)\boldsymbol{\Sigma}_\chi+\Lambda_{\varepsilon,22}(\tau)\boldsymbol{\ell}(\tau)\boldsymbol{\ell}(\tau)'+\boldsymbol{\Lambda}_{\nu\varepsilon}(\tau)\big)\Big). \tag{C.31}$$

As seen, the asymptotically dominating terms in $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,3}(\tau)$ and $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,4}(\tau)$ jointly converge to a Gaussian distribution (Lemma C.1). Moreover, note that the asymptotic covariance between $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,3}(\tau)$ and $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,4}(\tau)$ is determined by the covariance between the asymptotically dominating terms, i.e., $(NTh)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}e_{it}$ and $(NTh)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{x}_{it}e_{it}$. By Lemma C.1, we observe that $(NTh)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}e_{it}$ is asymptotically independent of $(NTh)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{\chi}_i e_{it}$ and $(NTh)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T}w_t^1(\tau)M_{it}\boldsymbol{\nu}_{it}e_{it}$, respectively. The asymptotic covariance between $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,3}(\tau)$ and $(NTh)^{-1/2}\boldsymbol{\Upsilon}_{NT,4}(\tau)$ is therefore fully determined by the covariance between $(NTh)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau)M_{it}e_{it}$ and $\boldsymbol{\ell}(\tau)(NTh)^{-1/2}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau)M_{it}e_{it}$. Claim (CL.2) is now immediate.

III. Proof of Claim (CL.3)

Using the symmetric property of $K(\cdot)$, we have $\nu_1 = \int_{-1}^{1} uK^2(u)\,du = 0$. With this identity, the remaining steps are similar to the proof for Lemma C.1 and are thus omitted. ■

Proof of Theorem 2 Recall the definition $\widehat{\boldsymbol{\theta}}(\tau)$ in (2.10), and $\boldsymbol{\Delta}_i^M(\tau) = \mathrm{diag}(\boldsymbol{m}_i)\boldsymbol{b}_i - \boldsymbol{Z}_i^M(\tau)\boldsymbol{\theta}(\tau)$ in Proposition C.2, where $\boldsymbol{b}_i = (g_1,\ldots,g_T)' + \left(\boldsymbol{x}_{i1}'\boldsymbol{\beta}_1,\ldots,\boldsymbol{x}_{iT}'\boldsymbol{\beta}_T\right)'$. Since

$$\boldsymbol{y}_i^M = \boldsymbol{Z}_i^M(\tau)\boldsymbol{\theta}(\tau) + \boldsymbol{m}_i\alpha_i + \boldsymbol{\Delta}_i^M(\tau) + \boldsymbol{e}_i^M, \tag{C.32}$$

we have

$$\widehat{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}(\tau) = \left(\sum_{i=1}^{N}\widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{Z}}_i^M(\tau)\right)^{-1}\left(\boldsymbol{\mathcal{A}}_{NT}(\tau) + \boldsymbol{\mathcal{B}}_{NT}(\tau) + \boldsymbol{\mathcal{D}}_{NT}(\tau)\right), \tag{C.33}$$

where

$$\boldsymbol{\mathcal{A}}_{NT}(\tau) = \sum_{i=1}^{N}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{K}_h(\tau)\boldsymbol{m}_i\alpha_i - \sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\boldsymbol{k}_h(\tau)'\left(\boldsymbol{m}_i\alpha_i - \omega_\tau^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{m}_i\alpha_i\right).$$

Moreover, $\boldsymbol{\mathcal{B}}_{NT}(\tau)$ and $\boldsymbol{\mathcal{D}}_{NT}(\tau)$ are defined in Propositions C.2 and C.3, respectively.

Consider $\boldsymbol{\mathcal{A}}_{NT}(\tau)$ first. Note that $\sum_{i=1}^{N}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{K}_h(\tau)\boldsymbol{m}_i\alpha_i = \sum_{i=1}^{N}\sum_{t=1}^{T}K\left(\frac{\tau_t-\tau}{h}\right)M_{it}\boldsymbol{z}_{it}(\tau)\alpha_i$. With the identity $\nu_{\tau,i}^{-1} = \boldsymbol{k}_h(\tau)'\boldsymbol{m}_i$ and the identification assumption $\sum_{i=1}^{N}\alpha_i = 0$, we obtain

$$\begin{aligned}
&\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\boldsymbol{k}_h(\tau)'\left(\boldsymbol{m}_i\alpha_i - \omega_\tau^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{m}_i\alpha_i\right)\\
&= \sum_{i=1}^{N}\nu_{\tau,i}\left\{\left(\sum_{t=1}^{T}K\left(\frac{\tau_t-\tau}{h}\right)M_{it}\boldsymbol{z}_{it}(\tau)\right)\left(\alpha_i\nu_{\tau,i}^{-1} - \omega_\tau^{-1}\sum_{i=1}^{N}\alpha_i\right)\right\} = \sum_{i=1}^{N}\sum_{t=1}^{T}K\left(\frac{\tau_t-\tau}{h}\right)M_{it}\boldsymbol{z}_{it}(\tau)\alpha_i.
\end{aligned}$$

Therefore, the two components in $\boldsymbol{\mathcal{A}}_{NT}(\tau)$ cancel out, leading to $\boldsymbol{\mathcal{A}}_{NT}(\tau) = \mathbf{0}$. As a result, Eq. (C.33) can further written as

$$\begin{aligned}
&\sqrt{NTh}\left[\widehat{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}(\tau) - \left(\frac{1}{NTh}\sum_{i=1}^{N}\widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{Z}}_i^M(\tau)\right)^{-1}\left(\frac{1}{NTh}\boldsymbol{\mathcal{B}}_{NT}(\tau)\right)\right]\\
&\qquad = \left(\frac{1}{NTh}\sum_{i=1}^{N}\widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{Z}}_i^M(\tau)\right)^{-1}\left(\frac{1}{\sqrt{NTh}}\boldsymbol{\mathcal{D}}_{NT}(\tau)\right),
\end{aligned} \tag{C.34}$$

where $\boldsymbol{\mathcal{B}}_{NT}(\tau)$ relates to the asymptotic bias, and $\boldsymbol{\mathcal{D}}_{NT}(\tau)$ determines the asymptotic distribution. By Propositions C.1 - C.3, we have Eq. (3.7). ■

Proposition C.4 *Recall $\boldsymbol{\mathcal{D}}_{NT}(\tau)$ as defined in Proposition C.3, and $\phi(\cdot)$ in Assumption A8. Under Assumptions A1, A3, A4, A5, A7(a) - (b), and A8,*

$$\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\boldsymbol{\mathcal{D}}_{NT}(\tau)\right\| = O_p\left(\sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}h}}\right). \tag{C.35}$$

Proof of Proposition C.4 Recall $\boldsymbol{x}_{it} = \boldsymbol{\chi}_i + \boldsymbol{\ell}(\tau_t) + \boldsymbol{\nu}_{it}$ and $e_{it} = \sigma_i(\tau_t)\varepsilon_{it}$, where $\sigma_i(\cdot) \in [\sigma_L, \sigma_U] \subset \mathbb{R}^+$ in Assumption A4, $i = 1, \ldots, N$. By the decomposition (C.25) and the corresponding discussions concerning $\boldsymbol{\Upsilon}_{NT,1}(\tau), \ldots, \boldsymbol{\Upsilon}_{NT,4}(\tau)$ provided below the equation, it is not hard to see that the asymptotic order of $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\boldsymbol{\mathcal{D}}_{NT}(\tau)\right\|$ is determined by

(a) $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\sum_{i=1}^N\sum_{t=1}^T w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}e_{it}\right\| + C\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\sum_{i=1}^N\sum_{t=1}^T w_t^0(\tau)M_{it}\left(Th\nu_{\tau,i}\right)e_{it}\right\|$;

(b) $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\sum_{i=1}^N\sum_{t=1}^T w_t^1(\tau)M_{it}e_{it}\right\|$;

(c) $\sup_{\tau\in[0,1]}\left\|(NTh)^{-1}\sum_{i=1}^N\sum_{t=1}^T w_t^1(\tau)M_{it}\boldsymbol{x}_{it}e_{it}\right\|$.

It is not hard to show that $\sup_{\tau\in[0,1]}\left|(NTh)^{-1}\sum_{i=1}^N\sum_{t=1}^T w_t^0(\tau)M_{it}\left(Th\nu_{\tau,i}\right)e_{it}\right|$ is asymptotically dominated by $\sup_{\tau\in[0,1]}\left|(NTh)^{-1}\sum_{i=1}^N\sum_{t=1}^T w_t^0(\tau)M_{it}\left[p_i(\tau)\right]^{-1}e_{it}\right|$. With the definition of $\boldsymbol{x}_{it}$, the asymptotic orders of the terms in Parts (a) - (c) boil down to linear combinations of the intermediate terms found in (C.36) - (C.38) below. More specifically, based on Theorem 1, the following bounds can be obtained:

$$\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T w_t^{k_1}(\tau)\boldsymbol{\ell}^{k_2}(\tau_t)M_{it}\left[p_i(\tau)\right]^{-1}e_{it}\right\| = O_p\left(\sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}h}}\right), \tag{C.36}$$

$$\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T w_t^{k}(\tau)M_{it}\boldsymbol{\nu}_{it}e_{it}\right\| = O_p\left(\sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}h}}\right), \tag{C.37}$$

$$\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\sum_{i=1}^N\sum_{t=1}^T w_t^{k}(\tau)M_{it}\boldsymbol{\chi}_{i}e_{it}\right\| = O_p\left(\sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}h}}\right), \tag{C.38}$$

where k, k_1, k_2 are nonnegative integers. Note that for (C.38), the process $\{M_{it}\boldsymbol{\chi}_i e_{it},\, t \geq 1\}$, $i = 1, \ldots, N$, is α-mixing (and thus NED) conditional on $\{\boldsymbol{\chi}_i,\, i = 1\ldots, N\}$, and therefore (C.38) immediately holds by Theorem 1 conditional on $\{\boldsymbol{\chi}_i,\, i = 1\ldots, N\}$. By Xiong and Li (2008, Theorem 3.3), it also holds unconditionally. Combining these results, we obtain (C.35). ■

Proof of Proposition 1 By Eq. (C.33),

$$\sup_{\tau\in[0,1]}\left\|\widehat{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}(\tau)\right\| \leq \sup_{\tau\in[0,1]}\left\|\left(\frac{1}{NTh}\sum_{i=1}^N\widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{Z}}_i^M(\tau)\right)^{-1}\right\| \times\left(\sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\boldsymbol{\mathcal{B}}_{NT}(\tau)\right\| + \sup_{\tau\in[0,1]}\left\|\frac{1}{NTh}\boldsymbol{\mathcal{D}}_{NT}(\tau)\right\|\right), \tag{C.39}$$

where $\boldsymbol{\mathcal{B}}_{NT}(\tau)$ and $\boldsymbol{\mathcal{D}}_{NT}(\tau)$ are defined in Propositions C.2 and C.3, respectively. Since $(\boldsymbol{A}+\boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} - \boldsymbol{A}^{-1}\boldsymbol{B}(\boldsymbol{A}+\boldsymbol{B})^{-1}$ for invertible matrices $\boldsymbol{A}$ and $\boldsymbol{A}+\boldsymbol{B}$, Proposition C.1 implies

$$\sup_{\tau\in[0,1]} \left\| \left(\frac{1}{NTh} \sum_{i=1}^{N} \widetilde{\boldsymbol{Z}}_i^M(\tau)' \widetilde{\boldsymbol{Z}}_i^M(\tau) \right)^{-1} \right\| = O_p(1). \tag{C.40}$$

Finally, we have $\sup_{\tau\in[0,1]} \left\|(NTh)^{-1}\boldsymbol{\mathcal{B}}_{NT}(\tau)\right\| = O_p(h^2)$ by Proposition C.2. Combining these results with Proposition C.4 leads to (3.6). ■

Proof of Corollary 1 By Eq. (C.32), we can write

$$\max_{1\le i\le N} |\widehat{\alpha}_i - \alpha_i| \le \max_{1\le i\le N} \sup_{\tau\in[0,1]} \left|\widehat{\alpha}_i^{\dagger}(\tau) - \alpha_i\right| \le \max_{1\le i\le N} \sup_{\tau\in[0,1]} \left|\nu_{\tau,i}\boldsymbol{k}_h(\tau)'\left(\boldsymbol{Z}_i^M(\tau) - \bar{\boldsymbol{Z}}^M(\tau)\right)\left(\widehat{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}(\tau)\right)\right|$$
$$+ \max_{1\le i\le N} \sup_{\tau\in[0,1]} \left|\nu_{\tau,i}\boldsymbol{k}_h(\tau)'\left(\boldsymbol{\Delta}_i^M(\tau) - \omega_\tau^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{\Delta}_i^M(\tau)\right)\right| + \max_{1\le i\le N} \sup_{\tau\in[0,1]} \left|\nu_{\tau,i}\boldsymbol{k}_h(\tau)'\left(\boldsymbol{e}_i^M - \omega_\tau^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{e}_i^M\right)\right|.$$

By applying Parts *(b)*, *(c)*, *(f)*, and *(h)* in Lemma B.4 and making use of Proposition 1, we obtain:

$$\max_{1\le i\le N} \sup_{\tau\in[0,1]} \left|\nu_{\tau,i}\boldsymbol{k}_h(\tau)'\left(\boldsymbol{Z}_i^M(\tau) - \bar{\boldsymbol{Z}}^M(\tau)\right)\left(\widehat{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}(\tau)\right)\right| \le \max_{1\le i\le N} \sup_{\tau\in[0,1]} \Bigg\{ (Th\nu_{\tau,i}) \times$$
$$\left\| \frac{1}{Th}\boldsymbol{k}_h(\tau)'\left(\boldsymbol{Z}_i^M(\tau) - \bar{\boldsymbol{Z}}^M(\tau)\right)\right\| \left\|\widehat{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}(\tau)\right\| \Bigg\} = O_p\left(h^2 + \sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}h}}\right). \tag{C.41}$$

Similarly, by Eq. (C.7), we have

$$\max_{1\le i\le N} \sup_{\tau\in[0,1]} \left|\nu_{\tau,i}\boldsymbol{k}_h(\tau)'\left(\boldsymbol{\Delta}_i^M(\tau) - \omega_\tau^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{\Delta}_i^M(\tau)\right)\right| = O_p(h^2). \tag{C.42}$$

Moreover, a straightforward modification of Lemma B.4*(d)* yields

$$\max_{1\le i\le N} \sup_{\tau\in[0,1]} \left|(Th)^{-1}\boldsymbol{k}_h(\tau)'\boldsymbol{e}_i^M\right| = O_p\left(\sqrt{\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}}\right), \tag{C.43}$$

leading to

$$\max_{1\le i\le N} \sup_{\tau\in[0,1]} \left|\nu_{\tau,i}\boldsymbol{k}_h(\tau)'\left(\boldsymbol{e}_i^M - \omega_\tau^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{e}_i^M\right)\right| = O_p\left(\sqrt{\frac{N^{\varpi(p_0)}\ln(NT)}{T^{\phi(p_0)}h}}\right). \tag{C.44}$$

Combining these results implies (3.8). ■

D Proofs of bootstrap validity

As in Gonçalves and Perron (2014, Appendix B), we will frequently use the property that $O_p^*(1)O_p(1) = O_p^*(1)O_p^*(1) = O_p^*(1)$ in probability, and $O_p^*(1)o_p(1) = O_p^*(1)o_p^*(1) = o_p^*(1)$ in probability, in the following proofs. A justification for this can be found in Lemma 3 of Cheng and Huang (2010).

Lemma D.1 *Under Assumptions A1 - A8, B1 - B2, for any fixed $\tau \in (0,1)$, $k = 0, 1, \ldots$, we have: as $(N,T) \to \infty$,*

$$\frac{1}{\sqrt{NTh}} \sum_{i=1}^{N} \sum_{t=1}^{T} w_t^k(\tau) M_{it} \begin{pmatrix} Th\nu_{\tau,i} \\ 1 \\ \boldsymbol{\chi}_i \\ \boldsymbol{\nu}_{it} \end{pmatrix} \xi_t^* \widetilde{e}_{it} \xrightarrow{d^*}_p \mathcal{N}\left(\mathbf{0}, \nu_{2k} \begin{pmatrix} \boldsymbol{\Lambda}_{\varepsilon}(\tau) & & \\ & \bar{\Lambda}_{\varepsilon}(\tau)\boldsymbol{\Sigma}_{\chi} & \\ & & \boldsymbol{\Lambda}_{\nu\varepsilon}(\tau) \end{pmatrix}\right), \tag{D.1}$$

which mimics the asymptotic distribution as given in Lemma C.1.

Proof of Lemma D.1 Similar to Lemma C.1, the Cramér-Wold device is employed to establish the asymptotic joint distribution. Specifically, for any unit vector $\boldsymbol{a} = (a_1, \boldsymbol{a}_2')' \in \mathbb{R}^{2(d+1)}$, we consider

$$\sum_{t=1}^{T} \widetilde{\mathcal{Y}}_{NT,t}^*(\tau) + a_1 \frac{1}{\sqrt{NTh}} \sum_{i=1}^{N} \sum_{t=1}^{T} w_t^k(\tau) M_{it} \left\{ (Th\nu_{\tau,i}) - [p_i(\tau)]^{-1} \right\} \xi_t^* \widetilde{e}_{it}, \tag{D.2}$$

where $\widetilde{\mathcal{Y}}_{NT,t}^*(\tau) = (NTh)^{-1/2} w_t^k(\tau) \sum_{i=1}^{N} M_{it} \boldsymbol{a}' \left[1/p_i(\tau), 1, \boldsymbol{\chi}_i', \boldsymbol{\nu}_{it}'\right]' \widetilde{e}_{it} \xi_t^*$.

Note that the sequence $\{\boldsymbol{w}_{it} := (\boldsymbol{\chi}_i, \boldsymbol{\nu}_{it}), i = 1, \ldots, N, t = 1, \ldots, T\}$ is not directly observable and therefore does not fall within our information set. To simplify the proof, we first condition on the data and $\{\boldsymbol{w}_{it}\}$, from which we derive the conditional asymptotic normality of $\sum_{t=1}^{T} \widetilde{\mathcal{Y}}_{NT,t}^*(\tau)$. Given that the limiting distribution does not depend on $\{\boldsymbol{w}_{it}\}$, our results also hold unconditionally with respect to the sequence.

I. Asymptotic mean and variance of $\sum_{t=1}^{T} \widetilde{\mathcal{Y}}_{NT,t}^*(\tau)$

Recall the notation defined below (C.13). Namely, $\boldsymbol{A}_{k_1k_2}(i, t_i; j, t_j) = \boldsymbol{a}_{k_1}(i, t_i)\boldsymbol{a}_{k_2}(j, t_j)'$, $1 \le k_1, k_2 \le 4$, where $\boldsymbol{a}_1(\ell, t_\ell) = 1/p_\ell(\tau)$, $\boldsymbol{a}_2(\ell, t_\ell) = 1$, $\boldsymbol{a}_3(\ell, t_\ell) = \boldsymbol{\chi}_\ell$, and $\boldsymbol{a}_4(\ell, t_\ell) = \boldsymbol{\nu}_{\ell t_\ell}$, for $\ell = i$ or $\ell = j$. Furthermore, let $\mathbb{E}_{\boldsymbol{w}}^*(\cdot) = \mathbb{E}\big(\,\cdot\,\big|\,\{(\boldsymbol{x}_{it}, y_{it}, M_{it}, \boldsymbol{w}_{it}), i = 1, \ldots, N, t = 1, \ldots, T\}\big)$. Since $\{\xi_t^*\}$ is independent of $\{(\boldsymbol{x}_{it}, y_{it}, M_{it}, \boldsymbol{w}_{it})\}$, it follows that $\mathbb{E}_{\boldsymbol{w}}^*(\xi_t^*) = 0$ and $\mathbb{E}_{\boldsymbol{w}}^*(\xi_t^*\xi_s^*) = \gamma^{|t-s|}$. As a result, we have $\mathbb{E}_{\boldsymbol{w}}^*\left(\sum_{t=1}^{T} \widetilde{\mathcal{Y}}_{NT,t}^*(\tau)\right) = 0$. Furthermore, by Eq. (C.13), we obtain $\mathbb{E}_{\boldsymbol{w}}^*\left(\sum_{t=1}^{T} \widetilde{\mathcal{Y}}_{NT,t}^*(\tau)\right)^2 = \boldsymbol{a}'\widetilde{\boldsymbol{\mathcal{E}}}_{NT}(\tau)\boldsymbol{a}$, where $\widetilde{\boldsymbol{\mathcal{E}}}_{NT}(\tau) = \left(\widetilde{\boldsymbol{\mathcal{E}}}_{NT,k_1k_2}(\tau),\ k_1,k_2 = 1,2,3,4\right)$ is a block matrix with the elements given by

$$\begin{aligned}
\widetilde{\boldsymbol{\mathcal{E}}}_{NT,k_1k_2}(\tau) = \frac{1}{NTh}\sum_{i,j=1}^{N}\Bigg\{&\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2 M_{it}M_{jt}\boldsymbol{A}_{k_1k_2}(i,t;j,t)\widetilde{e}_{it}\widetilde{e}_{jt} \\
&+\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\,w_{t+s}^k(\tau)\,M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)\widetilde{e}_{it}\widetilde{e}_{j(t+s)} \\
&+\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\,w_{t+s}^k(\tau)\,M_{i(t+s)}M_{jt}\boldsymbol{A}_{k_1k_2}(i,t+s;j,t)\widetilde{e}_{i(t+s)}\widetilde{e}_{jt}\Bigg\} \\
=:\ & \widetilde{\mathrm{I}}_{NT,k_1,k_2}(\tau)+\widetilde{\mathrm{II}}_{NT,k_1,k_2}(\tau)+\widetilde{\mathrm{III}}_{NT,k_1,k_2}(\tau).
\end{aligned} \tag{D.3}$$

Note that $\widetilde{e}_{it} = M_{it}\big[e_{it} + (\alpha_i - \widetilde{\alpha}_i) + (g_t - \tilde{g}_t) + \boldsymbol{x}_{it}'(\boldsymbol{\beta}_t - \widetilde{\boldsymbol{\beta}}_t)\big]$. Using Proposition 1 and Corollary 1, we can express $\widetilde{\mathrm{I}}_{NT,k_1,k_2}(\tau)$ as follows:

$$\widetilde{\mathrm{I}}_{NT,k_1,k_2}(\tau) = \frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{t=1}^{T}\left[w_t^k(\tau)\right]^2 M_{it}M_{jt}\boldsymbol{A}_{k_1k_2}(i,t;j,t)e_{it}e_{jt} + o_p(1). \tag{D.4}$$

We will maintain $\widetilde{\mathrm{I}}_{NT,k_1,k_2}(\tau)$ for now and proceed to simplify $\widetilde{\mathrm{II}}_{NT,k_1,k_2}(\tau)$ and $\widetilde{\mathrm{III}}_{NT,k_1,k_2}(\tau)$. The steps involved in simplifying $\widetilde{\mathrm{II}}_{NT,k_1,k_2}(\tau)$ and $\widetilde{\mathrm{III}}_{NT,k_1,k_2}(\tau)$ are similar, with the primary challenge lying in $\widetilde{\mathrm{II}}_{NT,k_1,k_2}(\tau)$. To proceed, we first split $\widetilde{\mathrm{II}}_{NT,k_1,k_2}(\tau)$ as follows:

$$\begin{aligned}
\widetilde{\mathrm{II}}_{NT,k_1,k_2}(\tau) =& \frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\,w_{t+s}^k(\tau)\,M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)e_{it}e_{j(t+s)} \\
&+\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\,w_{t+s}^k(\tau)\,M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)\left(\widetilde{e}_{it}\widetilde{e}_{j(t+s)} - e_{it}e_{j(t+s)}\right) \\
=:\ & \widetilde{\mathrm{II}}^{a}_{NT,k_1,k_2}(\tau)+\widetilde{\mathrm{II}}^{b}_{NT,k_1,k_2}(\tau).
\end{aligned} \tag{D.5}$$

We prove $\widetilde{\mathrm{II}}^{b}_{NT,k_1,k_2}(\tau) = o_p(1)$ first and address $\widetilde{\mathrm{II}}^{a}_{NT,k_1,k_2}(\tau)$ subsequently.

Recall the notation $\boldsymbol{\Upsilon}(\tau)$ in Proposition C.2, and define $\boldsymbol{\Upsilon}_t = \boldsymbol{\Upsilon}(\tau_t) = (g_t, \boldsymbol{\beta}_t')'$, $\widetilde{\boldsymbol{\Upsilon}}_t = \widetilde{\boldsymbol{\Upsilon}}(\tau_t) = (\tilde{g}_t, \widetilde{\boldsymbol{\beta}}_t')'$. We have

$$\begin{aligned}
M_{it}M_{j(t+s)}\left(\widetilde{e}_{it}\widetilde{e}_{j(t+s)} - e_{it}e_{j(t+s)}\right) =\ & M_{it}M_{j(t+s)}\Big[(\alpha_j-\widetilde{\alpha}_j)e_{it} + (\alpha_i - \widetilde{\alpha}_i)e_{j(t+s)} + (\alpha_i-\widetilde{\alpha}_i)(\alpha_j-\widetilde{\alpha}_j) \\
&+ (\alpha_i-\widetilde{\alpha}_i)(1,\boldsymbol{x}_{j(t+s)}')(\boldsymbol{\Upsilon}_{t+s}-\widetilde{\boldsymbol{\Upsilon}}_{t+s}) + (\alpha_j-\widetilde{\alpha}_j)(1,\boldsymbol{x}_{it}')(\boldsymbol{\Upsilon}_t-\widetilde{\boldsymbol{\Upsilon}}_t) + e_{it}(1,\boldsymbol{x}_{j(t+s)}')(\boldsymbol{\Upsilon}_{t+s}-\widetilde{\boldsymbol{\Upsilon}}_{t+s}) \\
&+ e_{j(t+s)}(1,\boldsymbol{x}_{it}')(\boldsymbol{\Upsilon}_t-\widetilde{\boldsymbol{\Upsilon}}_t) + (\boldsymbol{\Upsilon}_t-\widetilde{\boldsymbol{\Upsilon}}_t)'(1,\boldsymbol{x}_{it}')'(1,\boldsymbol{x}_{j(t+s)}')(\boldsymbol{\Upsilon}_{t+s}-\widetilde{\boldsymbol{\Upsilon}}_{t+s})\Big].
\end{aligned} \tag{D.6}$$

This decomposition results in $\widetilde{\mathrm{II}}^{b}_{NT,k_1,k_2}(\tau)$ being split into a summation of eight components, denoted as $\widetilde{\mathrm{II}}^{b,\ell}_{NT,k_1,k_2}(\tau)$ for $\ell = 1,\ldots,8$. The details of $\widetilde{\mathrm{II}}^{b,1}_{NT,k_1,k_2}(\tau),\ldots,\widetilde{\mathrm{II}}^{b,8}_{NT,k_1,k_2}(\tau)$ are provided below, i.e., Eqs. (D.9) - (D.16).

Before we proceed, it is worth noting that

$$\sum_{s=0}^{\infty}\gamma^s = \sum_{s=0}^{\infty}\theta^{s/\ell} = 1/(1-\theta^{1/\ell}) = -\ell/\ln(\theta) + o(\ell) = O(\ell), \qquad \forall\theta\in(0,1). \tag{D.7}$$

Furthermore, by applying the Cauchy-Schwarz inequality, we have:

$$\begin{aligned}\left(\mathbb{E}\left\|\sum_{i=1}^{N} M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)e_{it}\right\|^{\xi_0/2}\right)^{2/\xi_0} \\ \leq C\left(\mathbb{E}\left\|\sum_{i=1}^{N} M_{it}\sigma_i(\tau_t)\boldsymbol{a}_{k_1}(i,t)\varepsilon_{it}\right\|^{\xi_0}\right)^{1/\xi_0} \leq CN^{1/2},\end{aligned} \tag{D.8}$$

uniformly in $j,t,s\geq 1$, $\xi_0 = p_0/2$. The final inequality above is obtained from Assumption A7(a).

Using (D.7) and (D.8), and based on the exogeneity condition of $\{\varepsilon_{it}\}$ in Assumption A7(b), we can bound $\widetilde{\Pi}^{b,1}_{NT,k_1,k_2}(\tau)$ as follows:

$$\begin{aligned}\left\|\widetilde{\Pi}^{b,1}_{NT,k_1,k_2}(\tau)\right\| &= \left\|\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\, M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)(\alpha_j-\widetilde{\alpha}_j)e_{it}\right\| \\ &\leq \sum_{j=1}^{N}|\alpha_j-\widetilde{\alpha}_j|\sum_{s=1}^{T-1}\gamma^s\left\|\frac{1}{NTh}\sum_{i=1}^{N}\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\, M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)e_{it}\right\| \\ &= N\max_{1\leq j\leq N}|\alpha_j-\widetilde{\alpha}_j|\, O(\ell)\, O_p\left(\frac{1}{\sqrt{N}}\right) = o_p(1),\end{aligned} \tag{D.9}$$

where the second-to-last equality is deduced from the Triplex inequality (Jiang, 2009) and Assumption A8, following similar steps to the proof of Lemma C.1. Moreover, the final equality is due to Assumption B1. Similarly, we have

$$\widetilde{\Pi}^{b,2}_{NT,k_1,k_2}(\tau) = \frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\, M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)(\alpha_i-\widetilde{\alpha}_i)e_{j(t+s)} = o_p(1). \tag{D.10}$$

Moreover,

$$\begin{aligned}\left\|\widetilde{\Pi}^{b,3}_{NT,k_1,k_2}(\tau)\right\| &= \left\|\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\, M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)(\alpha_i-\widetilde{\alpha}_i)(\alpha_j-\widetilde{\alpha}_j)\right\| \\ &\leq \frac{1}{NTh}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\sum_{i,j=1}^{N}\left\|M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)(\alpha_i-\widetilde{\alpha}_i)(\alpha_j-\widetilde{\alpha}_j)\right\| \\ &\leq \left(\max_{1\leq j\leq N}|\alpha_j-\widetilde{\alpha}_j|\right)^2\frac{1}{NTh}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\sum_{i,j=1}^{N}\left\|\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)\right\| \\ &= \left(\max_{1\leq j\leq N}|\alpha_j-\widetilde{\alpha}_j|\right)^2 O_p(N\ell) = o_p(1),\end{aligned} \tag{D.11}$$

using Markov's inequality in the second-to-last step. For $\widetilde{\Pi}^{b,4}_{NT,k_1,k_2}(\tau)$, the application of Proposition 1 and Corollary 1 implies

$$\begin{aligned}
&\left\|\widetilde{\Pi}^{b,4}_{NT,k_1,k_2}(\tau)\right\| \\
&= \left\|\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\, M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)(\alpha_i-\widetilde{\alpha}_i)\left[\big(1,\boldsymbol{x}'_{j(t+s)}\big)(\boldsymbol{\Upsilon}_{t+s}-\widetilde{\boldsymbol{\Upsilon}}_{t+s})\right]\right\| \\
&\le \frac{1}{NTh}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\sum_{i,j=1}^{N}\left\|M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)(\alpha_i-\widetilde{\alpha}_i)\left[\big(1,\boldsymbol{x}'_{j(t+s)}\big)(\boldsymbol{\Upsilon}_{t+s}-\widetilde{\boldsymbol{\Upsilon}}_{t+s})\right]\right\| \\
&= \max_{1\le j\le N}|\alpha_j-\widetilde{\alpha}_j|\sup_{\tau\in[0,1]}\left\|\widetilde{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}(\tau)\right\| O_p(N\ell)=o_p(1). \qquad \text{(D.12)}
\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}
\widetilde{\Pi}^{b,5}_{NT,k_1,k_2}(\tau) &= \frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\, M_{it}M_{j(t+s)} \\
&\quad\times \boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)(\alpha_j-\widetilde{\alpha}_j)\left[\big(1,\boldsymbol{x}'_{it}\big)(\boldsymbol{\Upsilon}_t-\widetilde{\boldsymbol{\Upsilon}}_t)\right]=o_p(1). \qquad \text{(D.13)}
\end{aligned}$$

We now proceed to $\widetilde{\Pi}^{b,6}_{NT,k_1,k_2}(\tau)$. Note that for any $\boldsymbol{A}\in\mathbb{R}^{d_1\times d_2}$ and $\boldsymbol{a},\boldsymbol{b}\in\mathbb{R}^{d_3}$, $d_1,d_2,d_3\in\mathbb{Z}^+$, we have $\|\boldsymbol{A}(\boldsymbol{a}'\boldsymbol{b})\|=\|(\boldsymbol{A}\otimes\boldsymbol{a}')(\boldsymbol{I}_{d_2}\otimes\boldsymbol{b})\|\le\|\boldsymbol{A}\otimes\boldsymbol{a}'\|\,\|\boldsymbol{I}_{d_2}\otimes\boldsymbol{b}\|\le\|\boldsymbol{A}\otimes\boldsymbol{a}'\|\,\|\boldsymbol{b}\|$. Hence, $\widetilde{\Pi}^{b,6}_{NT,k_1,k_2}(\tau)$ can be bounded by

$$\begin{aligned}
&\left\|\widetilde{\Pi}^{b,6}_{NT,k_1,k_2}(\tau)\right\| \\
&= \left\|\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\, M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)e_{it}\big(1,\boldsymbol{x}'_{j(t+s)}\big)(\boldsymbol{\Upsilon}_{t+s}-\widetilde{\boldsymbol{\Upsilon}}_{t+s})\right\| \\
&= \left\|\frac{1}{NTh}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\left[\sum_{i,j=1}^{N}M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)e_{it}\big(1,\boldsymbol{x}'_{j(t+s)}\big)\right](\boldsymbol{\Upsilon}_{t+s}-\widetilde{\boldsymbol{\Upsilon}}_{t+s})\right\| \\
&\le \sup_{\tau\in[0,1]}\left\|\widetilde{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}(\tau)\right\|\frac{1}{NTh}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s} w_t^k(\tau)\, w_{t+s}^k(\tau)\left\|\sum_{i,j=1}^{N}M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)e_{it}\otimes\big(1,\boldsymbol{x}'_{j(t+s)}\big)\right\| \\
&= \sup_{\tau\in[0,1]}\left\|\widetilde{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}(\tau)\right\| O_p(\sqrt{N}\ell)=o_p(1), \qquad \text{(D.14)}
\end{aligned}$$

where the second-to-last equality is derived as follows. More specifically, by the Markov's inequality, for any $\epsilon > 0$, there exists $M_\epsilon < \infty$ such that

$$\mathbb{P}\left(\frac{1}{NTh}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\left\|\sum_{i,j=1}^{N}M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)e_{it}\otimes\big(1,\boldsymbol{x}_{j(t+s)}'\big)\right\|\geq\sqrt{N}\ell M_\epsilon\right)$$
$$\leq\frac{1}{M_\epsilon}\frac{1}{\ell}\frac{1}{Th}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\,\mathbb{E}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}M_{it}\boldsymbol{a}_{k_1}(i,t)e_{it}\frac{1}{N}\sum_{j=1}^{N}M_{j(t+s)}\boldsymbol{a}_{k_2}(j,t+s)'\otimes\big(1,\boldsymbol{x}_{j(t+s)}'\big)\right\|$$
$$\leq C\frac{1}{M_\epsilon}\frac{1}{\ell}\sum_{s=1}^{T-1}\gamma^s\frac{1}{Th}\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\left\{\frac{1}{N}\mathbb{E}\left(\sum_{i,j=1}^{N}M_{it}M_{jt}\boldsymbol{a}_{k_1}(i,t)'\boldsymbol{a}_{k_1}(j,t)e_{it}e_{jt}\right)\right\}^{1/2}\leq C/M_\epsilon\leq\epsilon,$$

where the second inequality follows from the Cauchy-Schwarz and c_r inequalities and Assumption A5, and the third one is due to Eq. (D.7) and *(b)*, *(d)* in Assumption A7. Similarly, one could obtain

$$\widetilde{\text{II}}_{NT,k_1,k_2}^{b,7}(\tau)=\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\,M_{it}M_{j(t+s)}$$
$$\times\boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)e_{j(t+s)}\big(1,\boldsymbol{x}_{it}'\big)(\boldsymbol{\varUpsilon}_t-\widetilde{\boldsymbol{\varUpsilon}}_t)=o_p(1).\quad\text{(D.15)}$$

Finally, we have

$$\left\|\widetilde{\text{II}}_{NT,k_1,k_2}^{b,8}(\tau)\right\|$$
$$=\left\|\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\,M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)(\boldsymbol{\varUpsilon}_t-\widetilde{\boldsymbol{\varUpsilon}}_t)'(1,\boldsymbol{x}_{it}')'\big(1,\boldsymbol{x}_{j(t+s)}'\big)(\boldsymbol{\varUpsilon}_{t+s}-\widetilde{\boldsymbol{\varUpsilon}}_{t+s})\right\|$$
$$\leq\left(\sup_{\tau\in[0,1]}\left\|\widetilde{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}(\tau)\right\|\right)^2\frac{1}{NTh}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\sum_{i,j=1}^{N}\left\|\boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)\right\|\left\|(1,\boldsymbol{x}_{it}')'\big(1,\boldsymbol{x}_{j(t+s)}'\big)\right\|$$
$$=\left(\sup_{\tau\in[0,1]}\left\|\widetilde{\boldsymbol{\theta}}(\tau)-\boldsymbol{\theta}(\tau)\right\|\right)^2O_p(N\ell)=o_p(1).\quad\text{(D.16)}$$

Combining (D.9) - (D.16) leads to $\widetilde{\text{II}}_{NT,k_1,k_2}^{b}(\tau)=o_p(1)$, and therefore

$$\widetilde{\text{II}}_{NT,k_1,k_2}(\tau)=\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\,M_{it}M_{j(t+s)}\boldsymbol{A}_{k_1k_2}\big(i,t;j,t+s\big)e_{it}e_{j(t+s)}+o_p(1).\quad\text{(D.17)}$$

Similarly, one could have

$$\widetilde{\text{III}}_{NT,k_1,k_2}(\tau)=\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{s=1}^{T-1}\gamma^s\sum_{t=1}^{T-s}w_t^k(\tau)\,w_{t+s}^k(\tau)\,M_{i(t+s)}M_{jt}\boldsymbol{A}_{k_1k_2}\big(i,t+s;j,t\big)e_{i(t+s)}e_{jt}+o_p(1).\quad\text{(D.18)}$$

By (D.3), (D.4), (D.17), (D.18), $\widetilde{\boldsymbol{\mathcal{E}}}_{NT,k_1k_2}(\tau)$ can be equivalently written as

$$\widetilde{\boldsymbol{\mathcal{E}}}_{NT,k_1k_2}(\tau)=\frac{1}{NTh}\sum_{i,j=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}w_t^k(\tau)\,M_{it}\boldsymbol{a}_{k_1}(i,t)e_{it}w_s^k(\tau)\,M_{js}\boldsymbol{a}_{k_2}(j,s)'e_{js}\gamma^{|t-s|}+o_p(1).$$

Since $\gamma^{|t-s|} = \theta^{|t-s|/\ell}$, we then have

$$\mathbb{E}_{\boldsymbol{w}}^{*}\left(\sum_{t=1}^{T}\widetilde{\mathcal{Y}}_{NT,t}^{*}(\tau)\right)^{2} = \boldsymbol{a}'\left\{\sum_{t=1}^{T}\sum_{s=1}^{T}\boldsymbol{X}_{NT,t}(\tau)\boldsymbol{X}_{NT,s}(\tau)'\theta^{|t-s|/\ell}\right\}\boldsymbol{a} + o_p(1), \qquad \forall\tau\in(0,1), \quad \text{(D.19)}$$

where

$$\boldsymbol{X}_{NT,t}(\tau) = \frac{1}{\sqrt{Th}}w_t^k(\tau)\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sigma_i(\tau_t)M_{it}\left[1/p_i(\tau), 1, \boldsymbol{\chi}_i', \boldsymbol{\nu}_{it}'\right]'\varepsilon_{it}\right), \qquad t = 1,\ldots,T. \quad \text{(D.20)}$$

Note that the function $x \mapsto \theta^{|x|}$ satisfies Assumption 1 of de Jong and Davidson (2000) for any $\theta\in(0,1)$. Moreover, define $\mathbb{E}_{\boldsymbol{\chi}}(\cdot) = \mathbb{E}\left(\cdot\,|\,\{\boldsymbol{\chi}_i\}_{i=1}^{N}\right)$ and recall $\mathcal{F}_{\xi,t_1}^{t_2} = \sigma\left(\boldsymbol{\xi}_{\bullet t},\, t_1\le t\le t_2\right)$. Then, for any $\tau\in(0,1)$,

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\chi}}\left\|\boldsymbol{X}_{NT,t}(\tau) - \mathbb{E}_{\boldsymbol{\chi}}\left(\boldsymbol{X}_{NT,t}(\tau)\,\middle|\,\mathcal{F}_{\xi,t-m}^{t+m}\right)\right\|^2 \\
=& \left(\frac{1}{\sqrt{Th}}w_t^k(\tau)\right)^2\mathbb{E}_{\boldsymbol{\chi}}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sigma_i(\tau_t)\left[M_{it}\left[1/p_i(\tau),1,\boldsymbol{\chi}_i',\boldsymbol{\nu}_{it}'\right]'\varepsilon_{it} - \mathbb{E}_{\boldsymbol{\chi}}\left(M_{it}\left[1/p_i(\tau),1,\boldsymbol{\chi}_i',\boldsymbol{\nu}_{it}'\right]'\varepsilon_{it}\,\middle|\,\mathcal{F}_{\xi,t-m}^{t+m}\right)\right]\right\|^2 \\
=& \left(\frac{1}{\sqrt{Th}}w_t^k(\tau)\right)^2\mathbb{E}_{\boldsymbol{\chi}}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sigma_i(\tau_t)M_{it}\varepsilon_{it}\left[\boldsymbol{\nu}_{it} - \mathbb{E}_{\boldsymbol{\chi}}\left(\boldsymbol{\nu}_{it}\,\middle|\,\mathcal{F}_{\xi,t-m}^{t+m}\right)\right]\right\|^2 \\
=& \left(\frac{1}{\sqrt{Th}}w_t^k(\tau)\right)^2\mathbb{E}_{\boldsymbol{\chi}}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sigma_i(\tau_t)M_{it}\varepsilon_{it}\left[\boldsymbol{\nu}_{it} - \mathbb{E}\left(\boldsymbol{\nu}_{it}\,\middle|\,\mathcal{F}_{\xi,t-m}^{t+m}\right)\right]\right\|^2 \\
\le& \left(\frac{1}{\sqrt{Th}}w_t^k(\tau)\right)^2\sup_{t\ge1}\mathbb{E}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sigma_i(\tau_t)M_{it}\varepsilon_{it}\left(\boldsymbol{\nu}_{it} - \boldsymbol{\nu}_{it}^{(m)}\right)\right\|^2 \le C\left(\frac{1}{\sqrt{Th}}w_t^k(\tau)\right)^2 m^{-\varsigma}, \qquad \text{(D.21)}
\end{aligned}$$

by Assumption B2 and the independence between $\{\boldsymbol{\chi}_i,\, i=1,\ldots,N\}$ and $\{\boldsymbol{\xi}_{\bullet t},\, t=1,\ldots,T\}$, where $\varsigma>1$. Therefore, for $\tau\in(0,1)$, conditional on $\{\boldsymbol{\chi}_i\}$, $\{\boldsymbol{X}_{NT,t}(\tau),\, t=1,\ldots,T\}$ fulfills Assumption 2 of de Jong and Davidson (2000) (by taking $d_{nt} = c_{nt} = (Th)^{-1/2}w_t^k(\tau)$ and $r = p_0/2$). Finally, Assumption B1 ensures that ℓ satisfies their Assumption 3. By employing (D.19) and applying Theorem 2.1 of de Jong and Davidson (2000), we can conclude that

$$\mathbb{E}_{\boldsymbol{w}}^{*}\left(\sum_{t=1}^{T}\widetilde{\mathcal{Y}}_{NT,t}^{*}(\tau)\right)^{2} = \boldsymbol{a}'\left\{\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}_{\boldsymbol{\chi}}\left[\boldsymbol{X}_{NT,t}(\tau)\boldsymbol{X}_{NT,s}(\tau)'\right] + o_p(1)\right\}\boldsymbol{a} + o_p(1), \qquad \text{(D.22)}$$

for $\tau\in(0,1)$. Here, we use the fact that if a random sequence is conditionally $o_p(1)$ then it is unconditionally $o_p(1)$, see, e.g., Xiong and Li (2008, Theorem 3.3). We next prove that

$$\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}_{\boldsymbol{\chi}}\left[\boldsymbol{X}_{NT,t}(\tau)\boldsymbol{X}_{NT,s}(\tau)'\right] = \mathbb{E}\left[\boldsymbol{\mathcal{E}}_{NT}(\tau)\right] + o_p(1), \qquad \forall\tau\in(0,1). \qquad \text{(D.23)}$$

where $\boldsymbol{\mathcal{E}}_{NT}(\tau)$ is defined in Lemma C.1. Eq. (D.23) can be obtained if the following two terms are asymptotically negligible for any $\tau \in (0,1)$:

$$\boldsymbol{\mathcal{K}}_{NT,1}(\tau) = \frac{1}{N}\sum_{i,j=1}^{N} K_{T,ij}(\tau)\left[\boldsymbol{\chi}_i\boldsymbol{\chi}_j' - \mathbb{E}\left(\boldsymbol{\chi}_i\boldsymbol{\chi}_j'\right)\right], \quad \boldsymbol{\mathcal{K}}_{NT,2}(\tau) = \frac{1}{N}\sum_{i,j=1}^{N} K_{T,ij}(\tau)\left[\boldsymbol{\chi}_i - \mathbb{E}\left(\boldsymbol{\chi}_i\right)\right] = \frac{1}{N}\sum_{i,j=1}^{N} K_{ij}(\tau)\boldsymbol{\chi}_i,$$

where $K_{T,ij}(\tau) = (Th)^{-1}\sum_{t,s=1}^{T} w_t^k(\tau)\, w_s^k(\tau)\,\sigma_i(\tau_t)\sigma_j(\tau_s)\mathbb{E}\left(M_{it}M_{js}\right)\mathbb{E}\left(\varepsilon_{it}\varepsilon_{js}\right)$. By Assumptions A7($c$) and B2, we have $\sup_{\tau\in[0,1]}\sum_{i,j=1}^{N}|K_{T,ij}(\tau)| = O(N)$ and $\sup_{\tau\in[0,1]}\sum_{i,j,m=1}^{N} K_{T,ij}(\tau)K_{T,im}(\tau) = o(N^2)$. Note that $\mathbb{E}\left[\boldsymbol{\mathcal{K}}_{NT,1}(\tau)\right] = \mathbf{O}$, $\mathbb{E}\left[\boldsymbol{\mathcal{K}}_{NT,2}(\tau)\right] = \mathbf{0}$, and $\max_{1\le i,j\le N}\sup_{\tau\in[0,1]}|K_{T,ij}(\tau)| \le C$. Moreover, since

$$\begin{aligned}
\mathbb{E}\left[\boldsymbol{\mathcal{K}}_{NT,1}(\tau)\boldsymbol{\mathcal{K}}_{NT,1}(\tau)'\right] &= \frac{1}{N^2}\sum_{i,j,m,n=1}^{N} K_{T,ij}(\tau)K_{T,mn}(\tau)\mathbb{E}\Big\{\left[\boldsymbol{\chi}_i\boldsymbol{\chi}_j' - \mathbb{E}\left(\boldsymbol{\chi}_i\boldsymbol{\chi}_j'\right)\right]\left[\boldsymbol{\chi}_m\boldsymbol{\chi}_n' - \mathbb{E}\left(\boldsymbol{\chi}_m\boldsymbol{\chi}_n'\right)\right]\Big\} \\
&= \begin{cases} O(N^{-1}), & i=j=m=n, \\ \mathbf{O}, & i=j=m\neq n, \text{ or } i=j\neq m=n, \text{ or } i\neq j\neq m\neq n, \\ O(N^{-1}), & i=m\neq j=n, \end{cases}
\end{aligned}$$

and $\mathbb{E}\left[\boldsymbol{\mathcal{K}}_{NT,2}(\tau)\boldsymbol{\mathcal{K}}_{NT,2}(\tau)'\right] = \boldsymbol{\Sigma}_\chi N^{-2}\sum_{i,j,m=1}^{N} K_{T,ij}(\tau)K_{T,im}(\tau) = o(1)$. Therefore, we deduce that $\boldsymbol{\mathcal{K}}_{NT,1}(\tau) = o_p(1)$ and $\boldsymbol{\mathcal{K}}_{NT,2}(\tau) = o_p(1)$, leading to result of Eq. (D.23). By (D.22), (D.23) and Lemma C.1, we have

$$\mathbb{E}_{\boldsymbol{w}}^*\left(\sum_{t=1}^{T}\widetilde{\mathcal{Y}}_{NT,t}^*(\tau)\right)^2 \xrightarrow{p} \boldsymbol{a}'\nu_{2k}\,\mathrm{diag}\left[\boldsymbol{\Lambda}_\varepsilon(\tau), \bar{\Lambda}_\varepsilon(\tau)\boldsymbol{\Sigma}_\chi, \boldsymbol{\Lambda}_{\nu\varepsilon}(\tau)\right]\boldsymbol{a}. \tag{D.24}$$

II. Asymptotic normality of $\sum_{t=1}^{T}\widetilde{\mathcal{Y}}_{NT,t}^*(\tau)$

Let $\widetilde{Q}_{N,t}(\tau) = N^{-1/2}w_t^k(\tau)\sum_{i=1}^{N} M_{it}\boldsymbol{a}'\left[1/p_i(\tau), 1, \boldsymbol{\chi}_i', \boldsymbol{\nu}_{it}'\right]'\widetilde{e}_{it}$ so that $\sum_{t=1}^{T}\widetilde{\mathcal{Y}}_{NT,t}^*(\tau) = (Th)^{-1/2}\sum_{t=1}^{T}\widetilde{Q}_{N,t}(\tau)\xi_t^*$. Note that $\{\xi_t^*\}$ admits an MA(∞) representation $\xi_t^* = \sum_{j=0}^{\infty}\gamma^j\nu_{t-j}^*$, where $\{\nu_t^*,\, t<1\}$ is defined similarly to $\{\nu_t^*,\, t\ge 1\}$. Choose an M such that $M/\ell \to \infty$ as $(N,T)\to\infty$, and truncate the bootstrap errors ξ_t^* at M lags, denoted by $\xi_{t,M}^* = \sum_{j=0}^{M}\gamma^j\nu_{t-j}^*$. Then

$$\sum_{t=1}^{T}\widetilde{\mathcal{Y}}_{NT,t}^*(\tau) = \frac{1}{\sqrt{Th}}\sum_{t=1}^{T}\widetilde{Q}_{N,t}(\tau)\xi_{t,M}^* + \frac{1}{\sqrt{Th}}\sum_{t=1}^{T}\widetilde{Q}_{N,t}(\tau)\left(\sum_{j=M+1}^{\infty}\gamma^j\nu_{t-j}^*\right). \tag{D.25}$$

Note that the second component in (D.25) is asymptotically negligible conditional on $\{\boldsymbol{w}_{it}\}$. To see this, observe that, for any $s\ge 0$ and $\theta\in(0,1)$,

$$\mathbb{E}_{\boldsymbol{w}}^*\left[\left(\sum_{j=M+1}^{\infty}\gamma^j\nu_{t-j}^*\right)\left(\sum_{j=M+1}^{\infty}\gamma^j\nu_{t+s-j}^*\right)\right] = \gamma^s(1-\gamma^2)\sum_{j=M+1}^{\infty}\gamma^{2j} = \gamma^s\theta^{2(M+1)/\ell} = o(\gamma^s). \tag{D.26}$$

By replacing γ^s with $o(\gamma^s)$ in (D.3) in Part I, it is immediate to obtain

$$\mathbb{E}_{\boldsymbol{w}}^* \left[\frac{1}{\sqrt{Th}} \sum_{t=1}^{T} \widetilde{Q}_{N,t}(\tau) \left(\sum_{j=M+1}^{\infty} \gamma^j \nu_{t-j}^* \right) \right]^2 = o_p(1), \qquad \forall \tau \in (0,1). \tag{D.27}$$

We now adopt the common blocking technique as in Friedrich et al. (2020, Proof of Theorem 2) and Friedrich and Lin (2022, Proof of Theorem 1) to establish asymptotic normality of $(Th)^{-1/2}\sum_{t=1}^{T}\widetilde{Q}_{N,t}(\tau)\xi_{t,M}^*$. We partition the index set $\{1,\dots,T\} = \cup_{j=1}^{k_T} B_j$, where $B_j = \{b_j+1,\dots,b_j+l_T\} \cup \{b_j+l_T+1,\dots,b_{j+1}\}$, $b_j = (j-1)(l_T+s_T)$, and $k_T = \lceil T/(l_T+s_T)\rceil$. We truncate the final block B_{k_T} to have T observations in total. Moreover, we require $1/l_T + l_T/(Th) \to 0$ and $1/s_T + s_T/l_T + M/s_T \to 0$ as $T \to \infty$. For instance, one can take $l_T = \lfloor Th^2 \rfloor$, $s_T = \lfloor (Th^2)^{1/2} \rfloor$, and $M = \lfloor (Th^2)^{1/4} \rfloor$. Then, we have $k_T \sim T/l_T$. By construction, each block B_j is divided into two subsets, one with a relatively large length (l_T) and the other with a small length (s_T). It leads to $(Th)^{-1/2}\sum_{t=1}^{T}\widetilde{Q}_{N,t}(\tau)\xi_{t,M}^* = \sum_{j=1}^{k_T} V_{N,T,j}^*(\tau) + \sum_{j=1}^{k_T} W_{N,T,j}^*(\tau)$, where

$$V_{N,T,j}^*(\tau) = \frac{1}{\sqrt{Th}} \sum_{t=b_j+1}^{b_j+l_T} \widetilde{Q}_{N,t}(\tau)\xi_{t,M}^*, \qquad W_{N,T,j}^*(\tau) = \frac{1}{\sqrt{Th}} \sum_{t=b_j+l_T+1}^{b_{j+1}} \widetilde{Q}_{N,t}(\tau)\xi_{t,M}^*. \tag{D.28}$$

We first show that $\sum_{j=1}^{k_T} W_{N,T,j}^*(\tau)$ is asymptotically negligible conditionally on $\{\boldsymbol{w}_{it}\}$. Note that $\left\{\xi_{t,M}^*\right\}$ is an M-dependent process conditionally on $\{\boldsymbol{w}_{it}\}$. As such, the blocks $W_{N,T,j}^*(\tau)$, $j = 1,\dots,k_T$, are conditionally independent for a sufficiently large T. Moreover, recall $M_{it}(\widetilde{e}_{it} - e_{it}) = M_{it}\big[(\alpha_i - \widetilde{\alpha}_i) + \big(1, \boldsymbol{x}_{it}'\big)(\boldsymbol{\Upsilon}_t - \widetilde{\boldsymbol{\Upsilon}}_t)\big]$. Then we can write $\widetilde{Q}_{N,t}(\tau) = Q_{N,t}(\tau) + R_{N,t}(\tau)$, where $Q_{N,t}(\tau) = N^{-1/2} w_t^k(\tau) \sum_{i=1}^{N} M_{it}\boldsymbol{a}' \left[1/p_i(\tau), 1, \boldsymbol{\chi}_i', \boldsymbol{\nu}_{it}'\right]' e_{it}$, and

$$\begin{aligned}
R_{N,t}(\tau) &= w_t^k(\tau) N^{-1/2} \sum_{i=1}^{N} \boldsymbol{a}' \left[1/p_i(\tau), 1, \boldsymbol{\chi}_i', \boldsymbol{\nu}_{it}'\right]' M_{it}(\widetilde{e}_{it} - e_{it}) \\
&\le w_t^k(\tau) N^{-1/2} \sum_{i=1}^{N} \left(C + \|\boldsymbol{\chi}_i\| + \|\boldsymbol{\nu}_{it}\|\right) \|M_{it}(\widetilde{e}_{it} - e_{it})\| \\
&\le C w_t^k(\tau) \left\{ \max_{1\le j\le N} |\alpha_j - \widetilde{\alpha}_j| + \sup_{\tau\in[0,1]} \left\|\widetilde{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}(\tau)\right\| \right\} \left\{ O_p(\sqrt{N}) + N^{-1/2} \sum_{i=1}^{N} \|\boldsymbol{\nu}_{it}\|^2 \right\},
\end{aligned} \tag{D.29}$$

where the term $O_p(\sqrt{N})$ is uniform in τ and t. By utilizing the results above and the identity $\left(\sum_{t=L}^{U} q_t\right)^2 =$

$\sum_{s=-(U-L)}^{U-L} \sum_{t=L}^{U-|s|} q_t q_{t+|s|}$, we obtain

$$\begin{aligned}
\mathbb{E}^*_{\boldsymbol{w}}\left(\sum_{j=1}^{k_T} W^*_{N,T,j}(\tau)\right)^2 &= \sum_{j=1}^{k_T} \mathbb{E}^*_{\boldsymbol{w}}\left(W^*_{N,T,j}(\tau)\right)^2 \\
&\leq \frac{1}{Th} \sum_{s=-s_T+1}^{s_T-1} \theta^{|s|/\ell} \sum_{j=1}^{k_T} \sum_{t=b_j+l_T+1}^{b_{j+1}-|s|} \left[\widetilde{Q}_{N,t}(\tau)\widetilde{Q}_{N,t+|s|}(\tau)\right] \\
&= \left[\frac{1}{Th} \sum_{s=-s_T+1}^{s_T-1} \theta^{|s|/\ell} \sum_{j=1}^{k_T} \sum_{t=b_j+l_T+1}^{b_{j+1}-|s|} Q_{N,t}(\tau)Q_{N,t+|s|}(\tau)\right] + o_p\left(\frac{s_T}{l_T}\right), \quad \text{(D.30)}
\end{aligned}$$

where the second step follows from

$$\mathbb{E}^*_{\boldsymbol{w}}\left(\xi^*_{t,M}\xi^*_{t+|s|,M}\right) = \gamma^{|s|}(1-\gamma^2)\sum_{j=0}^{M}\gamma^{2j} = \gamma^{|s|}\left(1-\gamma^{2(M+1)}\right) \leq \theta^{|s|/\ell}, \qquad \theta \in (0,1), \tag{D.31}$$

and the third step is due to (D.29) and the assumption $\sqrt{N}O(\ell)\left\{\max_{1\leq j\leq N}|\alpha_j - \widetilde{\alpha}_j| + \sup_{\tau\in[0,1]}\left\|\widetilde{\boldsymbol{\theta}}(\tau) - \boldsymbol{\theta}(\tau)\right\|\right\} = o_p(1)$. Since, by Assumption A7(d),

$$\begin{aligned}
\mathbb{E}\left(\frac{1}{Th}\sum_{s=-s_T+1}^{s_T-1}\theta^{|s|/\ell}\sum_{j=1}^{k_T}\sum_{t=b_j+l_T+1}^{b_{j+1}-|s|} Q_{N,t}(\tau)Q_{N,t+|s|}(\tau)\right) &\leq C\frac{k_T s_T}{T}\boldsymbol{a}'\Bigg\{N^{-1}\sum_{i,j=1}^{N}\mathcal{R}_{i,j}(\tau,\tau)\sigma_i(\tau)\sigma_j(\tau) \\
&\times \sum_{s=-s_T+1}^{s_T-1}\mathbb{E}\Big[\boldsymbol{A}(i,t;j,t+s)\varepsilon_{it}\varepsilon_{j(t+s)}\Big]\Bigg\}\boldsymbol{a} + o\left(\frac{s_T}{l_T}\right) = O\left(\frac{s_T}{l_T}\right), \quad \text{(D.32)}
\end{aligned}$$

where $\boldsymbol{A}(i,t;j,t+s)$ is a block matrix, and the $(k_1,k_2)_{th}$ block is given by $\boldsymbol{A}_{k_1k_2}(i,t;j,t+s)$. By combining (D.30) and (D.32), it follows that

$$\mathbb{E}^*_{\boldsymbol{w}}\left(\sum_{j=1}^{k_T} W^*_{N,T,j}(\tau)\right)^2 = \left(\frac{s_T}{l_T}\right) = o_p(1), \qquad \forall\tau\in(0,1). \tag{D.33}$$

Since $\mathbb{E}^*_{\boldsymbol{w}}\left(\sum_{j=1}^{k_T} W^*_{N,T,j}(\tau)\right) = 0$, we conclude that $\sum_{j=1}^{k_T} W^*_{N,T,j}(\tau)$ converges to zero in probability conditionally on $\{\boldsymbol{w}_{it}\}$.

Recall $V^*_{N,T,j}(\tau) = (Th)^{-1/2}\sum_{t=b_j+l_T+1}^{b_{j+1}} \widetilde{Q}_{N,t}(\tau)\xi^*_{t,M}$, $j = 1,\ldots,k_T$. To establish the asymptotic normality of $(Th)^{-1/2}\sum_{t=1}^{T}\widetilde{Q}_{N,t}(\tau)\xi^*_{t,M}$, it essentially involves proving the asymptotic normality of $\sum_{j=1}^{k_T} V^*_{N,T,j}(\tau)$. To this end, we will employ the Lindeberg central limit theorem (Davidson, 1994, Theorem 23.6). Note that $\mathbb{E}^*_{\boldsymbol{w}}\left(\sum_{j=1}^{k_T} V^*_{N,T,j}(\tau)\right) = 0$, and $\sum_{j=1}^{k_T} V^*_{N,T,j}(\tau)$ dominates the asymptotic behavior of $\sum_{t=1}^{T}\widetilde{\mathcal{Y}}^*_{NT,t}(\tau)$. By (D.19), (D.26), and (D.33), it becomes evident that $\mathbb{E}^*_{\boldsymbol{w}}\left(\sum_{j=1}^{k_T} V^*_{N,T,j}(\tau)\right)^2 \xrightarrow{p} \boldsymbol{a}'\nu_{2k}\,\mathrm{diag}\left[\boldsymbol{\Lambda}_\varepsilon(\tau), \bar{\Lambda}_\varepsilon(\tau)\boldsymbol{\Sigma}_\chi, \boldsymbol{\Lambda}_{\nu\varepsilon}(\tau)\right]\boldsymbol{a}$ for any $\tau\in(0,1)$. The final step is to verify the Lindeberg condition. That is, for any $\kappa > 0$, we shall show

$$\sum_{j=1}^{k_T} \mathbb{E}_{\boldsymbol{w}}^* \left(\frac{\left[V_{N,T,j}^*(\tau)\right]^2}{\omega_{NT}^{*2}} \mathbb{1}\left\{ \left| \frac{V_{N,T,j}^*(\tau)}{\omega_{NT}^*} \right| > \kappa \right\} \right) = o_p(1), \qquad \forall \tau \in (0,1), \tag{D.34}$$

where $\omega_{NT}^{*2} = \mathbb{E}_{\boldsymbol{w}}^* \left(\sum_{j=1}^{k_T} V_{N,T,j}^*(\tau) \right)^2$. Recall that $p_0 = 2(2+\delta)$ as indicated in Assumption A4, where $\delta > 0$. Given $\mathbb{E}_{\boldsymbol{w}}^*(\nu_i^*)^{p_0/2} < \infty$, $\widetilde{Q}_{N,t}(\tau)\xi_{t,M}^*$ forms an $L_{p_0/2}$-mixingale, conditional on $\{\boldsymbol{w}_{it}\}$, when $c_i = |\widetilde{Q}_{N,i}(\tau)| \left(\mathbb{E}_{\boldsymbol{w}}^*(\nu_i^*)^{p_0/2}\right)^{2/p_0}$ is taken in Definition 1 in Hansen (1991) (Davidson, 1994, Example 16.2, for instance). Since $\widetilde{Q}_{N,t}(\tau) = Q_{N,t}(\tau) + R_{N,t}(\tau)$, using Lemma 2 of Hansen (1991), the LHS of (D.34) can be bounded by

$$\begin{aligned}
&\frac{1}{\kappa^\delta} \sum_{j=1}^{k_T} \mathbb{E}_{\boldsymbol{w}}^* \left(\frac{\left[V_{N,T,j}^*(\tau)\right]^{p_0/2}}{\omega_{NT}^{*(p_0/2)}} \mathbb{1}\left\{ \left| \frac{V_{N,T,j}^*(\tau)}{\omega_{NT}^*} \right| > \kappa \right\} \right) \le \frac{1}{\kappa^\delta} \omega_{NT}^{*-(p_0/2)} \sum_{j=1}^{k_T} \mathbb{E}_{\boldsymbol{w}}^* \left(\left[V_{N,T,j}^*(\tau)\right]^{p_0/2} \right) \\
&\quad \le C \frac{1}{\kappa^\delta} \omega_{NT}^{*-(p_0/2)} \frac{1}{(Th)^{1+\delta/2}} \sum_{j=1}^{k_T} \left(\sum_{t=b_j+l_T+1}^{b_{j+1}} \left[\widetilde{Q}_{N,i}(\tau)\right]^2 \right)^{1+\delta/2} \\
&\quad \le C \frac{1}{\kappa^\delta} \omega_{NT}^{*-(p_0/2)} \frac{l_T^{\delta/2}}{(Th)^{1+\delta/2}} \sum_{j=1}^{k_T} \sum_{t=b_j+l_T+1}^{b_{j+1}} \left[\left|Q_{N,t}(\tau)\right|^{p_0/2} + \left|R_{N,t}(\tau)\right|^{p_0/2} \right] \\
&\quad \le C \frac{1}{\kappa^\delta} \omega_{NT}^{*-(p_0/2)} \left(\frac{l_T}{Th}\right)^{\delta/2} \frac{1}{Th} \sum_{t=1}^{T} \left[\left|Q_{N,t}(\tau)\right|^{p_0/2} + \left|R_{N,t}(\tau)\right|^{p_0/2} \right] \\
&\quad = O_p(1) \left(\frac{l_T}{Th}\right)^{\delta/2} \left\{ O_p\left(\frac{1}{Th} \sum_{t=1}^{T} \left[w_t^k(\tau)\right]^{p_0/2} \right) + o_p\left(\frac{1}{Th} \sum_{t=1}^{T} \left[w_t^k(\tau)\right]^{p_0/2} \right) \right\},
\end{aligned}$$

where the c_r-inequality is applied twice in the third step, while the final step is a result of Assumption A7(*a*), Eq. (D.29), and the fact that $\omega_{NT}^{*-4} = O_p(1)$. Since $l_T/(Th) \to 0$ as $T \to \infty$, we obtain (D.34).

III. Asymptotic order of the second component in (D.2)

Similar to the proof for $\widetilde{\mathrm{II}}_{NT,k_1,k_2}^b(\tau)$ in Part I, one can obtain

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{w}}^* \left| \frac{1}{\sqrt{NTh}} \sum_{i=1}^{N} \sum_{t=1}^{T} w_t^k(\tau) M_{it} \left\{ (Th\nu_{\tau,i}) - [p_i(\tau)]^{-1} \right\} \xi_t^* \widetilde{e}_{it} \right|^2 \\
&= \frac{1}{NTh} \sum_{i,j=1}^{N} \left\{ (Th\nu_{\tau,i}) - [p_i(\tau)]^{-1} \right\} \left\{ (Th\nu_{\tau,j}) - [p_j(\tau)]^{-1} \right\} \sum_{s,t=1}^{T} w_t^k(\tau) w_s^k(\tau) M_{it} M_{js} \widetilde{e}_{it} \widetilde{e}_{js} \gamma^{|t-s|} \\
&= \frac{1}{NTh} \sum_{i,j=1}^{N} \left\{ (Th\nu_{\tau,i}) - [p_i(\tau)]^{-1} \right\} \left\{ (Th\nu_{\tau,j}) - [p_j(\tau)]^{-1} \right\} \sum_{s,t=1}^{T} w_t^k(\tau) w_s^k(\tau) M_{it} M_{js} e_{it} e_{js} \gamma^{|t-s|} + o_p(1) \\
&= o_p(1),
\end{aligned}$$

where the final step follows from the arguments for Part (*b*) in the proof of Lemma C.1. Using Markov's inequality, the component $(NTh)^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} w_t^k(\tau) M_{it} \left\{ (Th\nu_{\tau,i}) - [p_i(\tau)]^{-1} \right\} \xi_t^* \widetilde{e}_{it}$ is asymptotically negligible.

The proof of the lemma is now complete. ■

Proposition D.1 *Let $\boldsymbol{\mathcal{D}}_{NT}^*(\tau)$ be the bootstrap counterpart of $\boldsymbol{\mathcal{D}}_{NT}(\tau)$ as defined in Eq. (C.23) in Proposition C.3. That is,*

$$\boldsymbol{\mathcal{D}}_{NT}^*(\tau) = \sum_{i=1}^{N} \boldsymbol{Z}_i^M(\tau)'\boldsymbol{K}_h(\tau)\boldsymbol{e}_i^* - \sum_{i=1}^{N} \nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\boldsymbol{k}_h(\tau)'\left(\boldsymbol{e}_i^* - \omega_\tau^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{e}_i^*\right). \tag{D.35}$$

Under Assumptions A1 - A8, B1 - B2, for any fixed $\tau \in (0,1)$, $(NTh)^{-1/2}\boldsymbol{\mathcal{D}}_{NT}^*(\tau)$ and $(NTh)^{-1/2}\boldsymbol{\mathcal{D}}_{NT}(\tau)$ share the same limiting distribution. Namely, for any fixed $\tau \in (0,1)$,

$$\frac{1}{\sqrt{NTh}}\boldsymbol{\mathcal{D}}_{NT}^*(\tau) \xrightarrow{d^*}_p \mathcal{N}\Big(\boldsymbol{0}, \boldsymbol{\Phi}_{\nu\varepsilon}(\tau)\Big), \qquad (N,T)\to\infty. \tag{D.36}$$

Proof of Proposition D.1 Recall $e_{it}^* = M_{it}\xi_t^*\widetilde{e}_{it}$ in Section 4.1 (Step 3). As in Proof of Proposition C.3, we split $\boldsymbol{\mathcal{D}}_{NT}^*(\tau)$ into four main blocks of vectors:

$$\begin{aligned}\boldsymbol{\mathcal{D}}_{NT}^*(\tau) &= \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{z}_{it}(\tau)\xi_t^*\widetilde{e}_{it} - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{z}_{it}(\tau)\right)\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\xi_t^*\widetilde{e}_{it}\right)\\ &+ \omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{z}_{it}(\tau)\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\xi_t^*\widetilde{e}_{it}\right) =: \begin{pmatrix}\boldsymbol{\Upsilon}_{NT,1}^*(\tau)\\ \vdots \\ \boldsymbol{\Upsilon}_{NT,4}^*(\tau)\end{pmatrix},\end{aligned} \tag{D.37}$$

where

$$\begin{aligned}\boldsymbol{\Upsilon}_{NT,1}^*(\tau) &= \omega_\tau^{-1}N\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\xi_t^*\widetilde{e}_{it}\right),\\ \boldsymbol{\Upsilon}_{NT,2}^*(\tau) &= \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\xi_t^*\widetilde{e}_{it} - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\xi_t^*\widetilde{e}_{it}\right)\\ &\quad + \omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\xi_t^*\widetilde{e}_{it}\right),\end{aligned}$$

and

$$
\begin{aligned}
\boldsymbol{\Upsilon}_{NT,3}^{*}(\tau) &= \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau)M_{it}\xi_t^*\widetilde{e}_{it} - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^1(\tau)M_{it}\right)\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\xi_t^*\widetilde{e}_{it}\right) \\
&\quad + \omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^1(\tau)M_{it}\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\xi_t^*\widetilde{e}_{it}\right), \\
\boldsymbol{\Upsilon}_{NT,4}^{*}(\tau) &= \sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\xi_t^*\widetilde{e}_{it} - \sum_{i=1}^{N}\nu_{\tau,i}\left(\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{t=1}^{T} w_t^0(\tau)M_{it}\xi_t^*\widetilde{e}_{it}\right) \\
&\quad + \omega_\tau^{-1}\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\right)\left(\sum_{i=1}^{N}\nu_{\tau,i}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\xi_t^*\widetilde{e}_{it}\right).
\end{aligned}
$$

Applying similar arguments as in Proposition C.3, we obtain

$$
\begin{aligned}
&\frac{1}{\sqrt{NTh}}\boldsymbol{\mathcal{D}}_{NT}^{*}(\tau) \\
&= \begin{pmatrix}
\left(\frac{Th}{N}\omega_\tau\right)^{-1}\left(\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}(Th\nu_{\tau,i})\xi_t^*\widetilde{e}_{it}\right) \\
\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}\boldsymbol{\nu}_{it}\xi_t^*\widetilde{e}_{it} + \boldsymbol{\ell}(\tau)\left(\frac{Th}{N}\omega_\tau\right)^{-1}\left(\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^0(\tau)M_{it}(Th\nu_{\tau,i})\xi_t^*\widetilde{e}_{it}\right) \\
\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau)M_{it}\xi_t^*\widetilde{e}_{it} \\
\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\sum_{t=1}^{T} w_t^1(\tau)M_{it}\boldsymbol{x}_{it}\xi_t^*\widetilde{e}_{it}
\end{pmatrix} + o_p^*(1),
\end{aligned}
$$

where the $o_p^*(1)$ term follows from utilizing (D.6) and employing similar steps in Lemma D.1. We then obtain (D.36) by Lemma D.1. ■

Proof of Theorem 3 As in Lemma 6 of Friedrich and Lin (2022), one can write

$$
\begin{aligned}
\sqrt{NTh}\left(\widehat{\boldsymbol{\theta}}^*(\tau) - \widetilde{\boldsymbol{\theta}}(\tau) - h^2\boldsymbol{b}(\tau)\right) &= \sqrt{NTh}\Big[\widehat{\boldsymbol{\theta}}^*(\tau) - \mathbb{E}^*\big(\widehat{\boldsymbol{\theta}}^*(\tau)\big)\Big] + \sqrt{NTh}\Big[\mathbb{E}^*\big(\widehat{\boldsymbol{\theta}}^*(\tau)\big) - \widetilde{\boldsymbol{\theta}}(\tau) - h^2\boldsymbol{b}(\tau)\Big] \\
&=: \boldsymbol{M}_{NT}^*(\tau) + \boldsymbol{R}_{NT}^*(\tau). \qquad \text{(D.38)}
\end{aligned}
$$

We shall prove: (i) $\boldsymbol{M}_{NT}^*(\tau)$ replicates the asymptotic distribution of the LLDV estimator, as established in Theorem 2; (ii) the remainder term $\boldsymbol{R}_{NT}^*(\tau)$ is asymptotically negligible for any $\tau \in [0,1]$.

For $\boldsymbol{M}_{NT}^*(\tau)$, by Proposition C.1, we obtain

$$
\boldsymbol{M}_{NT}^*(\tau) = \left(\frac{1}{NTh}\sum_{i=1}^{N}\widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{Z}}_i^M(\tau)\right)^{-1}\left(\frac{1}{\sqrt{NTh}}\sum_{i=1}^{N}\widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{e}}_i^*\right) = \left[\boldsymbol{\Phi}(\tau)^{-1} + o_p(1)\right]\boldsymbol{\mathcal{D}}_{NT}^*(\tau), \quad \text{(D.39)}
$$

where $\widetilde{\boldsymbol{e}}_i^* = \boldsymbol{K}_h^{1/2}(\tau)\boldsymbol{e}_i^* - \mathrm{diag}(\boldsymbol{m}_i)\boldsymbol{k}_h^{1/2}(\tau)\boldsymbol{k}_h(\tau)'\nu_{\tau,i}\left(\boldsymbol{e}_i^* - \omega_\tau^{-1}\sum_{i=1}^{N}\nu_{\tau,i}\boldsymbol{e}_i^*\right)$, $\boldsymbol{e}_i^* = (e_{i1}^*, \ldots, e_{iT}^*)'$, and $\boldsymbol{\mathcal{D}}_{NT}^*(\tau)$ is defined in Eq. (D.35). By Proposition D.1, $\boldsymbol{\mathcal{D}}_{NT}^*(\tau)$ shares the same limiting distribution of $\boldsymbol{\mathcal{D}}_{NT}(\tau)$. Overall, for any $\tau \in (0,1)$, we have

$$
\boldsymbol{M}_{NT}^*(\tau) \xrightarrow{d^*}_p \mathcal{N}\Big(\boldsymbol{0}, \boldsymbol{\Phi}(\tau)^{-1}\boldsymbol{\Phi}_{\nu\varepsilon}(\tau)\boldsymbol{\Phi}(\tau)^{-1}\Big), \qquad (N,T) \to \infty. \qquad \text{(D.40)}
$$

Next, we consider $\boldsymbol{R}_{NT}^*(\tau)$. Note that

$$
\begin{aligned}
\mathbb{E}^*\left(\widehat{\boldsymbol{\theta}}^*(\tau)\right) &= \left(\sum_{i=1}^N \widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{Z}}_i^M(\tau)\right)^{-1} \\
&\times \left[\sum_{i=1}^N \boldsymbol{Z}_i^M(\tau)'\boldsymbol{K}_h(\tau)\operatorname{diag}(\boldsymbol{m}_i)\widetilde{\boldsymbol{b}}_i - \sum_{i=1}^N \nu_{\tau,i}\boldsymbol{Z}_i^M(\tau)'\boldsymbol{k}_h(\tau)\boldsymbol{k}_h(\tau)'\operatorname{diag}(\boldsymbol{m}_i)\left(\widetilde{\boldsymbol{b}}_i - \omega_\tau^{-1}\sum_{i=1}^N \nu_{\tau,i}\widetilde{\boldsymbol{b}}_i\right)\right],
\end{aligned} \tag{D.41}
$$

where $\widetilde{\boldsymbol{b}}_i = (\widetilde{g}_1,\ldots,\widetilde{g}_T)' + \left(\boldsymbol{x}_{i1}'\widetilde{\boldsymbol{\beta}}_1,\ldots,\boldsymbol{x}_{iT}'\widetilde{\boldsymbol{\beta}}_T\right)'$. Now, let's rewrite the terms in the square brackets of Eq. (D.41). For $|\tau_t-\tau|\le h$, Proposition 1 implies that

$$
\begin{aligned}
\widetilde{g}(\tau_t) &= \widetilde{g}(\tau) + \widetilde{g}^{(1)}(\tau)(\tau_t-\tau) + \widetilde{g}(\tau_t) - \widetilde{g}(\tau) - \widetilde{g}^{(1)}(\tau)(\tau_t-\tau) \\
&= \widetilde{g}(\tau) + \widetilde{g}^{(1)}(\tau)(\tau_t-\tau) + \left[g(\tau_t) - g(\tau) - g^{(1)}(\tau)(\tau_t-\tau)\right] + O_p\left(\tilde{h}^2 + \sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}\tilde{h}}}\right) \\
&= \widetilde{g}(\tau) + \widetilde{g}^{(1)}(\tau)(\tau_t-\tau) + \left[\widetilde{g}^{(2)}(\tau) + O(h)\right]\frac{(\tau_t-\tau)^2}{2} + O_p\left(\tilde{h}^2 + \sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}\tilde{h}}}\right),
\end{aligned} \tag{D.42}
$$

where the O_p-terms are uniform in $\tau\in[0,1]$. Similarly, we have

$$
\widetilde{\boldsymbol{\beta}}(\tau_t) = \widetilde{\boldsymbol{\beta}}(\tau) + \widetilde{\boldsymbol{\beta}}^{(1)}(\tau)(\tau_t-\tau) + \left[\widetilde{\boldsymbol{\beta}}^{(2)}(\tau) + O(h)\right]\frac{(\tau_t-\tau)^2}{2} + O_p\left(\tilde{h}^2 + \sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}\tilde{h}}}\right). \tag{D.43}
$$

By (D.42) and (D.43), we obtain

$$
\begin{aligned}
\operatorname{diag}(\boldsymbol{m}_i)\widetilde{\boldsymbol{b}}_i &= \boldsymbol{Z}_i^M(\tau)\widetilde{\boldsymbol{\theta}}(\tau) + \frac{h^2}{2}\operatorname{diag}\left[\left(\frac{\tau_1-\tau}{h}\right)^2,\ldots,\left(\frac{\tau_T-\tau}{h}\right)^2\right] \\
&\quad \times \boldsymbol{Z}_i^M(\tau)\begin{pmatrix}\boldsymbol{\Upsilon}^{(2)}(\tau)+O(h) \\ \boldsymbol{0}_{d+1}\end{pmatrix} + O_p\left(\tilde{h}^2 + \sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}\tilde{h}}}\right) \\
&= \boldsymbol{Z}_i^M(\tau)\widetilde{\boldsymbol{\theta}}(\tau) + \boldsymbol{\Delta}_i^M(\tau) \\
&\quad + O(h^3)\operatorname{diag}\left[\left(\frac{\tau_1-\tau}{h}\right)^2,\ldots,\left(\frac{\tau_T-\tau}{h}\right)^2\right]\boldsymbol{Z}_i^M(\tau) + O_p\left(\tilde{h}^2 + \sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}\tilde{h}}}\right).
\end{aligned} \tag{D.44}
$$

where $\boldsymbol{\Upsilon}(\tau)$ and $\boldsymbol{\Delta}_i^M(\tau)$ are defined in Proposition C.2, and the $O(h^3)$ terms does not rely on $i\in\mathbb{Z}^+$ and $\tau\in[0,1]$. Putting (D.44) back to (D.41) leads to

$$
\begin{aligned}
\mathbb{E}^*\left(\widehat{\boldsymbol{\theta}}^*(\tau)\right) &= \widetilde{\boldsymbol{\theta}}(\tau) + \left(\frac{1}{NTh}\sum_{i=1}^N \widetilde{\boldsymbol{Z}}_i^M(\tau)'\widetilde{\boldsymbol{Z}}_i^M(\tau)\right)^{-1}\frac{1}{NTh}\boldsymbol{\mathcal{B}}_{NT}(\tau) + O_p(h^3) + O_p\left(\tilde{h}^2 + \sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}\tilde{h}}}\right) \\
&= \widetilde{\boldsymbol{\theta}}(\tau) + h^2\boldsymbol{b}(\tau) + O_p(h^3) + O_p\left(\tilde{h}^2 + \sqrt{\frac{\ln(NT)}{(NT)^{\phi(p_0/2)}\tilde{h}}}\right),
\end{aligned} \tag{D.45}
$$

where the second step is due to Propositions C.1 and C.2, and the $O_p(\cdot)$-terms are uniform in $\tau \in [0,1]$. By (D.45), we arrive at

$$\sup_{\tau\in[0,1]} \|\boldsymbol{R}^*_{NT}(\tau)\| = O_p\left(\sqrt{NTh^7}\right) + O_p\left(\sqrt{NTh\tilde{h}^4} + \sqrt{(NT)^{1-\phi(p_0/2)}\ln(NT)h/\tilde{h}}\right) = o_p(1), \quad \text{(D.46)}$$

under Assumption B3. The results from Eqs. (D.38), (D.40), and (D.46) jointly imply Eq. (4.6). ■

E Bandwidth selection

As previously observed in the literature, pointwise confidence intervals are not sensitive to bandwidth selection. This is in contrast to simultaneous confidence bands, which can be highly sensitive. The simultaneous coverage depends heavily on the local behavior of the parameter curves. For instance, if a curve exhibits large fluctuations in certain neighborhoods, a smaller bandwidth is typically needed to achieve the targeted coverage in those regions. Conversely, a flat region may only require a relatively large bandwidth. For additional insights, refer to the simulation study conducted in Friedrich and Lin (2022). Our initial investigation indicates that the commonly used leave-one-unit-out cross-validation method developed in Sun et al. (2009) tends to select an overly small bandwidth in our specific context. This might be attributed to the fact that it is specifically designed for panel data models that are cross-sectionally independent and without missing observations.

To maintain the structure of cross-sectional dependence and capture local features of parameter curves, we suggest extending the local modified-cross-validation procedure introduced by Friedrich and Lin (2022). This approach combines the modified cross-validation method proposed by Chu and Marron (1991) with the local cross-validation approach suggested by Vieu (1991), exhibiting superior performance as demonstrated in Friedrich and Lin (2022). More specifically, for a given $h > 0$, let $\left(\widehat{g}^{l,h}(\tau), \widehat{\boldsymbol{\beta}}^{l,h}(\tau)\right)$ denote the leave-$(2l+1)$-out LLDV estimator, which is constructed as in Eq. (2.10) but omitting all observations between time $\lfloor \tau T\rfloor - l$ and $\lfloor \tau T\rfloor + l$, namely $\left\{(y_{it}, \boldsymbol{x}_{it}),\ i = 1,\ldots,N,\ \lfloor \tau T\rfloor - l \le t \le \lfloor \tau T\rfloor + l\right\}$. For each $\tau \in (0,1)$ and $l \ge 0$, the locally optimal bandwidth $\hat{h}(\tau, l)$ minimizes the panel local modified cross-validation (PLMCV) criterion $\text{PLMCV}_{\tau,l}(h)$, $h \in [h_L, h_U] \subset (0,\infty)$, where

$$\text{PLMCV}_{\tau,l}(h) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} M_{it}\Big(y_{it} - \widehat{\alpha}_i - \widehat{g}^{l,h}(t/T) - \boldsymbol{x}_{it}'\widehat{\boldsymbol{\beta}}^{l,h}(t/T)\Big)^2 \omega_\tau(t/T),$$

and $\widehat{\alpha}_i$ is obtained without excluding any observations. Moreover, $\omega_\tau(\cdot)$ is some weight function. We follow Vieu (1991) to take $\omega_\tau(\cdot)$ as the density of $\mathcal{N}(\tau, 0.025)$. The final selector sets $\hat{h}_{\text{PLMCV}} = (TK)^{-1}\sum_{k=1}^{K}\sum_{t=1}^{T}\hat{h}(\tau_t, l_k)$, $l_k \in \{0,2,4,6\}$.

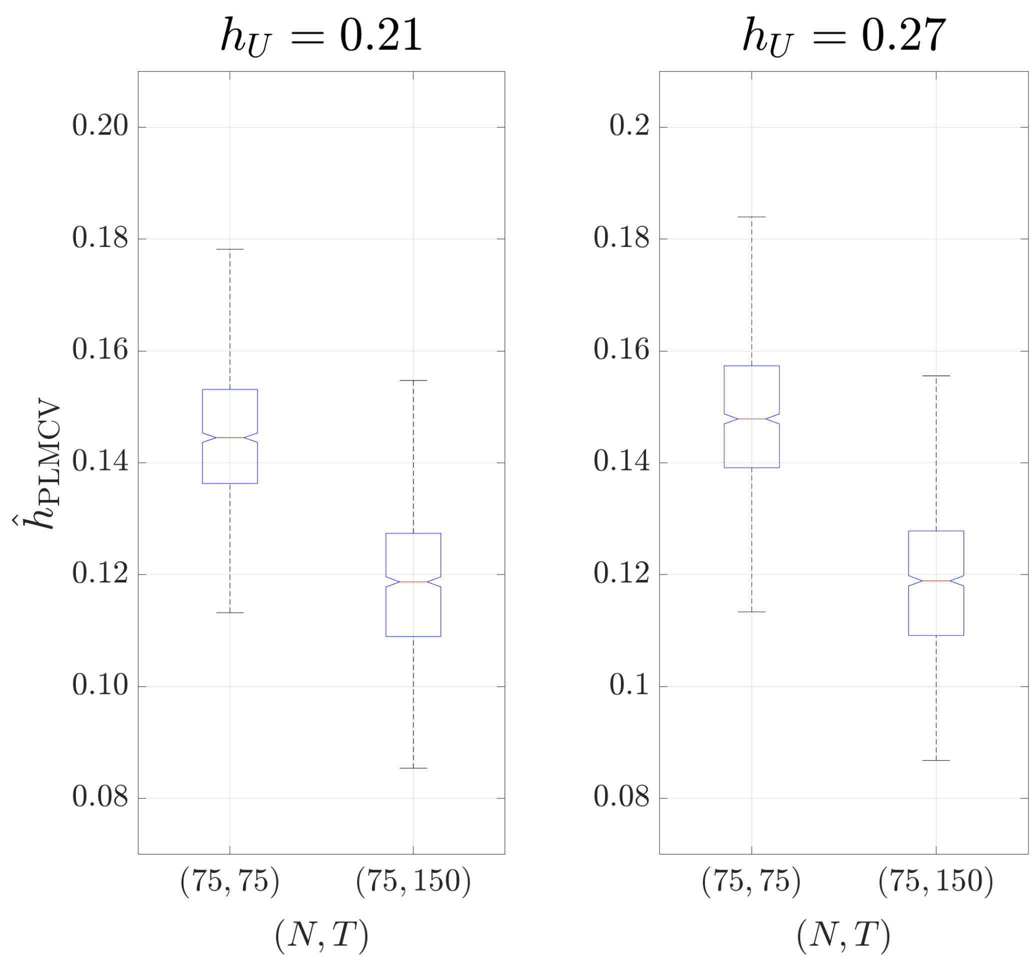


Figure E.1: Selected bandwidth by PLMCV with heteroskedastic errors and $\rho_\varepsilon = 0.3$.

E.1 Finite sample performance of PLMCV

In the simulation design outlined in Section 5, we also examine the impact of the chosen range during a grid search. Specifically, the grid spans from 0.06 to h_U, with steps of 0.015, where h_U takes values of 0.21 and 0.27. Figure E.1 shows the bandwidths selected by PLMCV for different h_U. As T increases, the selected bandwidth generally decreases. For a large $h_U = 0.27$, the PLMCV-selected bandwidth remains comparable to those under $h = 0.21$ on average.

F Bootstrap-corrected simultaneous confidence bands

In practice, one might be interested in understanding the overall variation of coefficient curves over a specific period. Simultaneous confidence bands serve this purpose. Specifically, for a given set of time points $G \subset [0, 1]$, simultaneous bands $I^G_{j,N,T,\alpha}(\cdot)$ should satisfy $\liminf_{T\to\infty, N\to\infty} \mathbb{P}\Big(\beta_j(\tau) \in I^G_{j,N,T,\alpha}(\tau),\ \forall \tau \in G\Big) \geq 1 - \alpha$. To address this, one may consider using the Bonferroni correction. However, it is well known that the Bonferroni correction can result in highly conservative results, particularly when dealing with an infinite number of points in G. Moreover, as of our best knowledge, there is currently no available asymptotic simultaneous bands for our models, whether with or without missing observations. Even if such bands were available, they might suffer from slow convergence speeds, as known in the time series case (Zhou and Wu, 2010). In contrast, the residual-based bootstrap correction has demonstrated superior performance in finite samples for time series time-varying models (Friedrich and Lin, 2022). We therefore adapt the three-step procedure outlined in Friedrich and Lin (2022), originally proposed by Bühlmann (1998), for nonparametric time series to large panels using the AWB described in Section 4.1.

This procedure produces confidence bands that are simultaneous within a finite union of neighborhoods $G = \cup_{i=1}^{m} U_i(h)$, where the neighborhoods $U_i(h)$ take the form $U_i(h) = [\tau_i - ah, \tau_i + bh]$, with $0 \leq a, b < \infty$ and $m < \infty$. Given the similarity in the construction of $g(\cdot)$, we will provide details for $\beta_j(\cdot)$ only.

Step 1 Compute the pointwise quantiles $\widehat{q}_{j,\alpha_p/2}(\tau)$, $\widehat{q}_{j,1-\alpha_p/2}(\tau)$ by varying $\alpha_p \in [1/B, \alpha]$, for $\tau \in G$, $j = 1, \ldots, d$.

Step 2 Choose $\hat{\alpha}_s = \hat{\alpha}_s(\alpha)$ as

$$\hat{\alpha}_s = \underset{\alpha_p \in [1/B, \alpha]}{\arg\min} \left| \mathbb{P}^* \left(\widehat{q}_{j,\alpha_p/2}(\tau) \leq \widehat{\beta}_j^*(\tau) - \widetilde{\beta}_j(\tau) \leq \widehat{q}_{j,1-\alpha_p/2}(\tau),\ \forall \tau \in G \right) - (1-\alpha) \right|.$$

Step 3 Given $\hat{\alpha}_s$ from Step 2, construct the simultaneous confidence bands as

$$I_{j,N,T,\hat{\alpha}_s}^{G*}(\tau) = \left[\widehat{\beta}_j(\tau) - \widehat{q}_{j,1-\hat{\alpha}_s/2}(\tau),\ \widehat{\beta}_j(\tau) - \widehat{q}_{j,\hat{\alpha}_s/2}(\tau) \right], \qquad \tau \in G.$$

Step 2 essentially chooses a level $\hat{\alpha}_s$ such that

$$\frac{\# \left\{ \hat{\beta}_j^*(\tau) - \tilde{\beta}_j(\tau) \in \left[\hat{q}_{j,\hat{\alpha}_s/2}(\tau), \hat{q}_{j,1-\hat{\alpha}_s/2}(\tau) \right],\ \forall \tau \in G \right\}}{B} \approx 1 - \alpha,$$

where $\#E$ counts how many times the event E occurs in bootstrap. It typically results in a level $\hat{\alpha}_s$ which is much smaller than α to ensure the simultaneous coverage is close to $1-\alpha$. This often leads to a level $\hat{\alpha}_s$ significantly smaller than α to ensure that the simultaneous coverage is close to $1-\alpha$. The theoretical justification of this procedure in the current context is highly challenging, as it necessitates the establishment of some strong approximation results in large panels with cross-sectional and serial dependence, accommodating nonstationarity. Addressing these complexities is a topic for future research. Nevertheless, we assess the finite-sample performance of this procedure through an extensive simulation study as described in Section 5.

F.1 Simulation results of bootstrap-corrected simultaneous bands

In addition to pointwise confidence intervals, it is informative to investigate the coverage of simultaneous confidence bands. For this, we count the number of times the true curve lies within the confidence bands for all points in subsets of $[0, 1]$. As in Friedrich and Lin (2022), we consider subsets of the form $G_{sub} = U_1(h) \cup U_4(h)$ and $G = \bigcup_{i=1}^{4} U_i(h)$, with $U_i(h) = \left\{ (i/5) - h + j/100,\ j = 0, \ldots, \lfloor 200h \rfloor \right\} \bigcap [0, 1]$.

We further investigate empirical simultaneous coverage over the full sample. The results are given in Tables F.1 to F.4. We make the following five observations.

(i) The empirical simultaneous coverage is mildly lower than 95%. The undercoverage issue is more pronounced over the full sample, as covering all points $\{1/T, 2/T, \ldots, T/T\}$ simultaneously is highly challenging. In contrast to the pointwise intervals, the results show that the serial correlation substantially affects the simultaneous coverage of $g(\cdot)$. For $\beta_1(\cdot)$ and $\beta_2(\cdot)$, however, the coverage remains close across ρ_ε.

(ii) The empirical simultaneous coverage for all three coefficient curves increases as the sample size increases, approaching the nominal level, especially when T increases.

(iii) The choice of bandwidth affects the simultaneous coverage. For example, for $\beta_1(\cdot)$, we observe that the coverage is lowest for $h = 0.15$ across different data generating processes, except for $(N, T, \rho_\varepsilon) = (75, 75, 0.1)$. In theory, h should shrink to zero as the sample sizes increase. We observe that a large h leads to under-coverage of simultaneous bands. As discussed in Friedrich and Lin (2022), a smaller h is preferred for simultaneous bands if the function has more local features to be captured. The results show that a careful bandwidth selection method is crucial for simultaneous bands.

(iv) The simultaneous empirical coverage results with h selected by PLMCV align with those for fixed bandwidths, indicating the proposed PLMCV is reasonable to use in practice. However, relying on data-driven methods completely may be troublesome in practice; see also the discussion by Friedrich and Lin (2022) on data-driven bandwidth selection methods.

(v) Table F.4 shows that the simultaneous empirical coverage is relatively robust to the value of γ.

Table F.1: Simultaneous empirical coverage over G_{sub} for various sample sizes and heteroskedastic errors, $\gamma = 0.2$.

		$\rho_\varepsilon = 0.1$			$\rho_\varepsilon = 0.3$		
	h	$N = 75$, $T = 75$	$N = 75$, $T = 150$	$N = 150$, $T = 150$	$N = 75$, $T = 75$	$N = 75$, $T = 150$	$N = 150$, $T = 150$
g	0.09	0.885	0.911	0.932	0.771	0.833	0.842
	0.12	0.895	0.916	0.920	0.819	0.822	0.870
	0.15	0.909	0.923	0.931	0.824	0.853	0.858
	$\hat{h}_{\text{PLMCV}}$	0.893	0.915	0.935	0.808	0.855	0.859
β_1	0.09	0.870	0.941	0.976	0.901	0.932	0.972
	0.12	0.882	0.916	0.960	0.920	0.913	0.952
	0.15	0.874	0.841	0.845	0.846	0.849	0.869
	$\hat{h}_{\text{PLMCV}}$	0.873	0.924	0.956	0.879	0.883	0.971
β_2	0.09	0.884	0.939	0.926	0.885	0.939	0.916
	0.12	0.902	0.915	0.927	0.903	0.916	0.919
	0.15	0.895	0.914	0.923	0.891	0.916	0.914
	$\hat{h}_{\text{PLMCV}}$	0.876	0.922	0.929	0.873	0.918	0.916

Table F.2: Simultaneous empirical coverage over G for various sample sizes and heteroskedastic errors, $\gamma = 0.2$.

		$\rho_\varepsilon = 0.1$			$\rho_\varepsilon = 0.3$		
	h	$N = 75$, $T = 75$	$N = 75$, $T = 150$	$N = 150$, $T = 150$	$N = 75$, $T = 75$	$N = 75$, $T = 150$	$N = 150$, $T = 150$
g	0.09	0.843	0.898	0.935	0.711	0.799	0.843
	0.12	0.897	0.921	0.935	0.802	0.829	0.886
	0.15	0.907	0.927	0.938	0.820	0.866	0.865
	$\hat{h}_{\text{PLMCV}}$	0.896	0.922	0.944	0.796	0.838	0.863
β_1	0.09	0.815	0.894	0.950	0.820	0.887	0.935
	0.12	0.829	0.871	0.903	0.873	0.866	0.893
	0.15	0.812	0.781	0.758	0.797	0.775	0.770
	$\hat{h}_{\text{PLMCV}}$	0.816	0.879	0.911	0.825	0.853	0.931
β_2	0.09	0.858	0.927	0.906	0.850	0.917	0.905
	0.12	0.895	0.909	0.926	0.887	0.907	0.915
	0.15	0.897	0.914	0.923	0.884	0.919	0.921
	$\hat{h}_{\text{PLMCV}}$	0.870	0.919	0.910	0.868	0.902	0.909

Table F.3: Simultaneous empirical coverage over the full sample $\{t/T,\, t = 1, \ldots, T\}$ for various sample sizes and heteroskedastic errors, $\gamma = 0.2$.

		$\rho_\varepsilon = 0.1$			$\rho_\varepsilon = 0.3$		
	h	$N = 75$, $T = 75$	$N = 75$, $T = 150$	$N = 150$, $T = 150$	$N = 75$, $T = 75$	$N = 75$, $T = 150$	$N = 150$, $T = 150$
g	0.09	0.772	0.861	0.894	0.601	0.733	0.784
	0.12	0.853	0.900	0.922	0.761	0.789	0.863
	0.15	0.900	0.929	0.939	0.813	0.866	0.866
	$\hat{h}_{\text{PLMCV}}$	0.894	0.886	0.921	0.792	0.810	0.821
β_1	0.09	0.786	0.879	0.946	0.797	0.874	0.930
	0.12	0.816	0.871	0.902	0.867	0.859	0.893
	0.15	0.816	0.787	0.764	0.797	0.776	0.771
	$\hat{h}_{\text{PLMCV}}$	0.810	0.878	0.908	0.824	0.852	0.929
β_2	0.09	0.787	0.878	0.863	0.795	0.881	0.861
	0.12	0.862	0.880	0.884	0.840	0.877	0.877
	0.15	0.873	0.903	0.904	0.865	0.907	0.910
	$\hat{h}_{\text{PLMCV}}$	0.848	0.891	0.870	0.850	0.874	0.874

Table F.4: Simultaneous empirical coverage (Cov.) and length (Lgth.) for $(\rho_\varepsilon, N, T) = (0.3, 75, 150)$ with the bandwidth $\hat{h}_{\text{PLMCV}}$ selected by PLMCV (Appendix E) and heteroskedastic errors.

		γ	0.15	0.2	0.25	0.3	0.35	0.4
Cov.	G_{sub}	g	0.849	0.855	0.856	0.833	0.843	0.835
		β_1	0.877	0.883	0.918	0.922	0.928	0.935
		β_2	0.915	0.918	0.914	0.919	0.909	0.900
	G	g	0.834	0.838	0.850	0.831	0.848	0.836
		β_1	0.849	0.853	0.889	0.876	0.888	0.895
		β_2	0.903	0.902	0.907	0.912	0.910	0.899
	Full Sample	g	0.804	0.810	0.813	0.799	0.819	0.825
		β_1	0.844	0.852	0.884	0.871	0.886	0.892
		β_2	0.877	0.874	0.883	0.884	0.870	0.869
Lgth.	G_{sub}	g	0.241	0.247	0.250	0.254	0.263	0.269
		β_1	0.209	0.213	0.217	0.221	0.229	0.233
		β_2	0.162	0.163	0.163	0.162	0.164	0.162
	G	g	0.252	0.258	0.261	0.265	0.275	0.281
		β_1	0.219	0.223	0.228	0.232	0.240	0.245
		β_2	0.170	0.170	0.171	0.169	0.171	0.170
	Full Sample	g	0.254	0.260	0.263	0.267	0.276	0.283
		β_1	0.220	0.225	0.229	0.233	0.241	0.246
		β_2	0.171	0.171	0.172	0.170	0.172	0.171

G Additional discussions on simulations

This section provides supplementary discussions and results for our simulation study. In Section G.1, we present the estimation accuracy results for the LLDV estimator under two levels of serial dependence for heteroskedastic errors. Section G.2 details the computation of empirical coverage and length of confidence intervals and bands. Additional results regarding empirical coverage and length of pointwise confidence intervals and simultaneous confidence bands are presented in Section G.3.

G.1 Accuracy results

For $N = T = 75$ in Table G.1, we observe that the average AMSE increases when ρ_ε increases. The average AMSE is lowest for coefficient $\beta_2(\cdot)$, which can be explained by its smoothness. The amount of smoothness of the coefficient curve is also crucial for the choice of bandwidth h. We see in Table G.1 that, for relatively smooth curves $g(\cdot)$ and $\beta_2(\cdot)$, increasing h leads to lower AMSE on average. However, for $\beta_1(\cdot)$, $h = 0.15$ results in too much smoothing, so that AMSE increases. Similar results can be found for the standard deviation of the AMSE, in Table G.1. The standard deviation increases for $\rho_\epsilon = 0.3$, and generally, larger bandwidth results in a lower standard deviation of AMSE for $g(\cdot)$ and $\beta_2(\cdot)$.

When we increase the sample size, we see that the average AMSE decreases for all coefficient functions,

Table G.1: Average AMSE and standard deviation of AMSE, for heteroskedastic errors.

	Average				Standard deviation			
				$N = 75$, $T = 75$				
ρ_ε	h	g	β_1	β_2	h	g	β_1	β_2
	0.09	38.775	26.665	24.811	0.09	20.265	11.932	11.852
0.1	0.12	30.436	23.774	19.095	0.12	17.107	11.696	10.699
	0.15	26.316	24.474	15.473	0.15	15.610	12.486	8.949
	0.09	53.212	28.236	25.585	0.09	26.436	14.425	12.170
0.3	0.12	43.194	24.958	20.838	0.12	23.253	12.366	10.975
	0.15	36.819	25.914	16.004	0.15	21.338	13.391	9.155
				$N = 75$, $T = 150$				
ρ_ε	h	g	β_1	β_2	h	g	β_1	β_2
	0.09	19.239	13.770	11.651	0.09	9.351	6.545	5.462
0.1	0.12	15.764	13.187	9.132	0.12	8.790	6.229	4.725
	0.15	14.011	16.653	7.484	0.15	7.834	7.735	4.137
	0.09	26.905	14.258	12.397	0.09	12.777	6.410	5.588
0.3	0.12	22.838	14.379	9.790	0.12	12.768	7.048	5.083
	0.15	23.420	18.811	9.438	0.15	14.180	8.519	5.443
				$N = 150$, $T = 150$				
ρ_ε	h	g	β_1	β_2	h	g	β_1	β_2
	0.09	9.853	7.301	5.891	0.09	4.776	3.217	2.751
0.1	0.12	8.271	8.664	4.636	0.12	4.226	3.871	2.498
	0.15	8.940	13.027	4.378	0.15	4.712	4.784	2.595
	0.09	13.776	7.767	6.258	0.09	7.236	3.486	2.992
0.3	0.12	12.414	9.561	5.504	0.12	6.686	4.222	2.868
	0.15	12.023	13.241	4.462	0.15	6.583	4.750	2.582

across different levels of serial dependence. These results confirm the consistency of the LLDV estimator in our model. Similarly, the standard deviation of the AMSE decreases as the sample size increases, suggesting that the LLDV estimates get more centered around the true value, as our theory predicts.

G.2 Computing empirical coverage and length

(*a*) Empirical pointwise coverage: For each $j = 1, \ldots, d$ and each Monte Carlo iteration, we calculate the percentage of $\beta_j(\tau_t)$ covered by the bootstrap intervals for $t = 1, \ldots, T$. The average of these percentages over the total of M iterations is then computed.

(*b*) Empirical simultaneous coverage: For each $j = 1, \ldots, d$ and each Monte Carlo iteration, we determine whether the set $\{\beta_j(\tau),\, \tau \in G\}$ is entirely contained within the confidence bands over G. The empirical simultaneous coverage is calculated as the success rate across all M iterations.

(*c*) Empirical length: For each Monte Carlo iteration, we compute the median length of intervals/bands across the time grid $\{1/T, 2/T, \ldots, T/T\}$. The average of these medians is then computed over M iterations.

G.3 Full results of empirical coverage and length

We report the complete set of empirical coverage and length for both pointwise intervals and simultaneous bands in Table G.2.

Table G.2: Empirical coverage and length of 95%-level pointwise confidence intervals and simultaneous confidence bands, $\gamma = 0.2$.

			$(N,T)=(75,75)$				$(N,T)=(75,150)$				$(N,T)=(150,150)$			
ρ_ϵ		h	Pointwise	G_{sub}	G	Full Sample	Pointwise	G_{sub}	G	Full Sample	Pointwise	G_{sub}	G	Full Sample
Empirical coverage														
0.1	g	0.09	0.945	0.885	0.843	0.772	0.957	0.911	0.898	0.861	0.968	0.932	0.935	0.894
		0.12	0.954	0.895	0.897	0.853	0.960	0.916	0.921	0.900	0.971	0.920	0.935	0.922
		0.15	0.959	0.909	0.907	0.900	0.964	0.923	0.927	0.929	0.970	0.931	0.938	0.939
	β_1	0.09	0.953	0.870	0.815	0.786	0.965	0.941	0.894	0.879	0.979	0.976	0.950	0.946
		0.12	0.945	0.882	0.829	0.816	0.948	0.916	0.871	0.871	0.957	0.960	0.903	0.902
		0.15	0.925	0.874	0.812	0.816	0.901	0.841	0.781	0.787	0.888	0.845	0.758	0.764
	β_2	0.09	0.939	0.884	0.858	0.787	0.953	0.939	0.927	0.878	0.950	0.926	0.906	0.863
		0.12	0.943	0.902	0.895	0.862	0.952	0.915	0.909	0.880	0.956	0.927	0.926	0.884
		0.15	0.947	0.895	0.897	0.873	0.947	0.914	0.914	0.903	0.954	0.923	0.923	0.904
0.3	g	0.09	0.913	0.771	0.711	0.601	0.928	0.833	0.799	0.733	0.943	0.842	0.843	0.784
		0.12	0.929	0.819	0.802	0.761	0.932	0.822	0.829	0.789	0.950	0.870	0.886	0.863
		0.15	0.929	0.824	0.820	0.813	0.942	0.853	0.866	0.866	0.948	0.858	0.865	0.866
	β_1	0.09	0.957	0.901	0.820	0.797	0.965	0.932	0.887	0.874	0.977	0.972	0.935	0.930
		0.12	0.954	0.920	0.873	0.867	0.947	0.913	0.866	0.859	0.956	0.952	0.893	0.893
		0.15	0.924	0.846	0.797	0.797	0.897	0.849	0.775	0.776	0.885	0.869	0.770	0.771
	β_2	0.09	0.942	0.885	0.850	0.795	0.954	0.939	0.917	0.881	0.949	0.916	0.905	0.861
		0.12	0.945	0.903	0.887	0.840	0.949	0.916	0.907	0.877	0.954	0.919	0.915	0.877
		0.15	0.943	0.891	0.884	0.865	0.951	0.916	0.919	0.907	0.954	0.914	0.921	0.910
Empirical length														
0.1	g	0.09	0.243	0.346	0.366	0.366	0.177	0.253	0.267	0.268	0.142	0.202	0.214	0.216
		0.12	0.227	0.325	0.340	0.342	0.163	0.234	0.245	0.246	0.138	0.198	0.207	0.208
		0.15	0.213	0.307	0.314	0.317	0.158	0.227	0.232	0.236	0.135	0.195	0.199	0.202
	β_1	0.09	0.223	0.317	0.336	0.337	0.161	0.228	0.242	0.243	0.136	0.192	0.204	0.205
		0.12	0.201	0.288	0.303	0.304	0.144	0.206	0.216	0.217	0.124	0.177	0.186	0.187
		0.15	0.184	0.263	0.271	0.273	0.134	0.192	0.197	0.200	0.116	0.165	0.171	0.173
	β_2	0.09	0.172	0.246	0.259	0.259	0.122	0.175	0.185	0.186	0.089	0.127	0.135	0.136
		0.12	0.153	0.219	0.230	0.231	0.108	0.155	0.163	0.163	0.081	0.115	0.121	0.122
		0.15	0.138	0.197	0.203	0.205	0.100	0.143	0.147	0.149	0.075	0.107	0.110	0.112
0.3	g	0.09	0.247	0.351	0.371	0.372	0.184	0.262	0.277	0.278	0.147	0.209	0.221	0.223
		0.12	0.231	0.330	0.346	0.348	0.169	0.242	0.253	0.255	0.141	0.202	0.212	0.213
		0.15	0.227	0.325	0.333	0.336	0.159	0.229	0.234	0.238	0.137	0.197	0.202	0.205
	β_1	0.09	0.227	0.322	0.342	0.342	0.164	0.233	0.248	0.249	0.138	0.195	0.207	0.209
		0.12	0.206	0.294	0.309	0.310	0.148	0.211	0.221	0.222	0.126	0.179	0.188	0.189
		0.15	0.194	0.277	0.286	0.287	0.134	0.192	0.198	0.201	0.117	0.166	0.171	0.173
	β_2	0.09	0.175	0.249	0.263	0.263	0.127	0.181	0.191	0.193	0.092	0.131	0.139	0.140
		0.12	0.156	0.224	0.235	0.236	0.112	0.161	0.168	0.169	0.083	0.119	0.125	0.126
		0.15	0.149	0.213	0.219	0.221	0.101	0.145	0.149	0.151	0.076	0.109	0.112	0.114

H Additional empirical results

This section presents additional empirical results and information about the data acquisition.

H.1 Additional results for mortality and $PM_{2.5}$

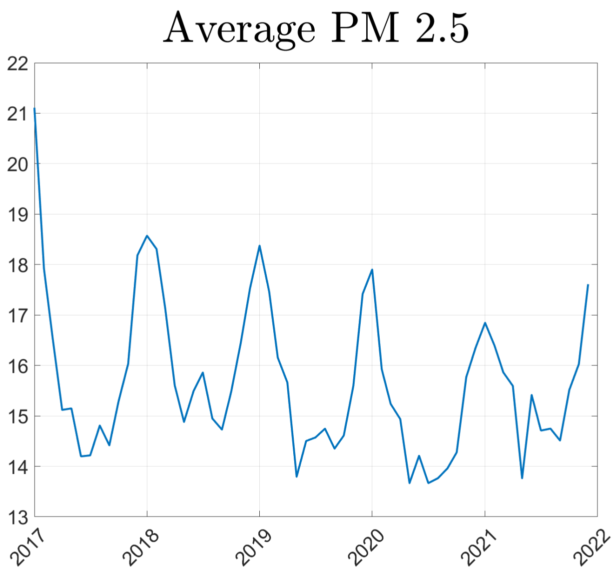


Figure H.1: Cross-sectional average for $PM_{2.5}$.

Figure H.1 displays the cross-sectional average for $PM_{2.5}$. The regressor $PM_{2.5}$ is slightly downward trending. In our theory, we allow for this type of trending pattern.

H.1.1 Additional estimation results

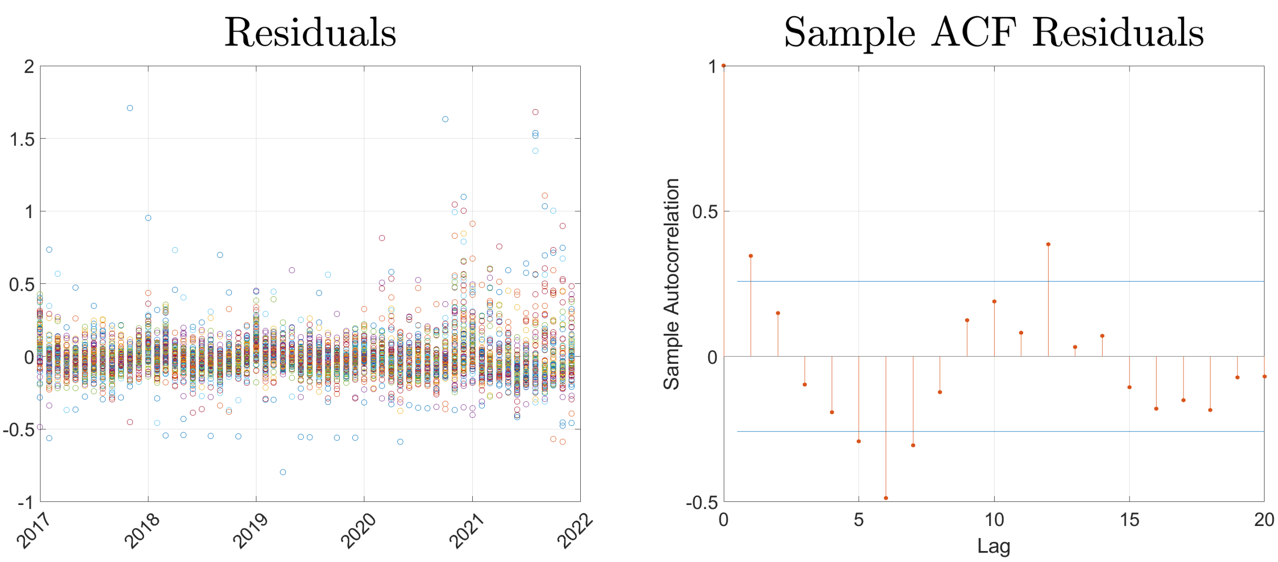


Figure H.2: The left figure displays the residuals after employing the LLDV estimator with bandwidth $\hat{h}_{\text{PLMCV}} = 0.2605$ and $\gamma = 0.2$. The right figure displays the sample autocorrelation function of the cross-sectionally averaged residuals.

In Figure H.2 (left) we show the residuals from employing the LLDV estimator on our data. We observe that the dispersion in the residuals varies per time point. We see that the dispersion in 2021 is larger than the dispersion in 2017 for instance. Our theory allows for this heteroskedasticity.

Moreover, our theory allows for serial correlation in the error process. In Figure H.2 (right) we display the sample autocorrelation function of the cross-sectionally averaged residual. We see that there is significant autocorrelation present in the residuals. In particular, the sample autocorrelation is significant for 1, 5, 6, and 12 months of lag.

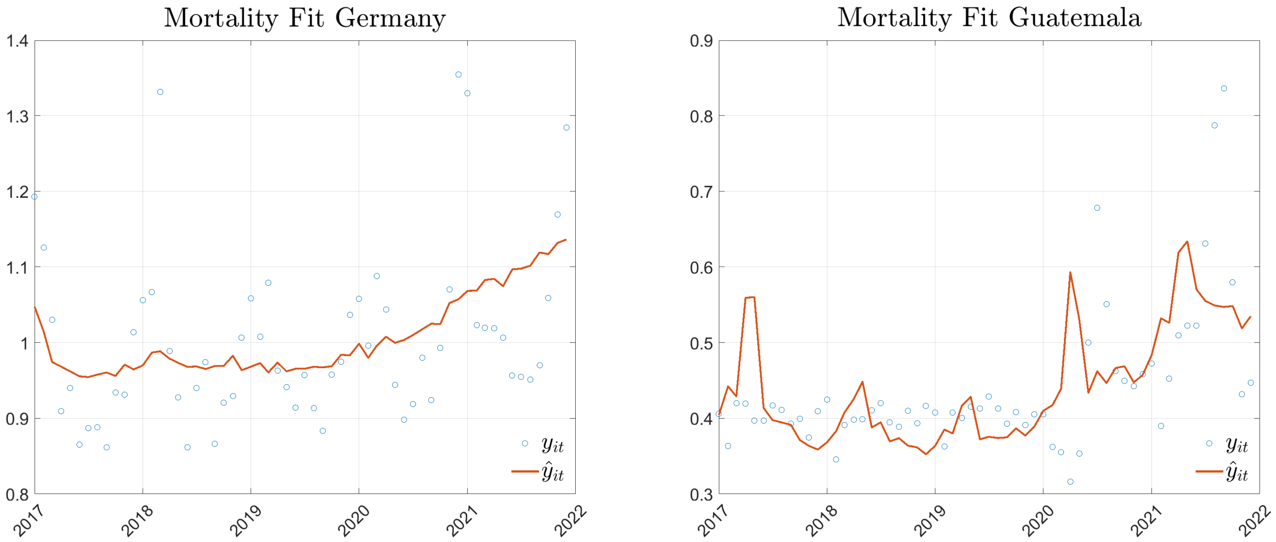
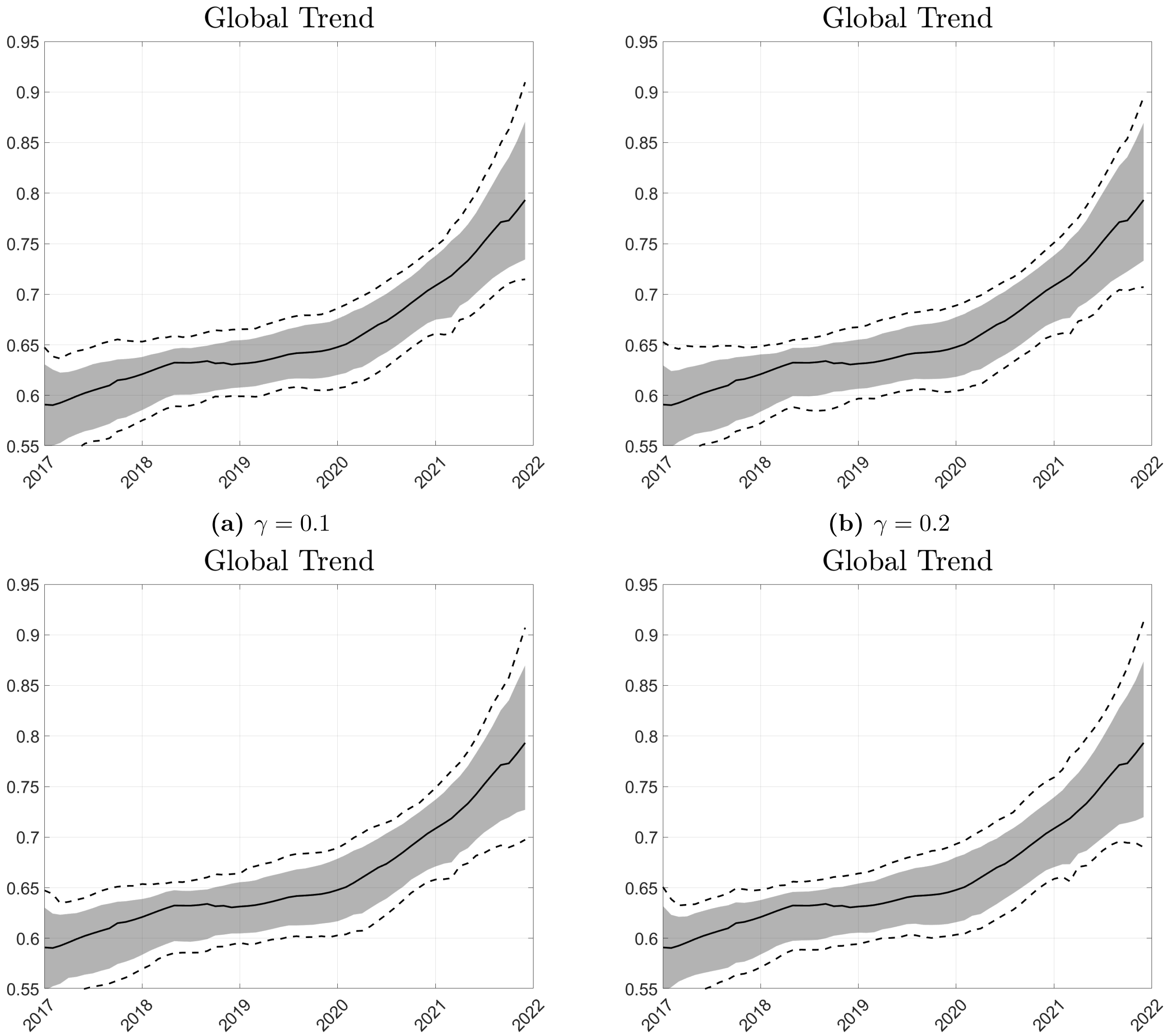


Figure H.3: Fit of our model and LLDV estimator for Germany and Guatemala. We use $\hat{h}_{\text{PLMCV}} = 0.2605$ and $\gamma = 0.2$.

Figure H.3 displays the estimated mortality compared to the monthly reported mortality for Germany and Guatemala. Given the substantial differences in economy, healthcare systems, and demographics between these countries, the fits illustrate the efficacy of the model and the LLDV estimator that allows for missing observations.

H.1.2 Robustness checks



(a) $\gamma = 0.1$

(b) $\gamma = 0.2$

(c) $\gamma = 0.3$

(d) $\gamma = 0.4$

Figure H.4: Robustness to choice of γ parameter: estimated global trend, and 95%-level pointwise intervals and simultaneous bands for $\gamma = 0.1, 0.2, 0.3, 0.4$.

In Figure H.4 we depict the global trend and estimated confidence intervals and bands for different values of the γ parameter. We fix the bandwidth to $h = 0.2638$, which corresponds to $\hat{h}_{\text{PLMCV}}$ for $\gamma = 0.2$. We observe that the width is similar for different values of γ.

H.2 Bandwidth selection for Ethane analysis

A large value of T intensifies the computational demands for bandwidth selection. In this instance, the computational time is long with $T = 13,394$ days. To circumvent this issue, we segment our data into non-overlapping blocks along the time dimension, each with a fixed cross-sectional dimension of $N = 11$. For each block, we calculate the bandwidth, and the final bandwidth is obtained by averaging the computed values across all blocks. Employing a window size of 250 days, we appropriately divide the

Station	Country	Location	Altitude	Obs.	Missing	Reference
Boulder	United States	40°N, 105°W	1,634m	788	81.61%	
Bremen	Germany	53°N, 9°E	27m	544	91.82%	Notholt et al. (2000)
Eureka	Canada	80°N, 86°W	610m	828	83.40%	Batchelor et al. (2009)
Jungfraujoch	Switzerland	47°N, 8°E	3,580m	3,171	76.26%	Franco et al. (2015)
Mauna Loa	United States	20°N, 156°W	3,397m	2,746	71.46%	
Ny-Ålesund	Norway	79°N, 12°E	15m	882	90.88%	Notholt et al. (1997)
Paramaribo	Suriname	6°N, 55°W	23m	102	96.89%	
Rikubetsu	Japan	43°N, 144°E	380m	1,078	89.10%	
Thule	Greenland	77°N, 69°W	225m	1,464	81.75%	
Toronto	Canada	44°N, 79°W	174m	2,377	67.78%	Yamanouchi et al. (2023)
Tsukuba	Japan	36°N, 140°E	31m	1,089	85.71%	

Table H.1: Information on the location of the FTIR measurement stations as well as the number of observations and missing data of the individual time series.

data. The last block accommodates the remaining observations, totaling 144 days. The final result gives $\hat{h}_{\mathrm{PLMCV}} = 0.1298$, and is relatively robust to different window sizes.

H.3 Additional information about data acquisition

Mortality and $PM_{2.5}$

The data of the Atmospheric Composition Analysis Group, Van Donkelaar et al. (2021), can be found at ACAG (`https://sites.wustl.edu/acag/datasets/surface-pm2-5/`). The mortality data can be found at UNdata (`https://unstats.un.org/unsd/demographic-social/products/dyb/index.cshtml`).

Ethane

With permission to use the data, we acquire ethane data from the following 11 stations: Thule, Boulder, Mauna Loa, Ny Alesund, Bremen, Paramaribo, Eureka, Toronto, Jungfraujoch, Rikubetsu, and Tsukuba. More information on the specific sites can be obtained from Table H.1. While the data is freely available on the NDACC website, for its use in publications or communications, it is essential to contact the NDACC principal investigators during the preparation phase to discuss potential collaboration and co-authorship.

References in Appendix

Batchelor, R. L., K. Strong, R. Lindenmaier, R. Mittermeier, Fast, H. Drummond, and P. Fogal (2009). A new bruker ifs 125hr ftir spectrometer for the polar environment atmospheric research laboratory at eureka, nunavut, canada: Measurements and comparison with the existing bomem da8 spectrometer. *Journal of Atmospheric and Oceanic Technology 26*.

Boswijk, H. P., G. Cavaliere, I. Georgiev, and A. Rahbek (2021). Bootstrapping non-stationary stochastic volatility. *Journal of Econometrics 224*(1), 161–180. Annals Issue: PI Day.

Bühlmann, P. (1998). Sieve bootstrap for smoothing in nonstationary time series. *The Annals of Statistics 26*, 48 – 83.

Cheng, G. and J. Z. Huang (2010). Bootstrap consistency for general semiparametric M-estimation. *The Annals of Statistics 38*(5), 2884 – 2915.

Chu, C.-K. and J. S. Marron (1991). Comparison of two bandwidth selectors with dependent errors. *The Annals of Statistics 19*(4), 1906–1918.

Davidson, J. (1994). *Stochastic Limit Theory*. Oxford University Press, Oxford.

de Jong, R. M. and J. Davidson (2000). Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices. *Econometrica 68*(2), 407–423.

Franco, B., W. Bader, G. C. Toon, C. Bray, A. Perrin, E. V. Fischer, et al. (2015). Retrieval of ethane from ground-based FTIR solar spectra using improved spectroscopy: Recent burden increase above Jungfraujoch. *Journal of Quantitative Spectroscopy and Radiative Transfer 160*, 36–49.

Friedrich, M. and Y. Lin (2022). Sieve bootstrap inference for linear time-varying coefficient models. *Journal of Econometrics (forthcoming)*.

Friedrich, M., S. Smeekes, and J.-P. Urbain (2020). Autoregressive wild bootstrap inference for nonparametric trends. *Journal of Econometrics 214*(1), 81–109. Annals Issue: Econometric Models of Climate Change.

Gao, J., K. Xia, and H. Zhu (2020). Heterogeneous panel data models with cross-sectional dependence. *Journal of Econometrics 219*(2), 329–353. Annals Issue: Econometric Estimation and Testing: Essays in Honour of Maxwell King.

Gine, E. and J. Zinn (1990). Bootstrapping general empirical measures. *The Annals of Probability 18*(2), 851–869.

Gonçalves, S. and B. Perron (2014). Bootstrapping factor-augmented regression models. *Journal of Econometrics 182*(1), 156–173. Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions.

Hansen, B. E. (1991). Strong laws for dependent heterogeneous processes. *Econometric Theory 7*, 213–221.

Jiang, W. (2009). On uniform deviations of general empirical risks with unboundedness, dependence, and high dimensionality. *Journal of Machine Learning Research 10*(36), 977–996.

Notholt, J., G. Toon, C. Rinsland, N. Pougatchev, N. Jones, B. Connor, et al. (2000). Latitudinal variations of trace gas concentrations in the free troposphere measured by solar absorption spectroscopy during a ship cruise. *Journal of Geophysical Research 105*, 1337–134.

Notholt, J., G. Toon, F. Stordal, S. Solberg, N. Schmidbauer, E. Becker, et al. (1997). Seasonal variations of atmospheric trace gases in the high Arctic at 79°N. *Journal of Geophysical Research 102*, 12,855–12,861.

Sun, Y., R. J. Carroll, and D. Li (2009). Semiparametric estimation of fixed-effects panel data varying coefficient models. In *Nonparametric econometric methods*. Emerald Group Publishing Limited.

Van Donkelaar, A., M. S. Hammer, L. Bindle, M. Brauer, J. R. Brook, M. J. Garay, et al. (2021). Monthly global estimates of fine particulate matter and their uncertainty. *Environmental Science & Technology 55*(22), 15287–15300.

Vieu, P. (1991). Nonparametric regression: Optimal local bandwidth choice. *Journal of the Royal Statistical Society: Series B (Methodological) 53*(2), 453–464.

White, H. (2001). *Asymptotic theory for econometricians* (Rev. ed.). Economic theory, econometrics, and mathematical economics. Bingley: Emerald.

Xiong, S. and G. Li (2008). Some results on the convergence of conditional distributions. *Statistics & Probability Letters 78*(18), 3249–3253.

Yamanouchi, S., S. Conway, K. Strong, O. Colebatch, E. Lutsch, S. Roche, et al. (2023). Network for the detection of atmospheric composition change (ndacc) fourier transform infrared (ftir) trace gas measurements at the university of toronto atmospheric observatory from 2002 to 2020. *Earth System Science Data 15*, 3387–3418.

Zhou, Z. and W. B. Wu (2010). Simultaneous inference of linear models with time varying coefficients. *Journal of the Royal Statistical Society. Series B (Statistical Methodology) 72*, 513–531.