Econometric Theory, **0**, 2026, 1–29.
doi:10.1017/S026646662510025X

UNIFORM INFERENCE FOR NONPARAMETRIC PANEL MODELS WITH FIXED EFFECTS

NAN LIU
Xiamen University

YANBO LIU
Shandong University

This article studies uniform inference on a function $g(\cdot)$ and its functionals in a nonparametric panel data model with fixed effects. The nonparametric panel model relaxes restrictions on time-series behavior by allowing for arbitrary types of stationary or nonstationary dependence (e.g., stationary mixingale, mildly stationary, or local-to-unity process). After removing the fixed effects via transformations, a sieve estimator is proposed, accompanied by Yurinskii's coupling principle of Gaussian processes and uniform confidence bands (UCBs) that rely on the sieve score bootstrap method to test for linear functionals of $g(\cdot)$. Under the asymptotic framework of an increasing cross-sectional dimension and either a fixed or diverging time dimension, we prove that the bootstrapping Kolmogorov–Smirnov (sup-type) test has asymptotic uniform size controls. This article shows that our uniform inference procedure can be extended to the two-way fixed-effects nonparametric panel model with stationary mixingale regressors. Extensive simulations confirm that our sieve estimators and their UCBs work well in finite samples. The present article further applies the above methods to empirical settings and finds some interesting results in nonlinear patterns of consumption concerning income shocks and asset holdings.

1. INTRODUCTION

It is acknowledged that evaluations of economic theory often depend on statistical inference on linear functionals in the nonparametric models. In the related literature, the associated discussions mainly focus on the pointwise testing procedures of these functionals by providing confidence intervals (CIs) evaluated at finite grid points. The pointwise CIs that ignore the guiding information on other parts of the support can lead to spurious statistical significance and misguide

The authors are indebted to the Editor, Peter C. B. Phillips, for very insightful advice and comments, which have considerably improved this work. The authors are grateful to the Co-Editor, Jia Li, and three anonymous referees for very helpful comments on the article and constructive guidance for revision. The authors gratefully acknowledge Jinyong Hahn, Arthur Lewbel, Zhipeng Liao, Whitney Newey, Ryo Okui, Yuya Sasaki, Liangjun Su, Zhentao Shi, Qiying Wang, Jun Yu, and Yichong Zhang for generous suggestions. Address correspondence to Yanbo Liu, School of Economics, Shandong University, Jinan, China, e-mail: yanbo.liu.metrics@outlook.com.

© The Author(s), 2026. Published by Cambridge University Press.

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

economic decisions, calling upon the alternative uniform confidence bands (UCBs) whose validity holds across the compact support (or a compact subset). Such concerns motivate us to develop uniform inference procedures for a broad class of nonparametric panel models.

To be specific, we consider a nonparametric panel data model with one-way fixed effects,

$$y_{it} = \gamma_i + g(x_{it}) + \epsilon_{it}, \quad i \in [n], \quad t \in [T],$$

in which $g(\cdot)$ is an unknown function that belongs to the function class $\mathcal{G}$, $x_{it} \in \mathcal{X} \subseteq \mathbb{R}^d$ is a d-dimensional vector of covariates, which follow arbitrary types of stationary or nonstationary dependence (e.g., stationary mixingale, mildly stationary (Phillips and Magdalinos, 2007), unit root, or local-to-unity (Phillips, 1987) processes), γ_i denotes the individual fixed effect, and ϵ_{it} is an idiosyncratic error component.

In order to introduce the UCBs for the unknown function $g(x)$, we first evaluate the upper bounds for estimating $g(x)$ and its functionals (e.g., derivatives) based on the panel within-group sieve estimator $\widehat{g}(\cdot)$ that will be introduced in Section 2; that is, we bound

$$\sup_{x\in\mathcal{C}} |\widehat{g}(x) - g(x)| \text{ and } \sup_{x\in\mathcal{C}} \left|\partial^j \widehat{g}(x) - \partial^j g(x)\right|,$$

in which $\partial^j \widehat{g}(\cdot)$ represents the jth-order partial derivatives of $g(\cdot)$ in terms of x and $\mathcal{C}$ is any compact subset of the support, $\mathcal{X}$. Specifically, we provide the sup-norm upper bound for our sieve estimator $\widehat{g}(\cdot)$. Next, we establish the strong Gaussian approximation for the estimated linear functionals of $g(x)$ in our panel model under various types of intertemporal dependence. Finally, we show the uniform probability coverage of the sieve score bootstrap UCBs for the linear functionals of $g(x)$ under primitive conditions.

In the asymptotic development, we approximate the empirical process by a Gaussian process of a growing dimension. The strong approximation allows us to perform inference across any compact subset $\mathcal{C}$ and to test for various functionals of $g(\cdot)$, which is infeasible given the pointwise limit theory. Our article extends Yurinskii's coupling to a nonparametric fixed-effects panel data model that allows for arbitrary types of stationary or nonstationary processes. To the best of our knowledge, the present article belongs to the first of including the near unit root process in establishing the coupling principle.

Moreover, this article extends the proposed framework to the nonparametric panel data model that incorporates both individual and time fixed effects, γ_i and f_t. We require that the considered regressors be stationary mixingale but allow the time fixed effects, f_t, to have any intertemporal dependence. We extend our strong approximation and bootstrapped UCBs to the two-way fixed-effects panel model and prove the validity.

Extensive simulations compare the performance of the bootstrapping UCBs with pointwise CIs and verify the good behavior of the bootstrapping UCBs in

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

the finite sample. In particular, the widths of the pointwise CIs are narrower than UCBs, showing that severe size distortions of rejecting the uniform null hypotheses can arise when the empirical analysis is based on pointwise CIs. Various specifications are examined, including the one-way fixed-effects panel data model with a univariate regressor, the partial linear panel model[1] and the two-way fixed-effects panel model. Also, sensitivity analysis is conducted regarding the order of sieve expansion, showing the robust uniform size controls of our UCBs.

In addition, we present an empirical example to illustrate the usefulness of UCBs for analyzing income–consumption dynamics. We provide the bootstrapping UCBs and pointwise CIs for consumption concerning income shocks and asset holdings. The empirical results show that the UCBs of each covariate are wider than its pointwise CIs, coinciding with the findings of numerical simulations.

This article is related to three strands of literature. First, the article contributes to the literature on coupling and high-dimensional central limit theorems (CLTs). For the independent random vectors, the high-dimensional CLTs and bootstrap theory are provided in Chernozhukov, Chetverikov, and Kato (2014). Moreover, high-dimensional CLTs are discussed in the case of time series and network/spatial data (Li and Liao, 2020; Kurisu, Kato, and Shao, 2024; Li, Liao, and Zhou, 2024). Our article extends Yurinskii's coupling principle to the panel data model that incorporates nonstationary processes and provides UCBs for the unknown functionals in a nonparametric panel data model.

Second, the present article contributes to the literature on sieve estimations of the nonstationary time-series models. Due to the wandering behavior of nonstationary processes, the range of a unit root process, $\mathcal{X}$, is expanded to the real number line $\mathbb{R}$, violating the assumption of compact support in much of the literature on sieve estimation. As an alternative, Dong, Gao, and Peng (2019) considered the orthogonal Hermite polynomial system defined on the real number line and further established the inference procedures relying on sieve estimations on the unbounded support. To the best of our knowledge, no strong approximation theory or uniform inference procedure is available for the series/sieve estimators that depend on the Hermite polynomial system.

The present article also contributes to the literature on specification testing in the nonparametric panel model. Discussions of estimations and specification tests for nonparametric/semi-nonparametric panel data models that will be introduced later focus more on pointwise asymptotic theory and the Cramér–von Mises test (Su and Lu, 2013; Lee, 2014). Though econometricians can still achieve consistent specifications via the Cramér–von Mises tests, specific types of tests can generate better power performance against a particular class of alternatives. So, it is essential to apply the test to a suitable setting and examine the panel data sup-test, which remains to be investigated. This article contributes to this field by filling this void.

[1] The discussions of the one-way fixed-effects panel data models with additive and nonseparable functional forms are given in the Supplementary Material.

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

The structure of the present article is as follows. Section 2 presents our nonparametric panel data model with one-way individual fixed effects and discusses the sieve estimator and its properties. Section 3 derives the uniform limit theory for the sieve estimation. Section 4 provides the sieve score bootstrap UCBs and proves their uniform size controls. Section 5 extends our bootstrapping uniform inference procedure to the nonparametric panel model with two-way fixed effects. Section 6 examines the finite-sample performance of the sieve score bootstrapping UCBs using Monte Carlo simulations. Section 7 applies the uniform inference procedure to empirical data. Section 8 concludes. The Supplementary Material collects all the proofs and additional simulation studies.

In what follows, the ℓ_2-norm of the vector v is denoted by $\|v\|$ and the operator norm for a matrix M is given by $\|M\|$. For a matrix-valued function $Q(x)$ with $x \in \mathcal{X}$, the notation $\|Q\|_{\mu,q} := (\int_{x\in\mathcal{X}} \|Q(x)\|^q \mu(x)dx)^{1/q}$, in which $\mu(\cdot)$ is the Lebesgue measure. The space $L_2(\mathbb{R}, \mu(x))$ represents the class of functions with $\|Q(x)\|_{\mu,2} < \infty$. The notation $a \lesssim b$ denotes $a \leq cb$ for some constant c, while the notation $a \lesssim_p b$ represents $a = O_p(b)$ or $a = o_p(b)$. The notation $a \asymp b$ means that $a \lesssim b$ and $b \lesssim a$. The notations $\to_p$ and $\Rightarrow$ denote convergence in probability and weak convergence. The notation $\mathbf{1}_{1\times d}$ refers to a d-dimensional row vector of ones, the notation $\mathbf{0}_{d\times 1}$ denotes a d-dimensional column vector of zeros, and the quantity $\mathbf{0}_{d\times d}$ represents a $d \times d$ matrix of zeros. We use standard notations $\mathbb{E}_n[\cdot]$, $\mathbb{E}_T[\cdot]$, $\mathbb{E}^*_T[\cdot]$, $\mathbb{G}_n[\cdot]$, and $\mathbb{G}_T[\cdot]$ in the empirical process literature as $\mathbb{E}_n[f] = (1/n)\sum_{i=1}^n f(w_i)$, $\mathbb{E}_T[f] = (1/a_T^2)\sum_{t=1}^T f(w_t)$, $\mathbb{E}^*_T[f] = (1/T)\sum_{t=1}^T f(w_t)$, $\mathbb{G}_n[f] = (1/\sqrt{n})\sum_{i=1}^n (f(w_i) - \mathbb{E}(f(w_i)))$, and $\mathbb{G}_T[f] = (1/a_T)\sum_{t=1}^T (f(w_t) - \mathbb{E}(f(w_t)))$, where a_T is the stabilization rate (i.e., $a_T = \sqrt{T}$ for stationary mixingale case, $a_T = \sqrt{T^{1-\zeta/2}}$ for mildly stationary case, and $a_T = \sqrt[4]{T}$ for the unity/local-to-unity case).[2] Write $i \in [n]$ to denote $i = 1, \ldots, n$ and use similar notations as $t \in [T]$.

2. MODEL SETUP AND SIEVE ESTIMATION

2.1. Model

We restate the additive nonparametric panel data model with individual fixed effects of the introduction. The dependent variable y_{it} is generated according to

$$y_{it} = \gamma_i + g(x_{it}) + \epsilon_{it}, \quad \mathbb{E}\left[\epsilon_{it} | \{x_{it}\}_{t\in[T]}\right] = 0, \quad i \in [n] \text{ and } t \in [T], \tag{1}$$

in which $g(\cdot)$ is an unknown function of the space $\mathcal{G}$, x_{it} is a d-dimensional vector of covariates taking values in the set $\mathcal{X} \subseteq \mathbb{R}^d$, γ_i denotes the individual fixed effect that is correlated with regressors x_{it}, and ϵ_{it} is an error term. We discuss a panel model with a large cross-sectional dimension n but impose no restrictions on the time-series dimension T; that is, T stays fixed or grows to infinity.

[2]Discussions of a_T in all three cases can be found in Section C.i of the Supplementary Material. The notation ζ is a generic parameter that is defined for the mildly stationary process and will be introduced later.

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

In equation (1), we impose cross-sectional independence for variates $(x_{it}', \epsilon_{it})'$ but incorporate various forms of time-series correlations for $(x_{it}', \epsilon_{it})'$. In particular, the error ϵ_{it} is a stationary mixingale that forms a general class of time-series behaviors, including martingale differences, linear processes, and various types of mixing and near-epoch-dependent processes as special cases (Li and Liao, 2020). Moreover, the covariate x_{it} follows an arbitrary type of stationary or nonstationary process which incorporates stationary mixingale, mildly stationary, unit root, or local-to-unit root process. Specifically, we impose the following conditions on x_{it} and ϵ_{it} to summarize the requirements.

Assumption 1. For each $i \in [n]$, the variates $\upsilon_{it} = \left(x_{it}', \epsilon_{it}\right)'$ that satisfy equation (1) are mutually independent across $i \in [n]$. Regression errors $\{\epsilon_{it}\}$ satisfy the uniform upper bound condition: $\sup_{x \in \mathcal{X}} \mathbb{E}\left[|\epsilon_{it}|^m | \{x_{it}\}_{t \in [T]}\right] \lesssim 1$, for some $m > 6$.

Assumption 2. For each $i \in [n]$, the variates $\{\upsilon_{it}\}$ follow a $(d+1)$-dimensional stationary L_q-mixingale array ($q \geq 3$) with respect to a filtration $\{\mathcal{F}_{i,t}\}$ that satisfies the following conditions: for any $l \geq 0$, with $\upsilon_{T,it} := \upsilon_{it}/\sqrt{T}$ and F as a cumulative distribution function (CDF) for $\{\upsilon_{it}\}$, it follows that

$$\left\| \mathbb{E}\left[\upsilon_{T,it} \middle| \mathcal{F}_{i,t-l} \right] \right\|_q \leq c_{i,t} \cdot \nu_l, \quad \left\| \upsilon_{T,it} - \mathbb{E}\left[\upsilon_{T,it} \middle| \mathcal{F}_{i,t+l} \right] \right\|_q \leq c_{i,t} \cdot \nu_{l+1}, \tag{2}$$

with $\|\upsilon\|_q = (\mathbb{E}\|\upsilon\|^q)^{1/q}$ for some positive sequence $\bar{c}_i$, $\sup_{t \in [T]} |c_{i,t}| \leq \bar{c}_i/\sqrt{T} = O(1)$ for each $i \in [n]$[3] and $\sum_{l=1}^{\infty} \nu_l < \infty$. Assume that v_{it} has sub-Gaussian tails.

Assumption 3. For each $i \in [n]$, the error $\{\epsilon_{it}\}$ follows the formulation of a mixingale process in Assumption 2, while the univariate ($d = 1$) regressor $\{x_{it}\}$ follows a near unit root process[4] as $x_{it} = (1 + C/T^{\zeta})x_{i,t-1} + u_{x,it}$ in which $\zeta \in (0, 1]$ and $u_{x,it} \sim \text{iid}(0, \sigma_x^2)$. The scalar quantity C is any constant scalar value when $\zeta = 1$ (unity/local-to-unity; Phillips, 1987) while is negative when $\zeta \in (0, 1)$ (mildly stationary; Phillips and Magdalinos, 2007).

Several remarks are given in order. First, in the present article, we impose cross-sectional independence across i to facilitate the strong approximation in the later discussions. The attempt to relax the restrictions of cross-sectional independence in x_{it} is meaningful and is left for future studies. Second, Assumptions 2 and 3 take into account two cases for $\{x_{it}\}$ and $\{\epsilon_{it}\}$. In Assumption 2, the covariates $\{x_{it}\}$ are stationary mixingales and uncorrelated with the instantaneous error $\{\epsilon_{it}\}$, similar as the setting of regressors and error term in Li and Liao (2020). While in Assumption 3, the covariate $\{x_{it}\}$ is a scalar near unit root process ($d = 1$) whose autoregressive root lies inside a shrinking neighborhood of unity, including the mildly stationary, unit root, and local-to-unity processes, which are

[3] This condition that ensures a stationary mixingale is also provided in Li and Liao (2020, Assumption 5) and allows us to approximate the partial sum of mixingale using a martingale. Our condition is also comparable to equation (2.10) of Li et al. (2024) before its cross-sectional averaging.

[4] The present article classifies the mildly stationary, unit root, and local-to-unity processes to the class of near-unit root processes as their autoregressive roots are in the shrinking neighborhood of unity.

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

widely used in the econometric development of bubble detection and predictability testing. The above two conditions cover a large class of regressors, showing the flexibility of incorporating the mixed-root phenomenon. The mixed-root features in $\{x_{it}\}$ also challenge our asymptotic establishment, which will be extensively investigated in the Supplementary Material. In later discussions, we define $\overline{\sigma}^2 := \sup_{x\in\mathcal{X}} \mathbb{E}\left[\epsilon_{it}^2 | x_{it} = x\right]$ and $\underline{\sigma}^2 := \inf_{x\in\mathcal{X}} \mathbb{E}\left[\epsilon_{it}^2 | x_{it} = x\right]$.

The pointwise inference results of the above nonparametric setup are detailedly studied in the literature (see Lee, 2014 for stationary covariates and Dong et al., 2019 for nonstationary regressors). However, the limit null distributions of these tests are pointwise that hold valid on the chosen grid points of $\mathcal{X}$ rather than uniformly over the support, limiting their applicability. To fill the void, this article provides UCBs for the $g(\cdot)$ that stay valid over any compact set of $\mathcal{X}$. Specifically, in this article, we propose a sieve estimator $\widehat{g}(\cdot)$ that will be introduced later and construct a UCB such that for a given confidence level $100(1-\alpha)\%$, the confidence band $[\widehat{g}_l(\cdot), \widehat{g}_u(\cdot)]$ provides uniform probability coverage as

$$\Pr\left(\widehat{g}_l(x) \le g(x) \le \widehat{g}_u(x) \text{ for all } x \in \mathcal{C}\right) = 1 - \alpha + o(1), \tag{3}$$

where the set $\mathcal{C}$ is any compact subset of $\mathcal{X}$.

To facilitate the identification of $g(\cdot)$, we impose the following assumption.

Assumption 4. The normalization condition holds that $g(\mathbf{0}_{d\times 1}) = 0$.

Assumption 4 assumes the identification condition for $g(\cdot)$. The normalization requires $g(\cdot)$ to pass through the origin. If the true function $g(\mathbf{0}_{d\times 1}) \neq 0$, we reparameterize the fixed effect as $\gamma_i + g(\mathbf{0}_{d\times 1})$ and the unknown function as $g(x) - g(\mathbf{0}_{d\times 1})$ such that the true data-generating process (DGP) still holds. We refer interested readers to Lee (2014) for more discussion of this assumption.

2.2. Construction of Sieve Estimator

We approximate the unknown function $x \mapsto g(x)$ with $x \in \mathcal{X}$ by sieve expansions $x \mapsto p(x)'b$, in which $x \mapsto p(x) := (p_1(x), p_2(x), \ldots, p_k(x))'$, is a vector of basis functions whose dimension k diverges with n. To stay consistent with the identification condition in Assumption 4, we further remove the constant value of $p(\cdot)$ and ensure that the normalization requirement is satisfied (e.g., $p(\mathbf{0}_{d\times 1}) = \mathbf{0}_{k\times 1}$). Based on the sieve approximation, we define

$$p_{it} := p(x_{it}) = (p_1(x_{it}), p_2(x_{it}), \ldots, p_k(x_{it}))'.$$

When $\{x_{it}\}$ are stationary mixingales (Assumption 2), their compact support $\mathcal{X}$ (Duffy, 2020) ensures that all sieve basis functions for the functional space of compact supports stay valid, including polynomial basis, spline series, Cohen–Daubechies–Vial (CDV) wavelets, etc. (Chen, 2007). When the near-unity regressor $\{x_{it}\}$ is used (Assumption 3), the support of $g(\cdot)$ is unbounded ($\mathcal{X} = \mathbb{R}$), failing the condition of a compact support. Therefore, the basis over compact supports

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

does not apply to the case of a near-unity regressor. Instead, for any $g(x) \in L_2(\mathbb{R})$ satisfying $\int_{\mathbb{R}} g^2(x)dx < \infty$, we consider the Hermite basis function as

$$p_j(x) = \frac{1}{\sqrt[4]{\pi}\sqrt{2^j j!}} H_j(x) \cdot \exp(-x^2/2), \ \ j \in [k], \tag{4}$$

in which $H_j(x)$ is the Hermite polynomial of the space $L_2(\mathbb{R}, \exp(-x^2))$, as

$$H_j(x) = (-1)^j \exp(x^2)(d^j \exp(-x^2)/dx^j), \ \ j \in [k]. \tag{5}$$

Then, for any $g(\cdot) \in L_2(\mathbb{R})$, the Hermite sieve expansion (Dong, Gao, and Zhu, 2021) follows

$$g(x) = \sum_{j=1}^{\infty} b_j p_j(x), \text{ with } b_j = \int_{\mathbb{R}} g(x) p_j(x) dx.$$

Remark 1. The Hermite polynomials, $H_j(\cdot)$ of (5), are the orthogonal basis of $L_2(\mathbb{R}, \exp(-x^2))$, satisfying the orthogonal properties regarding their Lebesgue measure (Dong et al., 2021):

$$\langle H_{j_1}(x), H_{j_2}(x) \rangle = \int_{x \in \mathbb{R}} H_{j_1}(x) H_{j_2}(x) \exp(-x^2) dx = \sqrt{\pi} 2^{j_1} j_1 ! \, \delta_{j_1, j_2}, \tag{6}$$

in which δ_{j_1, j_2} is the Kronecker delta that is unity when $j_1 = j_2$ and equals zero otherwise. Correspondingly, the Hermite basis functions of (4) can further serve as the orthogonal basis functions of the functional space $L_2(\mathbb{R})$ and facilitate the Hermite sieve expansions regarding unbounded supports.

The space $L_2(\mathbb{R})$ can help resolve the challenge brought by the unbounded supports of functionals that use nonstationary regressors, including the unit root and local-to-unity regressors in our one-way fixed-effects panel model. For instance, when x_{it} is a unit root regressor, then $x_{it} = O_p(\sqrt{T})$ and its functionals belong to $L_2(\mathbb{R})$. Since the functionals of the near-unity regressors belong to $L_2(\mathbb{R})$, we can apply the Hermite basis, $p_j(\cdot)$ with $j \in [k]$ of (4), to approximate these functionals. In addition, the functional space $L_2(\mathbb{R})$ is large enough to incorporate any square-integrable functions, allowing for the smooth functions of stationary mixingales, which are defined on $L_2(\mathcal{X}, \partial F)$[5] and have bounded supports.

Based on the above discussions, the Hermite basis is flexible enough to include stationary mixingales and near-unity regressors, proposing unified machinery to obtain sieve basis expansions, though the mixed-root regressor that demonstrates asymptotic heterogeneity is considered in the present framework. Also, the Hermite basis function of (5) provides a unified representation system, enabling a unified inference method that stays robust to mixed-root regressors. In what follows, we will apply the Hermite basis functions to obtain a unified theory of coupling that relies on this unified representation of the Hermite basis functions.

[5] $\partial F(x)$ is the partial differential of the CDF in terms of x.

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

Also, we apply the Hermite basis of the space $L_2(\mathbb{R})$ to all cases in our simulation and empirical studies.

To measure the accuracy of sieve approximations, we define the approximation error by $r_g(x) := g(x) - p(x)'\beta$ and denote the approximation error for $i \in [n]$ and $t \in [T]$ by

$$r_{it} := r_g(x_{it}) = g(x_{it}) - p'_{it}\beta. \tag{7}$$

Based on the sieve approximation, we then attain a linear panel data regression model as

$$y_{it} = \gamma_i + p'_{it}\beta + u_{it}, \;\; u_{it} := r_{it} + \epsilon_{it}. \tag{8}$$

We define the demeaned variates as $\widetilde{y}_{it}(:= y_{it} - \mathbb{E}^*_T[y_{is}])$ and $\widetilde{p}_{it}(:= p_{it} - \mathbb{E}^*_T[p_{is}])$. Then, the demeaned version of equation (8) follows

$$\widetilde{y}_{it} = \widetilde{p}'_{it}\beta + \widetilde{u}_{it}, \;\; \widetilde{u}_{it} := u_{it} - \mathbb{E}^*_T[u_{is}]. \tag{9}$$

Correspondingly, the least squares estimator $\widehat{\beta}$ of the pseudo-true parameter β is given by

$$\widehat{\beta} := \arg\min_{b\in\mathbb{R}^k} \mathbb{E}_n\left[\mathbb{E}^*_T\left[\left(\widetilde{y}_{it} - \widetilde{p}'_{it}b\right)^2\right]\right] = \widehat{Q}^{-1} \cdot \mathbb{E}_n\left[\mathbb{E}_T\left[\widetilde{p}_{it}\widetilde{y}_{it}\right]\right], \tag{10}$$

where $\widehat{Q} := \mathbb{E}_n[\mathbb{E}_T[\widetilde{p}_{it}\widetilde{p}'_{it}]]$. The least squares estimator $\widehat{\beta}$ of (10) induces the sieve estimator

$$\widehat{g}(x) := p(x)'\widehat{\beta}, \tag{11}$$

for the $g(x)$ on $x \in \mathcal{X}$. Then, by (7), we decompose the error in estimating the $g(\cdot)$ as

$$\widehat{g}(x) - g(x) = p(x)'\left(\widehat{\beta} - \beta\right) - r_g(x),$$

in which the first term $p(x)'(\widehat{\beta} - \beta)$ is the estimation error that generates a strong Gaussian approximation in the L_∞-space, which will be discussed in Section 3, and the second term $r_g(x)$ is the approximation error.

2.3. Approximation Properties of Sieve Estimator

Next, we introduce the approximation properties of the sieve estimator $\widehat{g}(x)$. To facilitate our discussions, we define a uniform upper bound of the basis function:

$$\tau_k := \sup_{x\in\mathcal{X}} \|p(x)\|.$$

With stationary mixingale regressors (Assumption 2), we consider the basis functions of Belloni et al. (2015). Examples of their upper bounds include $\tau_k \lesssim k$ when polynomial series are considered, and $\tau_k \lesssim k^{1/2}$ when spline series and CDV wavelets are considered. With mildly stationary or local-to-unity regressors

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

(Assumption 3), we use the Hermite basis (Dong and Gao, 2019) with their upper bound $\tau_k \lesssim k^{1/2}$ since

$$\sup_{x\in\mathbb{R}} \|p(x)\| \lesssim \sup_{x\in\mathbb{R}} \left\|H(x)\cdot \exp(-x^2/2)\right\| \lesssim k^{1/2},$$

ensured by Dong et al. (2016, Thm. 3.3). Similar to Belloni et al. (2015, Assumption A.3), we state the properties of basis functions as follows.

Assumption 5. For each n, T, and k, there are c_k and ℓ_k such that for each $g \in \mathcal{G}$,

$$\left\|r_g\right\|_{\mu,2} \le c_k, \text{ and } \sup_{x\in\mathcal{X}} \left|r_g(x)\right| \le \ell_k c_k,$$

in which c_k denotes the approximation error in terms of the L_2 distance ($L_2(\mathcal{X}, \partial F)$ for stationary case and $L_2(\mathbb{R})$ for the near-unity case) and $\ell_k c_k$ models the approximation error concerning the uniform distance.

We further impose the following regularity conditions on the signal matrix to facilitate discussions of asymptotic developments in estimation and inference.

Assumption 6. We assume that all eigenvalues of Q are bounded above from infinity and below from zero, uniformly over k:

(i) When Assumption 2 holds, it follows that $Q(:= \mathbb{E}[(p_{it} - \mathbb{E}[p_{it}])(p_{it} - \mathbb{E}[p_{it}])'])$.
(ii) When Assumption 3 and $C = \mathbf{0}_{d\times d}$ hold, it follows that $Q := \sqrt{2/(\pi\sigma_x^2)}$ $(\int_{-\infty}^{\infty} p(x)p(x)'dx)$.
(iii) When Assumption 3, $\zeta = 1$, and $C \neq \mathbf{0}_{d\times d}$ hold, it follows that $Q := \mathbb{E}[L_{c,i}(1,0)](\int_{-\infty}^{\infty} p(x)p(x)'dx)$, in which $L_{c,i}(\cdot,\cdot)$ is the local time for the Ornstein–Uhlenbeck process, $dJ_{c,i}(r) = CJ_{c,i}(r)dr + \sigma_x dW_{x,i}(r)$ with $W_{x,i}(r)$ being the standard Brownian motion driven by $u_{x,it}$.
(iv) When Assumption 3 and $\zeta \in (0,1)$ hold, it follows that $Q := \sqrt{-C/(\pi\sigma_x^2)}$ $(\int_{-\infty}^{\infty} p(x)p(x)'dx)$.

Assumption 6 eliminates the case in which $p_1(x_{it}), p_2(x_{it}), \ldots, p_k(x_{it})$ are co-linear and ensures the invertibility of signal matrix, Q. Assumption 6(i) considers stationary mixingale. Conditions of Assumption 6(ii) and 6(iii) discuss the case in which the Hermite basis functions apply to the unity and local-to-unity processes. Assumption 6(iv) discusses the mildly stationary process. In summary, Assumption 6 ensures that the signal matrix Q can be consistently estimated when x_{it} follows an arbitrary type of time-series dependence. Some theoretical results on approximating the population term Q via the sample moment $\widehat{Q}$ are provided in Section C.i of the Supplementary Material.

Then, we introduce the consistency of our sieve estimator in the L_∞ space. Define the normalized basis $\alpha(x) := p(x)/\|p(x)\|$. Define $\tau_k^L := \sup_{x,x'\in\mathcal{X}:x\neq x'} \|\alpha(x) - \alpha(x')\|/\|x - x'\|$. Then, we can have a uniform linearization that follows Belloni et al. (2015, Thm. 4.3):

$$\sqrt{n}a_T\alpha(x)'\left(\widehat{\beta} - \beta\right) = \alpha(x)'Q^{-1}\mathbb{G}_n[\mathbb{G}_T[\widetilde{p}_{it}\epsilon_{it}]] + R_{1,nT}(\alpha(x)) + R_{2,nT}(\alpha(x)), \quad \textbf{(12)}$$

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

in which $R_{1,nT}(\alpha(x))$ is bounded by

$$R_{1,nT}(\alpha(x)) \lesssim_p \left(\sqrt{\frac{\tau_k^2 \log k}{n}} + \frac{1}{k}\right)\left((nT)^{1/m}\sqrt{\log k} + \sqrt{k}\cdot \ell_k c_k\right) =: \overline{R}_{1,nT}, \tag{13}$$

and the remainder $R_{2,nT}(\alpha(x)) \lesssim_p \sqrt{k}\cdot \ell_k c_k =: \overline{R}_{2,nT}$, uniformly over $x \in \mathcal{C}$. Based on the representation in (12), we derive the uniform convergence rate of sieve estimation:

$$\sup_{x\in\mathcal{X}} |\widehat{g}(x) - g(x)| \lesssim_p \frac{\tau_k}{\sqrt{n}\cdot a_T}\left(\sqrt{\log k} + \overline{R}_{1,nT} + \overline{R}_{2,nT}\right) + \ell_k c_k. \tag{14}$$

The expressions of (12) demonstrate the empirical process of the leading term to which we apply the strong approximation, which will be detailedly introduced in the next section.

3. UNIFORM LIMIT THEORY

After giving the uniform linear approximation of $\widehat{g}(\cdot)$, we apply Yurinskii's coupling principle (Belloni et al., 2015) to show its strong Gaussian approximation. Before introducing the details of coupling, we first define variance matrices.

When the covariates $\{x_{it}\}$ follow stationary mixingales (Assumption 2), we define the population variances of different instances, $\widetilde{\Omega}(:= Q^{-1}\widetilde{\Sigma}Q^{-1})$ and $\Omega_0(:= Q^{-1}\Sigma_0 Q^{-1})$, where

$$\widetilde{\Sigma} := \sum_{j=-\infty}^{\infty} \mathbb{E}\Big[(\epsilon_{it} + r_{it})\,(p_{it} - \mathbb{E}[p_{it}])\left(p_{i,t-j} - \mathbb{E}\left[p_{i,t-j}\right]\right)'\left(\epsilon_{i,t-j} + r_{i,t-j}\right)\Big], \tag{15}$$

$$\Sigma_0 := \sum_{j=-\infty}^{\infty} \mathbb{E}\Big[(p_{it} - \mathbb{E}[p_{it}])\left(p_{i,t-j} - \mathbb{E}\left[p_{i,t-j}\right]\right)'\epsilon_{it}\epsilon_{i,t-j}\Big]. \tag{16}$$

When the regressor $\{x_{it}\}$ follows a near-unity process (Assumption 3), we similarly define the variances, $\widetilde{\Omega}(:= Q^{-1}\widetilde{\Sigma}Q^{-1})$ and $\Omega_0(:= Q^{-1}\Sigma_0 Q^{-1})$, in which

$$\widetilde{\Sigma} := \lim_{(n,T)\to\infty} \mathbb{E}_n\left[\sum_{j=-\infty}^{\infty} \mathbb{E}_T\left[\widetilde{p}_{it}\widetilde{p}'_{i,t-j}\,(\epsilon_{it} + r_{it})\left(\epsilon_{i,t-j} + r_{i,t-j}\right)\right]\right], \tag{17}$$

$$\Sigma_0 := \lim_{(n,T)\to\infty} \mathbb{E}_n\left[\sum_{j=-\infty}^{\infty} \mathbb{E}_T\left[\widetilde{p}_{it}\widetilde{p}'_{i,t-j}\epsilon_{it}\epsilon_{i,t-j}\right]\right], \tag{18}$$

for any $k \in \mathbb{N}$. To be specific, the validity of the joint limit in (17) and (18) follows the joint convergence in Phillips and Moon (1999, Lemma 5): for a fixed $k \in \mathbb{N}$, when $(n,T) \to \infty$,

$$\mathbb{E}_n\left[\sum_{j=-\infty}^{\infty} \mathbb{E}_T\left[\widetilde{p}_{it}\widetilde{p}'_{i,t-j}\,(\epsilon_{it} + r_{it})\left(\epsilon_{i,t-j} + r_{i,t-j}\right)\right]\right] \to_p \Sigma_0,$$

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

where the sequential limit theory ($T \to \infty$ then $n \to \infty$) coincides with the joint limit.

Though the covariates $\{x_{it}\}$ under Assumptions 2 and 3 differ significantly in the limit theory, our work intends to provide a unified system of representations and test statistics for all considered scenarios, facilitating a unified inference procedure that remains robust to the presence of nonstationary roots. In particular, based on (12), we can attain a unified theory of Gaussian approximation for the nonparametric estimator.

THEOREM 3.1 (Strong Approximation). *Suppose that Assumptions 1, 2 (or 3), 4, 5, and 6 hold. In addition, we assume the conditions that (i)* $\overline{R}_{1,nT} = o_p(c_{nT}^{-1})$, *(ii)* $1 \lesssim \underline{\sigma}^2$, *and (iii)* $c_{nT}^6 a_T^6 k^4 \tau_k^2 (1 + \ell_k^3 c_k^3)^2 \log^2 n/n \to 0$. *Then, for some* $\mathcal{N}_k \sim \mathcal{N}(\mathbf{0}_{k\times 1}, I_k)$ *and* c_{nT} *as some sequence of positive numbers satisfying* $c_{nT} \to \infty$, *if* $\sup_{x\in\mathcal{X}} \sqrt{n} a_T |r_g(x)|/\|s(x)\| = o(c_{nT}^{-1})$ *with* $s(x) = \Omega^{1/2} p(x)$, *then*

$$\sqrt{n} a_T \cdot \frac{\widehat{g}(x) - g(x)}{\|s(x)\|} = \frac{s(x)'}{\|s(x)\|} \mathcal{N}_k + o_p(c_{nT}^{-1}), \quad \text{in } L_\infty(\mathcal{X}), \tag{19}$$

where we set $\Sigma = \widetilde{\Sigma}$ *and* $\Omega = \widetilde{\Omega}$. *If* $\overline{R}_{2,nT} = o_p(c_{nT}^{-1})$, *then we can set* $\Sigma = \Sigma_0$ *and* $\Omega = \Omega_0$.

To address uniform inference of $g(x)$ over $x \in \mathcal{C}$ with $\mathcal{C}$ being any compact subset of $\mathcal{X}$, we rely on the following strong Gaussian approximation implied by (19):

$$\left| \sqrt{n} a_T \frac{\widehat{g}(x) - g(x)}{\|s(x)\|} - \frac{s(x)'}{\|s(x)\|} \mathcal{N}_k \right| = o_p(c_{nT}^{-1}) \text{ in } L_\infty(\mathcal{C}).$$

Since $c_{nT} \to \infty$, the sieve estimator is asymptotically approximated by a Gaussian process whose distribution only depends on the data by the covariance matrix Ω. Therefore, it is feasible to perform uniform inference procedures using the bootstrapped critical values, which are further discussed in Section 5.

Remark 2. Consider the score vector $S_{nT} (:= \mathbb{G}_n[\mathbb{G}_T[\Upsilon_{it}]])$, where $\Upsilon_{it} = \widetilde{p}_{it}\epsilon_{it}$ in the context of panel data. This article extends Yurinskii's coupling to the fixed-effects panel models by "stacking" the time-series units into a time-series average, $\Upsilon_{i\cdot}$,

$$S_{nT} = \mathbb{G}_n[\Upsilon_{i\cdot}], \quad \text{with } \Upsilon_{i\cdot} = \mathbb{G}_T[\Upsilon_{it}], \tag{20}$$

and then applying Yurinskii's coupling to the cross-sectional averaging that involves the generated variable, $\Upsilon_{i\cdot}$. Guided by this logic, we extend Yurinskii's coupling principle to a widely used panel setting and use it to construct the UCBs for the unknown function $g(\cdot)$. Then, the strong approximation of the score vector $S_{nT} (:= \mathbb{G}_n[\Upsilon_{i\cdot}])$ is given by

$$\left\| S_{nT} - \widetilde{S}_{nT} \right\| = o_p(\delta_n), \quad \text{with } S_{nT} = \mathbb{G}_n[\Upsilon_{i\cdot}], \tag{21}$$

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

for $\delta_n \to 0$ with an approximating variable $\widetilde{S}_{nT}$ having a Gaussian $\mathcal{N}(\mathbf{0}_{k\times 1}, \Sigma_{nT})$ distribution with some $k \times k$ covariance matrix. In summary, Yurinskii's coupling principle utilizes the independence across the generated variables $\{\Upsilon_{i\cdot}\}$ to obtain a Gaussian approximation in the nonparametric panel data model with fixed effects.

This machinery is closely related to a recent paper, Li et al. (2024), in which the mixingale coupling principle developed in Li and Liao (2020) applies to the nonparametric panel model by "stacking" the cross-sectional units into a cross-sectional average, $\{\Upsilon_{\cdot t}\}$,

$$S_{nT} = \mathbb{G}_T[\Upsilon_{\cdot t}], \quad \text{with} \quad \Upsilon_{\cdot t} = \mathbb{G}_n[\Upsilon_{it}], \tag{22}$$

and applying the coupling of Li and Liao (2020) to the time-series averaging that involves the generated variable, $\{\Upsilon_{\cdot t}\}$. As the coupling of Li et al. (2024) is driven by time-series averaging and no CLTs are applied to the cross-sectional dimension, any arbitrary type of cross-sectional dependence can be incorporated. However, the unit root or local-to-unity process cannot be accommodated in their framework. Therefore, the present article contributes to the literature on uniform inference by filling the void and discussing the case of near unit root regressors in the context of panel data.

Remark 3. Regarding the nature of our coupling principle that makes use of the iid cross-sectional individuals and the fact that no time-series CLTs are applied, the intensity of intertemporal correlations is unrestricted, and a large class of time-series dependence structures is allowable within our inference framework. Examples include the stationary mixingale (Li and Liao, 2020), mildly stationary, unit root, and local-to-unity processes. Since stationary mixingales are used in the analysis of panel data (Hahn, Kuersteiner, and Mazzocco, 2022; Li et al., 2024), mildly stationary processes are applied to the panel analysis of bubbles and local-to-unity roots belong to one of the main workhorses for detecting predictability, our uniform inference method has the full generality of incorporating various types of time-series dependence, enabling its wide applicability.

Remark 4. Even though Assumptions 2 and 3 separately addressed the setting of stationary mixingale regressors and the case of a near-unity regressor, our coupling principle in Theorem 3.1 and uniform testing method that will be introduced later provide a unified scheme that allows for both cases. Such arrangements emphasize the novelty of our uniformity results to include near-unity regressors, which we believe are first mentioned in the literature. In particular, this article gives a specific type of basis functions (Hermite polynomials) that suit the unbounded support of near-unity regressors and examines their approximation characteristics for square-integrable nonparametric functional forms. Meanwhile, our work also shows that when all the regressors are confirmed to be stationary mixingales, a large class of basis functions still works (Chen, 2007), demonstrating the generality of our framework. It is worth mentioning that when no prior information is available for the persistence properties of regressors, Hermite polynomials

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

can be applied without costs, revealing its robustness in consistently estimating and uniformly testing the nonparametric quantity of interest for mixed-root regressors.

Even though the stationary and near-unity regressors are examined separately under parallel conditions, mixed-root regressors with various degrees of persistence are allowed in our uniform inference method, which will be introduced later. Specifically, Section C of the Supplementary Material and Remark 7 of the main article introduce three concrete examples to allow for mixed-root regressors: nonseparable, additive, and partial linear panel models. The nonseparable model uses the tensor-form basis functions, while the additive and panel linear models rely on the basis functions of additive forms (see (C.7) and (C.14) of the Supplementary Material for more details). For the nonseparable, additive, and partial linear models, when the degrees of persistence in regressors are unobserved, we can construct their tensor-form and additive-form basis functions by relying on the Hermite polynomials. In addition, when our models contain one unity and multiple stationary regressors, our tensor-form and additive-form sieve approximations can still guarantee the uniform limit theory of our estimation method and the uniform probability coverage of our uniform testing procedure.

In all, our uniform limit theory and uniform testing procedure allow for very general and flexible square-integrable functional forms (of $L_2(\mathcal{X}, \partial F)$ or $L_2(\mathbb{R}^d)$) that include the nonseparable, additive, and partial linear models as examples. Also, the mixed-root setting of one near-unity and multiple stationary mixingale regressors can be naturally included. The setup with multiple unity regressors in the nonseparable functional forms can be challenging since the underlying cointegrating relationship disturbs our asymptotic development. Further developments for multiple unit-root regressors are left for future research due to this potential complication.

4. UNIFORM INFERENCE FOR LINEAR FUNCTIONALS

This section derives UCBs for linear functionals $\theta(x)$ of $g(x)$, which is approximated as

$$\theta(x) = l_\theta(x)'\beta + r_\theta(x), \quad \text{for } x \in \mathcal{X}, \tag{23}$$

in which $l_\theta(x)'\beta$ denotes a sieve approximation and $r_\theta(x)$ represents the approximation error. Examples of $\theta(x)$ that were discussed in Belloni et al. (2015) are also provided as:

(i) an original function: $\theta(x) = g(x)$, so that $l_\theta(x) = p(x)$ and $r_\theta(x) = r_g(x)$;
(ii) a differential functional: $\theta(x) = \partial^j g(x)$, so that $l_\theta(x) = \partial^j p(x)$ and $r_\theta(x) = \partial^j r_g(x)$.

We define the plug-in nonparametric sieve estimator $\widehat{\theta}(x) := l_\theta(x)'\widehat{\beta}$ in which $\widehat{\beta}$ is given in (10). To further construct UCBs for the functional, $\theta(x)$, we define

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

several estimates for the signal and variance matrices:

$$\widehat{Q} = \mathbb{E}_n[\mathbb{E}_T[\widetilde{p}_{it}\widetilde{p}_{it}']],\ \widehat{\Omega} = \widehat{Q}^{-1}\widehat{\Sigma}\widehat{Q}^{-1},$$
$$\widehat{\Sigma} = \begin{cases} \mathbb{E}_n\left[\mathbb{G}_T\left[\widetilde{p}_{it}\widehat{\widetilde{\epsilon}}_{it}\right]\cdot\mathbb{G}_T\left[\widetilde{p}_{it}'\widehat{\widetilde{\epsilon}}_{it}\right]\right] & \text{when} \quad T \text{ is small,} \\ \mathbb{E}_n\left[\sum_{j=-(M-1)}^{(M-1)} K\left(\frac{j}{M}\right)\mathbb{E}_T\left[\widetilde{p}_{it}\widetilde{p}_{i,t-j}'\widehat{\widetilde{\epsilon}}_{it}\widehat{\widetilde{\epsilon}}_{i,t-j}\right]\right] & \text{when} \quad T \to \infty, \end{cases} \tag{24}$$

where $\widehat{\widetilde{\epsilon}}_{it} := \widetilde{y}_{it} - \widetilde{p}_{it}'\widehat{\beta}$ and $K(\cdot)$ denotes the kernel weight. In what follows, we consider the following restrictions on the kernel function $K(\cdot)$ and the bandwidth parameter M.

Assumption 7. For the bandwidth parameter M, the following restriction is satisfied:

$$M\sqrt{T^{2/m}\tau_k^2 \log k/(n^{1-2/m})} \to 0.$$

The kernel function $K(\cdot)$ is a twice continuously differentiable even function with (i) $K(0) = 1$, $\partial K(0) = 0$, $\partial^2 K(0) \neq 0$; (ii) $\int_{-1}^{1} K(r)dr = 1$; and (iii) $K(x) = 0$, when $|x| > 1$ with $\lim_{|x|\to 1} K(x)/(1-|x|)^2 < \infty$.

We can show that the variance estimate $\widehat{\Omega}$ and the sample moment $\widehat{Q}$ consistently estimate the population variance and signal matrices of increasing dimensions, as provided in Proposition A.1 of the Supplementary Material.

Remark 5. One major contribution of this work is the allowance for unit-root regressors whose asymptotic properties are only valid in the large-T setting. To our best knowledge, the uniform nonparametric inference method for panel data models with persistent regressors was not discussed in the literature, and our results belong to the first attempt. Accordingly, the long-run variance estimates are needed to stabilize the test statistics in this case since the simple variance estimates cannot work for the long-panel case (Newey and West, 1994). To facilitate the application of our uniform testing method in empirical studies, we propose a simple rule to select which type of variance estimate should be used. Given our simulation studies, when $T \geq 10$, we suggest using the Newey–West variance estimate. However, the simple variance estimation needs to be considered for a smaller sample size.

Based on the estimates of (24), we introduce the variance estimate for the plug-in sieve estimator $\widehat{\theta}(x) := l_\theta(x)'\widehat{\beta}$. In particular, we define the variance estimator $\widehat{\sigma}_\theta^2(x)$ for $\widehat{\theta}(x)$,

$$\widehat{\sigma}_\theta^2(x) = l_\theta(x)'\widehat{\Omega}l_\theta(x)/(na_T^2), \tag{25}$$

which estimates $\sigma_\theta^2(x) := l_\theta(x)'\Omega l_\theta(x)/(na_T^2)$ with $\widehat{\Omega}$ of (24). Based on the matrix LLN for $\widehat{\Omega}$ in Proposition A.1 of the Supplementary Material, we show the consistency of $\widehat{\sigma}_\theta^2(\cdot)$:

$$|\widehat{\sigma}_\theta(x)/\sigma_\theta(x) - 1| = o_p(1/\log(k)),$$

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

uniformly over $x \in \mathcal{C}$. Then, we provide the following t-statistic process as

$$\left\{ t_\theta(x) = \frac{\widehat{\theta}(x) - \theta(x)}{\widehat{\sigma}_\theta(x)}, \ \text{ for } x \in \mathcal{C} \right\}, \tag{26}$$

and $\widehat{\theta}(x)$ is a sieve estimator. To have UCBs, we define the uniform upper bounds $\tau_{k,\theta} := \sup_{x \in \mathcal{X}} \|l_\theta(x)\|$, $\tau^L_{k,\theta} = \sup_{x,\widetilde{x} \in \mathcal{X}: x \neq \widetilde{x}} \|l_\theta(x) - l_\theta(\widetilde{x})\| / \|x - \widetilde{x}\|$ and introduce Assumption 8.

Assumption 8. The following bounds hold: (i) $\sup_{x \in \mathcal{X}} 1/\|l_\theta(x)\| \lesssim 1$ and (ii) $\log \tau^L_{k,\theta} \lesssim \log k$.

We apply the sieve score bootstrap (Chen and Christensen, 2018) to compute critical values so that the UCBs can attain uniform probability coverage over $\mathcal{C}$. In what follows, we introduce the sieve score bootstrap method and prove its uniform size controls for the cases of stationary mixingale and near-unity regressors imposed by Assumptions 2 and 3.

Let $\{\omega_i\}_{i \in [n]}$ be iid standard Gaussian variates ($\omega_i \sim \mathcal{N}(0,1)$) which are independent of data $(y_{it}, x'_{it}, \epsilon_{it})'$. Then, we define the sieve score bootstrapping t-statistic process as

$$t^b_\theta(x) := \frac{l_\theta(x)' \left(\mathbb{E}_n[\mathbb{E}_T[\widetilde{p}_{it}\widetilde{p}'_{it}]] \right)^{-1} \left\{ \mathbb{G}_n \left[\omega_i \cdot \mathbb{G}_T \left[\widetilde{p}_{it} \widehat{\widetilde{\epsilon}}_{it} \right] \right] \right\}}{\widehat{\sigma}_\theta(x)}, \tag{27}$$

for each $x \in \mathcal{C}$ with $\widehat{\widetilde{\epsilon}}_{it}$ defined in calculating (24). To compute the critical values $cv^b_n(1-\alpha)$ of UCBs, one calculates $\sup_{x \in \mathcal{C}} |t^b_\theta(x)|$ based on a large number of iid draws of $\{\omega_i\}_{i \in [n]}$ which are independent from data. Therefore, the critical value $cv^b_n(1-\alpha)$ can be constructed by the $(1-\alpha)$ quantile of $\sup_{x \in \mathcal{C}} |t^b_\theta(x)|$ over the multiple random draws of the perturbation processes $\{\omega_i\}_{i \in [n]}$, as

$$cv^b_n(1-\alpha) := (1-\alpha) \text{ quantile of } \sup_{x \in \mathcal{C}} |t^b_\theta(x)| \text{ given the data and } \{\omega_i\}_{i \in [n]}. \tag{28}$$

Further, we can develop the UCBs for the linear functional $\theta(x)$ by

$$\left[\widehat{\theta}^b_l(x), \widehat{\theta}^b_u(x) \right] = \left[\widehat{\theta}(x) - cv^b_n(1-\alpha) \cdot \widehat{\sigma}_\theta(x), \ \widehat{\theta}(x) + cv^b_n(1-\alpha) \cdot \widehat{\sigma}_\theta(x) \right], \ \text{ for } x \in \mathcal{C}, \tag{29}$$

in which $cv^b_n(1-\alpha)$ ensures that $\theta(x) \in [\widehat{\theta}^b_l(x), \widehat{\theta}^b_u(x)]$ for all $x \in \mathcal{C}$ with the asymptotic confidence level $100(1-\alpha)\%$ for both the cases of the stationary mixingale and near unit root regressors imposed by Assumptions 2 and 3.

To facilitate the empirical applications of the bootstrapping UCBs for linear functionals, we detail its construction in the following algorithm.

Algorithm 1 (UCBs for Linear Functionals).

(i) Step 1: Compute the nonparametric sieve estimator $\widehat{\theta}(\cdot)$ in Section 4 of the main article; Compute the sample moment and variances, $\widehat{Q}$, $\widehat{\Sigma}$, and $\widehat{\Omega}$ provided in (24).

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

(ii) Step 2: Draw the iid standard Gaussian variates $\{\omega_i\}_{i\in[n]}$ which are independent of data $(y_{it}, x'_{it}, \epsilon_{it})'$ and simulate the bootstrap sup-test $\sup_{x\in\mathcal{C}} |t^b_\theta(x)|$ with $t^b_\theta(x)$ of (27).
(iii) Step 3: Repeat Step 2 for B times and calculate the $(1-\alpha)$ quantile of the sup-test, $\sup_{x\in\mathcal{C}} |t^b_\theta(x)|$; Report the $(1-\alpha)$ confidence level two-sided UCB of (29).

The uniform size controls of the proposed UCBs are proved in the following theorem.

THEOREM 4.1 (Uniform Inference of Sieve Score Bootstrap). *Suppose that Assumptions 1, 2 (or 3), 4, 5, 6, 7, and 8 and other conditions of Theorem 3.1 hold. Assume that $cv^b_n(1-\alpha)$ is constructed as in (28). Then, the confidence bands defined in (29) satisfy*

$$\Pr\left\{\theta(x) \in \left[\widehat{\theta}^b_l(x), \widehat{\theta}^b_u(x)\right], \text{for all } x \in \mathcal{C}\right\} = 1-\alpha+o(1). \tag{30}$$

Theorem 4.1 applies the sieve score bootstrap in Chen and Christensen (2018) to the fixed-effects nonparametric panel data model that incorporates either the stationary mixingale or near unit root process and proves the validity of our proposed UCBs. In such a case, though the unit root process that can result in nonstandard distributions exists, we can still apply the cross-sectional averaging to retain the Gaussian approximation and prove the uniform probability coverage of the sieve score bootstrap UCBs.

Remark 6. The uniform testing method that involves the unit root process requests a strong approximation theory to the empirical process of the test statistic. However, the sample moment and covariance generated from the nonstationary time series (Phillips, 1987) conventionally call upon the invariance principle and random weak limits, which cannot be compatible with the matrix LLN and coupling principles. To deal with the challenge that arises from the nonstandard limit theory of the unit root process, we use panel data and obtain the standard Gaussian limit theory by focusing on the additional averaging over cross-sectional individuals. That is, we treat the scaled time-series averaging as the generated variables and only count on the high-dimensional CLT over i. In this manner, Yurinskii's coupling and the matrix LLN can still be applied to the present context.

Accordingly, the sieve score bootstrap method of Chen and Christensen (2018) is modified to accommodate the fixed-effects nonparametric panel model of (1). The underlying intuition of the sieve score bootstrap is randomizing the independent component (generated variable) by the perturbation process to recover the uniform null distribution of the test statistic. As mentioned above, the independent generated variable of the current context is the time-series averaging component. So, our proposed bootstrapping method randomizes the time-series averages rather than the unit-specific element for each i and t. Therefore, the article makes another contribution to the literature by showing that the modified sieve score bootstrap method can help attain the uniform probability coverage for the nonparametric

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

exogenous shocks in both the conventional individual fixed-effects panel model and the two-way fixed-effects panel model that is discussed in Section 5.

Remark 7. When the dimension of regressors is large (d is large), the efficiency loss and computational burdens of the nonparametric estimation method increase dramatically. This is the so-called "curse of dimensionality." In the literature, Stone (1980), which gave the minimax optimal rate of the nonparametric estimator, addressed this issue. We quote their main finding for illustration: Define (X, Y) to be iid with $X \in \mathbb{R}^d$ valued and Y being scalar. Stone (1980) showed that $n^{-s/(2s+d)}$ is the L_2 minimax optimal rate for estimators of regression functions which rely on a random sample of size n. Here, the quantity s denotes the order of smoothness of regression functions. It is evident that as d (the number of regressors) increases, the optimal rate will be monotonically reduced. This rate shows the negative effects of an increasing number of regressors on the optimal convergence rate. This result also holds for all nonparametric/machine-learning methods and can be intentionally extended to panel data.

One possible solution strategy to resolve this issue is introducing the semi-nonparametric approach, which relies on the partial linear panel model as

$$y_{it} = \gamma_i + g(x_{it}) + z_{it}'\eta + \epsilon_{it}, \quad \mathbb{E}[\epsilon_{it}|x_{it}, z_{it}] = 0, \tag{31}$$

in which x_{it} is the one-dimensional regressor of interest and z_{it} is a d-dimensional control variable. To help improve the estimation efficiency, we put the control variates into the parametric part while investigating the nonlinear effects of the independent variable of interest. We apply the sieve expansion to $g(\cdot)$ and obtain the approximate linear model:

$$y_{it} = \gamma_i + p_{it}'\beta + z_{it}'\eta + u_{it}, \quad u_{it} := r_{it} + \epsilon_{it}. \tag{32}$$

Then, we derive the least squares estimators $\widehat{\beta}$ and $\widehat{\eta}$ and obtain the nonparametric estimator $\widehat{g}(\cdot)$. Further steps for constructing the bootstrapping UCBs follow Algorithm 1, and the uniform size controls can be shown by Theorem 4.1. It is also worth mentioning that the semi-nonparametric bootstrapping UCB for (31) is a special case of the uniform inference procedure in the additive nonparametric panel model given in Section C.ii of the Supplementary Material. Simulation studies examine both the cases of a partial linear functional form and of an additive functional form and show the robustness of our proposed bootstrapping UCBs, which can help attain uniform probability coverage. In addition, the empirical analysis also applies both the methods that rely on the partial linear form and the additive functional form to the real data.

Remark 8. This remark discusses the choice of the tuning parameter k. Our uniform inference procedure relies on uniform asymptotics, with a rate of convergence being derived in terms of the L_∞ norm. Comparatively, the cross-validation criterion approximates the mean squared error (MSE) that balances the squared bias and variance components. This data-driven method works by relying on the L_2 norm and can be treated as a valid model selection method of tuning parameters

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

in pointwise asymptotics rather than uniform asymptotics. Given this reason, the cross-validation method may not be a suitable candidate for our uniform inference framework.

The uniform convergence rate of our article can be challenging to approximate since the exact approximation error is difficult to evaluate. So, the number of sieve basis functions cannot be directly estimated given this theoretical result. As an alternative, this article uses simulation studies to search for a suitable choice that yields the desirable finite-sample performance. Simulation results show that once the number of basis functions is bigger than 6, the finite-sample probability coverage of our UCBs can be close to the nominal level. This choice of basis functions will also be applied to our empirical studies, which will be introduced later.

Similarly, the related discussions in Li and Liao (2020) and Li et al. (2024) rely on simulation findings to choose the number of basis functions and show the robustness of this choice. So, their results corroborate the feasibility of our rule of selecting tuning parameters.

5. EXTENSIONS TO TWO-WAY FIXED-EFFECTS PANEL MODEL

Econometricians widely apply the two-way fixed-effects panel for estimating causal effects in empirical work, mainly for two reasons: first, by including the fixed effects in both time and cross-sectional dimensions, we can control unobserved confounders and remove endogeneity; second, the two-way fixed-effects panel model reduces to the difference-in-difference panel model under the simplest setting with two groups and two time periods.

This section further extends the uniform inference procedure to the case of the nonparametric panel model with two-way fixed effects:

$$y_{it} = \gamma_i + f_t + g(x_{it}) + \epsilon_{it}, \ \ \mathbb{E}\left[\epsilon_{it} | \{x_{it}\}_{t\in[T]}\right] = 0, \ \ i \in [n] \text{ and } t \in [T], \tag{33}$$

in which x_{it}, ϵ_{it}, and γ_i follow the setup of (1) and f_t denotes the time fixed effect that is correlated with regressor x_{it}. Different from the setup of the model that only includes the individual fixed effects, we allow for a large panel in which both n and T diverge.

Based on the sieve expansion, we similarly obtain a linear panel data model as

$$y_{it} = \gamma_i + f_t + p_{it}'\beta + u_{it}, \ \ u_{it} := r_{it} + \epsilon_{it}. \tag{34}$$

With a slight abuse of notation, we define a new approach of demeaning as

$$\widetilde{y}_{it} := y_{it} - \mathbb{E}_T^*[y_{is}] - \mathbb{E}_n[y_{is}] + \mathbb{E}_n[\mathbb{E}_T^*[y_{is}]], \tag{35}$$

$$\widetilde{p}_{it} := p_{it} - \mathbb{E}_T^*[p_{is}] - \mathbb{E}_n[p_{is}] + \mathbb{E}_n[\mathbb{E}_T^*[p_{is}]], \tag{36}$$

$$\widetilde{u}_{it} := y_{it} - \mathbb{E}_T^*[u_{is}] - \mathbb{E}_n[u_{is}] + \mathbb{E}_n[\mathbb{E}_T^*[u_{is}]]. \tag{37}$$

Then, we follow the identification condition of Assumption 4 and compute the least squares estimate $\widehat{\beta}$ of (10) and the nonparametric estimator $\widehat{g}(\cdot)$ of (11) by

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

using the demeaned variates $\widetilde{y}_{it}$, $\widetilde{x}_{it}$, and $\widetilde{\epsilon}_{it}$ defined in (35)–(37). To construct the UCBs for linear functionals $\theta(x)$ of $g(x)$, we also compute the sample moment and variance estimates $\widehat{Q}$, $\widehat{\Sigma}$, and $\widehat{\Omega}$ based on $\widetilde{y}_{it}$ and $\widetilde{x}_{it}$ of (35) and (36) and the residual estimate $\widehat{\widetilde{\epsilon}}_{it} := \widetilde{y}_{it} - \widetilde{p}'_{it}\widehat{\beta}$. Then, we can attain the variance estimate $\widehat{\sigma}^2_\theta(x)$ based on (25). To implement the sieve score bootstrap method, we calculate the simulated empirical process of (27) and take $cv^b_n(1-\alpha)$ in (28) to build the UCBs with the $100(1-\alpha)\%$ confidence level

$$\left[\widehat{\theta}^b_l(x), \widehat{\theta}^b_u(x)\right] = \left[\widehat{\theta}(x) - cv^b_n(1-\alpha)\cdot\widehat{\sigma}_\theta(x),\ \widehat{\theta}(x) + cv^b_n(1-\alpha)\cdot\widehat{\sigma}_\theta(x)\right], \tag{38}$$

for any compact subset $\mathcal{C} \subseteq \mathcal{X}$. The computational procedure follows the details in Algorithm 1 of Section 4.

Similar to Theorem 4.1, we prove the uniform size control of the UCBs given by (38) in the following theorem.

THEOREM 5.1 (Uniform Inference of Sieve Score Bootstrap). *Suppose that Assumptions 1, 2, and 4–8 and other conditions of Theorem 3.1 hold. Then,*

$$\Pr\left\{\theta(x) \in \left[\widehat{\theta}^b_l(x), \widehat{\theta}^b_u(x)\right], \textit{for all } x \in \mathcal{C}\right\} = 1 - \alpha + o(1).$$

It is worth mentioning that the time fixed effect f_t allows for various types of persistence, including not only the stationary mixingale, mildly stationary, unit root, and local-to-unity processes but also the deterministic trends, while the regressor $\{x_{it}\}$ follows the stationary mixingale condition of Assumption 2. The robustness of our inference method enables the broad applicability of our framework in a wide class of panel models.

Though the time effect, f_t, introduces cross-sectional dependence, Theorem 5.1 utilizes the independence of the population moments that involve f_t in the score vector to motivate Yurinskii's coupling principle and uniform inference. More detailed discussions and proofs given in the Supplementary Material will be omitted in the main article for simplicity. Extensive simulation studies that examine the finite-sample performance of the UCBs in the two-way fixed-effects panel model (33) are also collected in the Supplementary Material.

6. MONTE CARLO SIMULATIONS

This section evaluates the finite sample performance of our bootstrapping UCBs and pointwise CIs that use the Gaussian critical values. Namely, we consider four types of models, including the one-way and two-way fixed-effects panel models with a univariate regressor (Models 1 and 3; Equations (1) and (33)), the one-way fixed-effects partial linear panel model with multivariate regressors (Model 2-1; Equation (31)), and the one-way fixed-effects panel data model with an additive functional form (Model 2-2; Equation (C.5) of the Supplementary Material).

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

6.1. Simulation Setup

The simulations reported below involve four models (Model 1, Model 2-1, Model 2-2, and Model 3) which are designed as follows.

For the univariate model with one-way fixed effects (Model 1), to ensure the intertemporal correlations for regressors and also the correlation between fixed effect and regressors, we set the regressor $x_{it} = \phi_x x_{i,t-1} + v_{it}^*$ and the fixed effects $\gamma_i = (1/T)\sum_{t=1}^{T} \Phi\left(w_{it}^*\right)$, in which the function Φ is the Gaussian CDF and $\left(w_{it}^*, v_{it}^*\right)$ are iid draws from a joint normal distribution with zero means, unity variances, and covariance terms $\mathbb{E}[w_{it}^* v_{it}^*] = 0.5$. The term $\{u_{\epsilon,it}\}$ is randomly drawn from $\mathcal{N}(0,1)$ that is independent of $(w_{it}^*, v_{it}^*)'$ and the $\{\epsilon_{it}\}$ follows $\epsilon_{it} = \phi_\epsilon \epsilon_{i,t-1} + u_{\epsilon,it}$. The autoregressive slope $\phi_x \in \{0.5, 1 - 1/T^{0.5}, 0.92, 1, 1.03\}$ and $\phi_\epsilon = 0.2$. It is evident that our DGPs cover the cases of stationary ($\phi_x = 0.5$), mildly stationary ($\phi_x = 1 - 1/T^{0.5}$), and unity/local-to-unity ($\phi_x = 1$), staying consistent with the theoretical derivations. In addition, parts of the simulated autoregressive roots are taken from empirical estimates in the next section; for instance, the estimated roots, 0.9217 and 1.0317, are captured from the dynamics of asset holding and income.

For the multivariate panel data model with one-way fixed effects (Models 2-1 and 2-2), we use the bivariate regressors $x_{1,it} = \phi_{x1} x_{1,i,t-1} + v_{1,it}^*$, $x_{2,it} = \phi_{x2} x_{2,i,t-1} + v_{2,it}^*$ and the individual fixed effects $\gamma_i = (1/T)\sum_{t=1}^{T} \Phi\left(w_{it}^*\right)$, in which the function $\Phi(\cdot)$ is the Gaussian CDF and $(w_{it}^*, v_{1,it}^*, v_{2,it}^*)$ are iid draws from the joint Gaussian distribution with zero means, unity variances, and covariance terms $\mathbb{E}[w_{it}^* v_{1,it}^*] = \mathbb{E}[w_{it}^* v_{2,it}^*] = 0.5$ and $\mathbb{E}[v_{1,it}^* v_{2,it}^*] = 0.25$. The error term ϵ_{it} is identical to the univariate panel in Model 1. The autoregressive coefficients are set as $\phi_{x1}, \phi_{x2} \in \{0.5, 0.92, 1, 1.03\}$, in which the autoregressive roots $\phi_{x1}, \phi_{x2} \in \{0.92, 1.03\}$ are estimated from empirical analysis.

In the univariate panel model with two-way fixed effects (Model 3), we set the regressor $x_{it} = \phi_x x_{i,t-1} + v_{it}^*$, the individual fixed effects $\gamma_i = (1/T)\sum_{t=1}^{T} \Phi\left(w_{1,it}^*\right)$, and the time fixed effects $f_t = \phi_f f_{t-1} + w_{2t}^*$ with $w_{2t}^* = (1/n)\sum_{i=1}^{n} w_{2,it}^*$, in which the function $\Phi(\cdot)$ is the Gaussian CDF and $\left(w_{1,it}^*, w_{2,it}^*, v_{it}^*\right)$ are iid draws from the joint Gaussian distribution with zero means, unity variances, and the covariances $\mathbb{E}[w_{1,it}^* v_{it}^*] = \mathbb{E}[w_{2,it}^* v_{it}^*] = 0.5$ and $\mathbb{E}[w_{1,it}^* w_{2,it}^*] = 0$. The error ϵ_{it} is set identically as the univariate panel model with one-way fixed effects. The autoregressive parameters $\phi_x \in \{0.5, 0.92\}$ and $\phi_f = 0.5$.

We conduct simulation studies with functional forms given by

$$\text{F1: } g(x) = \sin(x)\exp(-x^2/2), \tag{39}$$

$$\text{F2: } g(x_1, x_2) = \sin(x_1)\exp(-x_1^2/2) + x_2, \tag{40}$$

$$\text{F3: } g(x_1, x_2) = \sin(x_1)\exp(-x_1^2/2) + x_2\exp(-x_2^2/2). \tag{41}$$

To be specific, the functional form F1 is applied to the univariate models, Models 1&3, and the forms F2 and F3 correspond to the multivariate models with partial linear form (Model 2-1) and additive functional form (Model 2-2).

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

We generate 10,000 simulation paths with the schemes of sample size $\{n = 100, T = 500\}$ and $\{n = 792, T = 6\}$ for each model.[6] For each simulated path, we implement our method using the Hermite basis functions to approximate the unknown functional, the first-order partial derivative regarding the entry of interest ($\partial g(\cdot)$ for F1 and $\partial_{x_1} g(\cdot,\cdot)$ for F2 and F3). Also, we use different values for the order of Hermite sieve expansions, k ($k \in \{6, 7, 8\}$), to check the robustness of simulation performance. We choose the bandwidth for the HAC estimation by using $\lfloor T^{1/3} \rfloor$ when $T \geq 10$ and use the conventional simple variance estimator when $T < 10$.

In this simulation study, we care about the shape of $g(\cdot)$, particularly the first-order partial derivative, $\partial g(\cdot)$ (or $\partial_{x_1} g(\cdot)$). For each experiment, we calculate the 90%, 95%, and 99% UCBs for $\partial g(\cdot)$ (or $\partial_{x_1} g(\cdot)$) over the chosen set $[\underline{x}, \bar{x}]$[7] with 1,000 bootstrap replications. To implement the uniform inference methods, we randomly draw the bootstrap innovations ω_i from the standard normal distribution to construct the critical values of sup-test. We then calculate the probability coverage of our bootstrapping UCBs of interest[8] and compare it with that of the pointwise CIs based on the Gaussian critical values.

The details of the bootstrapping UCBs for the one-way fixed-effects univariate panel model are presented in Table 1, giving the empirical probability coverage under the uniform null hypothesis. By comparing the estimated probability coverage levels with their nominal values (90%, 95%, and 99%), we can find that the UCBs attain the nominal level, illustrating that the sup-type statistics can control over-rejections nicely across all $x \in \mathcal{C}$ and with stationary/mildly stationary/local-to-unity regressors. Comparatively, the pointwise CIs attain severe size distortions at all significance levels, again demonstrating the risk of spurious statistical significance.

Further sensitivity analysis for the choice of sieve expansion order, k, is given in Table 1, in which we examine the order of Hermite basis expansions, namely, $k = 6$, 7, and 8. Our bootstrapping UCBs show satisfying size controls across these cases, so it is reasonable to apply these choices of k to empirical analysis. Among these choices, it is well evident that the bootstrapping UCBs with $k = 7$ demonstrate the empirical rejection rates closest to the nominal levels. By this observation, we use the tuning parameter $k = 7$ in our empirical analysis of the next section.

Simulation results for the one-way fixed-effects multivariate model with the partial linear form are in Table 2, providing the empirical sizes of our bootstrapping UCBs and pointwise CIs. Overall, all these simulations show that the sieve score bootstrapping method provides uniform probability coverage for the partial linear panel model in finite samples, corroborating that the bootstrapping UCBs can help reduce spurious significance regarding the uniform null hypothesis robustly. This robust performance holds across nonstationary regressors and various choices of

[6] More simulation results of $n \in \{100, 200\}$ and $T \in \{5, 15, 25, 50, 100\}$ are given in the Supplementary Material. The sample size $\{n = 792, T = 6\}$ used in the main article mimics our empirical data.

[7] We use the 20% and 80% quantiles of the regressor x_{it}, as $\underline{x}$ and $\bar{x}$.

[8] The discussion of UCBs for additive functionals is given in Section C.ii of the Supplementary Material.

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

TABLE 1. Probability coverage (PC) of uniform and pointwise confidence bands for one-way fixed-effects univariate panel models (Model 1) with the function form (F1)

	Uniform confidence bands									Pointwise CIs		
	$k=6$			$k=7$			$k=8$			$k=7$		
ϕ_x	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
$n=100, T=50$												
0.5	0.894	0.947	0.988	0.893	0.947	0.992	0.892	0.945	0.988	0.558	0.734	0.921
0.85	0.892	0.949	0.988	0.894	0.946	0.989	0.887	0.942	0.988	0.470	0.662	0.898
1	0.894	0.947	0.989	0.897	0.951	0.989	0.894	0.950	0.990	0.322	0.519	0.840
0.92	0.897	0.952	0.989	0.902	0.952	0.990	0.899	0.950	0.991	0.434	0.622	0.890
1.03	0.886	0.941	0.988	0.884	0.943	0.987	0.886	0.939	0.986	0.314	0.509	0.816
$n=792, T=6$												
0.92	0.906	0.956	0.989	0.902	0.950	0.987	0.904	0.951	0.988	0.440	0.622	0.866
1.03	0.900	0.948	0.992	0.907	0.954	0.990	0.905	0.954	0.989	0.398	0.589	0.851

Note: [a]The ϕ_x values 0.5, $0.85(=1-1/50^{0.5})$ and 1 denote stationary, mildly stationary, and unity roots. [b]The ϕ_x values 0.92 and 1.03 are estimated from the empirical data of asset holding and income. [c]The simulated size $n=792,\ T=6$ mimics the sample size of empirical data. [d]The kth-order Hermite basis expansion applies to the nonparametric component ($k=6,7,8$). [e]The uniform and pointwise confidence bands are computed for $\partial_x g(\cdot)$.

tuning parameter k. It is well evident that the empirical size performance in the partial linear panel model is best controlled around the nominal level when $k=7$, which is suggested for our empirical study. Similar patterns can also be found in the bivariate panel data models with additive functional forms, as presented in Table 3.

Table 4 gives the empirical size behaviors of our UCBs in the two-way fixed-effects panel data model. With stationary regressors, the probability coverage of UCBs is very close to the nominal level, showing the flexibility of our method in allowing for the two-way fixed-effects panel model. Furthermore, when an autoregressive root close to unity (e.g., 0.92) is included in the simulation study, the degree of robustness is well shown; for instance, when across all levels of significance, the empirical size performances of UCBs are well controlled. In comparison, the pointwise CIs that use the Gaussian critical values encounter severe size distortions in specifying the uniform null hypothesis. Similarly, by evaluating the behaviors of different k, we find that the choice of $k=7$ helps attain the best-controlled size performance.

7. EMPIRICAL ILLUSTRATION

In this section, we use the data of the Panel Survey of Income Dynamics (PSID) from 1999 to 2009 (six waves) to examine the influence of financial status on

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

TABLE 2. Probability coverage (PC) of uniform and pointwise confidence bands for one-way fixed-effects multivariate panel models (Model 2-1) with a partial linear form (F2)

		Uniform confidence bands									Pointwise CIs		
		$k=6$			$k=7$			$k=8$			$k=7$		
ϕ_{x1}	ϕ_{x2}	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
$n=100, T=50$													
0.50	1.00	0.893	0.946	0.987	0.892	0.941	0.987	0.892	0.943	0.989	0.559	0.728	0.918
1.00	0.5	0.893	0.944	0.987	0.889	0.944	0.989	0.891	0.944	0.988	0.311	0.514	0.829
0.92	1.03	0.886	0.943	0.990	0.890	0.944	0.987	0.892	0.944	0.988	0.422	0.617	0.876
1.03	0.92	0.885	0.942	0.987	0.887	0.940	0.989	0.886	0.943	0.989	0.314	0.513	0.824
$n=792, T=6$													
0.92	1.03	0.895	0.944	0.987	0.891	0.942	0.987	0.890	0.945	0.988	0.423	0.601	0.851
1.03	0.92	0.891	0.946	0.990	0.892	0.944	0.989	0.896	0.945	0.989	0.378	0.567	0.836

Note: [a]The kth-order Hermite basis expansion applies to the nonparametric component of x_1 ($k = 6, 7, 8$). [b]The uniform and pointwise confidence bands are computed for $\partial_{x_1} g(\cdot, \cdot)$.

TABLE 3. Probability coverage (PC) of uniform and pointwise confidence bands for one-way fixed-effects multivariate panel models (Model 2-2) with an additive function form (F3)

		Uniform confidence bands									Pointwise CIs		
		$k=6$			$k=7$			$k=8$			$k=7$		
ϕ_{x1}	ϕ_{x2}	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
$N=100, T=50$													
0.50	1.00	0.894	0.945	0.988	0.895	0.945	0.988	0.894	0.947	0.989	0.551	0.725	0.922
1.00	0.5	0.886	0.941	0.989	0.885	0.940	0.989	0.884	0.942	0.988	0.319	0.517	0.827
0.92	1.03	0.892	0.944	0.989	0.888	0.943	0.990	0.892	0.946	0.990	0.413	0.612	0.878
1.03	0.92	0.885	0.943	0.989	0.884	0.942	0.986	0.881	0.939	0.987	0.311	0.507	0.819
$N=792, T=6$													
0.92	1.03	0.895	0.949	0.989	0.892	0.948	0.990	0.892	0.946	0.988	0.422	0.599	0.853
1.03	0.92	0.894	0.945	0.989	0.894	0.946	0.990	0.900	0.950	0.989	0.385	0.565	0.839

Note: [a]The kth-order Hermite basis expansion applies to the nonparametric component of both x_1 and x_2 ($k = 6, 7, 8$). [b]The uniform and pointwise confidence bands are computed for $\partial_{x_1} g(\cdot, \cdot)$.

consumption. We estimate consumption responses to income and asset holdings and establish the UCBs to investigate the nonlinear relationships.

Similar to Arellano, Blundell, and Bonhomme (2017), we proceed to estimate a model that allows for nonstationary measures of assets in a sample of homogeneous households from the PSID data. For instance, the households hold a

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

TABLE 4. Probability coverage (PC) of uniform and pointwise confidence bands for two-way fixed-effects panel models with the function form (F1)

	Uniform confidence bands									Pointwise CIs		
	$k=6$			$k=7$			$k=8$			$k=7$		
ϕ_x	90%	95%	99%	90%	95%	99%	90%	95%	99%	90%	95%	99%
$n=100, T=50$												
0.5	0.887	0.946	0.989	0.892	0.944	0.987	0.890	0.943	0.988	0.551	0.723	0.919
0.92	0.889	0.944	0.989	0.888	0.941	0.989	0.887	0.944	0.989	0.410	0.607	0.872
$n=792, T=6$												
0.5	0.896	0.949	0.990	0.891	0.947	0.990	0.891	0.946	0.989	0.482	0.654	0.873
0.92	0.899	0.951	0.989	0.900	0.950	0.989	0.897	0.948	0.989	0.437	0.616	0.862

Note: [a]The kth-order Hermite basis expansion applies to the nonparametric component of x ($k=6,7,8$).
[b]The uniform and pointwise confidence bands are computed for $\partial g(\cdot)$.

one-period bond whose return $\{a_{it}\}$ follows a nonstationary relationship with an autoregressive root $(1+r)>1$ as

$$a_{i,t+1} = (1+r)\,a_{it} + i_{it} - c_{it}, \tag{42}$$

in which a_{it}, i_{it}, and c_{it} denote the asset, disposable income, and consumption, respectively. Since a_{it} follows a nonstationary process whose root is close to unity, it is worth mentioning that our uniform inference procedures can naturally allow for nonstationary regressors of the unit root type.

The baseline model in our discussion is formulated as

$$c_{it} = \gamma_i + g\,(a_{it}, i_{it}) + \epsilon_{it}, \tag{43}$$

in which c_{it} is the consumption of households and the covariates a_{it} and i_{it} denote the asset and income. By applying the within-group sieve estimator that relies on the Hermite basis functions, we intend to find the presence of nonlinear impacts of assets and income on the consumption of households.

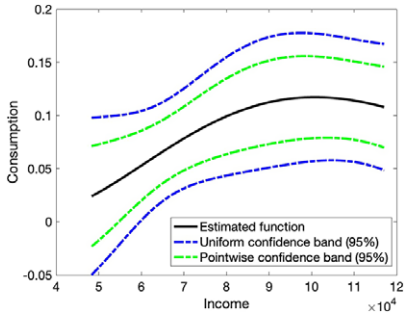
To give the UCBs for functionals $\partial_a g(\cdot,\cdot)$ and $\partial_i g(\cdot,\cdot)$, we discuss two types of specifications for its functional form: the partial linear form and the additive functional form:

(i) $g(a_{it}, i_{it}) = m(i_{it}) + \eta_1 a_{it}$; (ii) $g(a_{it}, i_{it}) = m(a_{it}) + \eta_2 i_{it}$;
and (iii) $g(a_{it}, i_{it}) = m_1(a_{it}) + m_2(i_{it})$.

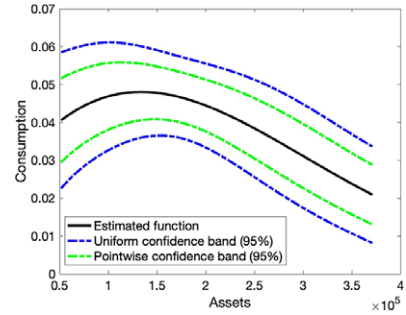
We utilize these specifications and discuss their UCBs. Computations and asymptotics of the sieve score bootstrapping UCBs for the partial linear model and the additive functional form are given in Remark 7 of the main article and Section C.ii of the Supplementary Material.

The sieve expansion order $k=7$ for both these models, as suggested by simulation studies. Based on (24), we apply the variance estimator that is designed for the case of a small T, since $T=6$. The sup-type statistics to test for nonlinear

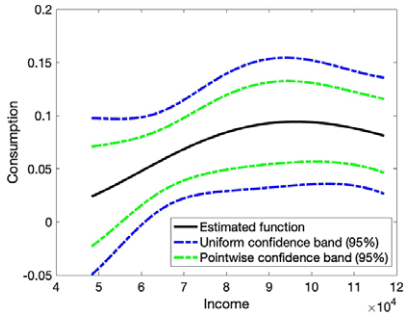
https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press



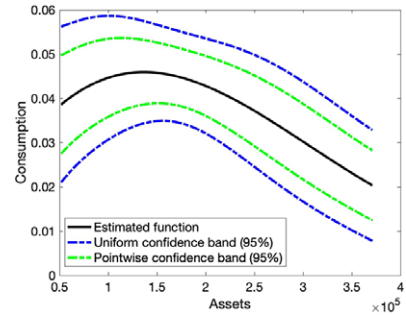
(a) Partial linear (i): Income



(b) Partial linear (ii): Assets holdings



(c) Additive (iii): Income



(d) Additive (iii): Assets holdings

FIGURE 1. UCBs for consumption responses.
[a]Figure 1(a) uses Model (i) and tests $\partial_l g(\cdot,\cdot)$. Figure 1(b) uses Model (ii) and tests $\partial_a g(\cdot,\cdot)$.
[b] Figures 1(c) and 1(d) use Model (iii) and test $\partial_l g(\cdot,\cdot)$ and $\partial_a g(\cdot,\cdot)$.

patterns are computed on the 200 evenly chosen grid points between 20% and 80% quantiles of the variable of interest. Figures 1(a) and 1(c) show the nonlinear dynamics of consumption concerning income shocks and show the results for both the partial linear and additive functional forms. It is worth mentioning that the 95% pointwise CIs show that households tend to increase consumption when positive changes are expected for nearly all levels. In comparison, it is evident that the 95% UCBs have larger widths and reveal that only positive changes that occur for the income of a higher level will increase consumption. Also, it is shown that the approaches for the additive and the partial linear functional forms produce similar empirical findings.

Figures 1(b) and 1(d) discuss the dynamics of consumption concerning asset holdings and show the empirical results that rely on the partial linear and additive functional forms. Our bootstrapping UCBs show that the nonlinear dynamics in the consumption concerning asset holdings are statistically significant on the 5% level. Also, it is revealed that the UCBs have wider widths than the pointwise confidence intervals that use the Gaussian critical values. The UCBs that depend

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

on both types of panel models show similar patterns, demonstrating the robustness of our findings.

8. CONCLUSION

We develop uniform inference procedures for the nonparametric/semi-nonparametric panel models with fixed effects. In the framework of the present article, we allow for general forms of intertemporal correlations in regressors, including stationary mixingale, mildly stationary, or local-to-unity process. We extend this uniform inference method to the nonparametric panel model with two-way fixed effects. While the pointwise inference problem of the nonparametric panel models has been extensively explored in the literature, the sup-type test statistic has not been discussed in the literature. The present article contributes significantly to this area by filling the void.

To the best of our knowledge, our work makes the first attempt to consider the uniform inference procedure (sup-test) in the nonparametric panel model with individual fixed effects. Further extension of our framework can consider cross-sectional correlations represented as interactive fixed-effect models or social interaction models. Such discussions are beyond the scope of the present article and will be left for our future research.

COMPETING INTEREST STATEMENT

The authors declare no competing interests exist.

FUNDING STATEMENT

Yanbo Liu acknowledges that this project is supported by the National Natural Science Foundation of China (NSFC) under Grant No. 72303130. Nan Liu acknowledges that this project is supported by the National Natural Science Foundation of China (NSFC) under Grant No. 72103172.

SUPPLEMENTARY MATERIAL

Liu, N., & Liu, Y. (2025): Supplement to "Uniform Inference for Nonparametric Panel Models with Fixed Effects," Econometric Theory Supplementary Material. To view, please visit https://doi.org/10.1017/S026646662510025X.

REFERENCES

Arellano, M., Blundell, R., & Bonhomme, S. (2017). Earnings and consumption dynamics: A nonlinear panel data framework. *Econometrica*, 85(3), 693–734.

Belloni, A., Chernozhukov, V., Chetverikov, D., & Kato, K. (2015). Some new asymptotic theory for least squares series: Pointwise and uniform results. *Journal of Econometrics*, 186(2), 345–366.

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. *Handbook of Econometrics*, 6, 5549–5632.
Chen, X., & Christensen, T. M. (2018). Optimal sup-norm rates and uniform inference on nonlinear functionals of nonparametric IV regression. *Quantitative Economics*, 9(1), 39–84.
Chernozhukov, V., Chetverikov, D., & Kato, K. (2014). Gaussian approximation of suprema of empirical processes. *Annals of Statistics*, 42(4), 1564–1597.
Dong, C., & Gao, J. (2019). Expansion and estimation of Lévy process functionals in nonlinear and nonstationary time series regression. *Econometric Reviews*, 38(2), 125–150.
Dong, C., Gao, J., & Peng, B. (2019). Estimation in a semiparametric panel data model with nonstationarity. *Econometric Reviews*, 38(8), 961–977.
Dong, C., Gao, J., & Tjøstheim, D. (2016). Estimation for single-index and partially linear single-index integrated models. *Annals of Statistics*, 44(1), 425–453.
Dong, C., Gao, J., & Zhu, P. (2021). Recent developments on nonstationary time series model estimation and testing driven by orthogonal series method. *China Journal of Econometrics*, 1(3), 479.
Duffy, J. A. (2020). Asymptotic theory for kernel estimators under moderate deviations from a unit root, with an application to the asymptotic size of nonparametric tests. *Econometric Theory*, 36(4), 559–582.
Hahn, J., Kuersteiner, G., & Mazzocco, M. (2022). Joint time-series and cross-section limit theory under mixingale assumptions. *Econometric Theory*, 38(5), 942–958.
Kurisu, D., Kato, K., & Shao, X. (2024). Gaussian approximation and spatially dependent wild bootstrap for high-dimensional spatial data. *Journal of the American Statistical Association*, 119(547), 1820–1832.
Lee, Y. (2014). Nonparametric estimation of dynamic panel models with fixed effects. *Econometric Theory*, 30(6), 1315–1347.
Li, J., & Liao, Z. (2020). Uniform nonparametric inference for time series. *Journal of Econometrics*, 219(1), 38–51.
Li, J., Liao, Z., & Zhou, W. (2024). Uniform nonparametric inference for spatially dependent panel data. *Journal of Business & Economic Statistics*, 42(2), 654–664.
Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61(4), 631–653.
Phillips, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika*, 74(3), 535–547.
Phillips, P. C. B., & Magdalinos, T. (2007). Limit theory for moderate deviations from a unit root. *Journal of Econometrics*, 136(1), 115–130.
Phillips, P. C. B., & Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Econometrica*, 67(5), 1057–1111.
Stone, C. J. (1980). Optimal rates of convergence for nonparametric estimators. *Annals of Statistics*, 8(6), 1348.
Su, L., & Lu, X. (2013). Nonparametric dynamic panel data models: Kernel estimation and specification testing. *Journal of Econometrics*, 176(2), 112–133.

A. PROOFS

A.1. Proof of Theorem 3.1

In order to apply the coupling method, we enhance the probability space in which the standard Gaussian variate $\mathcal{N}_k$ is measurable and apply Yurinskii's coupling principle (Lemma A.2 of the Supplementary Material) to the empirical process:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i, \quad \xi_i := \Sigma^{-1/2}\mathbb{G}_T\left[\widetilde{p}_{it}\left(\epsilon_{it}+r_{it}\right)\right].$$

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

By definition, it follows that $\mathbb{E}[\xi_i] = \mathbf{0}_{k\times 1}$ for each individual i. Another sufficient condition to verify the validity of Yurinskii's coupling is given as

$$\begin{aligned}\mathbb{E}\left\|\xi_i\right\|^3 &\lesssim \mathbb{E}\left\|\mathbb{G}_T\left[\widetilde{p}_{it}\left(\epsilon_{it}+r_{it}\right)\right]\right\|^3 \\ &\lesssim \mathbb{E}\left\|\mathbb{G}_T\left[\widetilde{p}_{it}\right]\right\|^2 a_T\tau_k\left(1+\ell_k^3 c_k^3\right) \\ &\lesssim k\cdot a_T^3\tau_k\left(1+\ell_k^3 c_k^3\right),\end{aligned}$$

in which we use the facts $\sup_{x\in\mathcal{X}}\mathbb{E}\left[\left|\epsilon_{it}\right|^3 | x_{it}=x\right] \lesssim 1$ and $\max_{i\in[n]}\max_{t\in[T]}\left|r_{it}\right| \leq \sup_{x\in\mathcal{X}}\left|r_g(x)\right| \lesssim \ell_k c_k$. The strong approximation in Lemma A.2 of the Supplementary Material yields that

$$\begin{aligned}\Pr\left\{\left\|\frac{\sum_{i=1}^n \xi_i}{\sqrt{n}} - \mathcal{N}_k\right\| \geq 3\delta c_{nT}^{-1}\right\} &\lesssim \frac{k^2 a_T^3\tau_k\left(1+\ell_k^3 c_k^3\right)}{\left(\delta c_{nT}^{-1}\right)^3\sqrt{n}} \\ &\left(1+\frac{\log\left(k^2\cdot a_T\tau_k\left(1+\ell_k^3 c_k^3\right)\right)}{k}\right) \to 0, \qquad \textbf{(A.1)}\end{aligned}$$

due to the rate restriction that $c_{nT}^6 k^4 a_T^6 \tau_k^2 (\log n)^2/n \to 0$.

Based on the linear approximation of (12), we obtain the following:

$$\begin{aligned}&\left\|\sqrt{n}a_T\cdot\alpha(x)'\left(\widehat{\beta}-\beta\right) - \alpha(x)'\Omega^{1/2}\mathcal{N}_k\right\| \\ &\quad\leq \left\|\mathbb{G}_n[\alpha(x)'\Omega^{1/2}\xi_i] - \alpha(x)'\Omega^{1/2}\mathcal{N}_k\right\| + \overline{R}_{1,nT} + \overline{R}_{2,nT} = o_p\left(c_{nT}^{-1}\right), \qquad \textbf{(A.2)}\end{aligned}$$

over $x\in\mathcal{C}$, by the stochastic boundedness of $\|\alpha(x)'\Omega^{1/2}\|$, Proposition A.1 of the Supplementary Material, (A.1) and the conditions imposed for $\overline{R}_{1,nT}$ and $\overline{R}_{2,nT}$ in Theorem 3.1. Therefore, under the condition $\sup_{x\in\mathcal{X}} n^{1/2}a_T|r_g(x)|/\|s(x)\| = o_p\left(c_{nT}^{-1}\right)$, equation (19) holds. ■

A.2. Proof of Theorem 4.1

By Proposition A.1 of the Supplementary Material, it follows that

$$\left|\frac{\widehat{\sigma}_\theta(x)}{\sigma_\theta(x)} - 1\right| \lesssim_p M\cdot(v_{nT}\vee 1 + \ell_k c_k)\left(\sqrt{\frac{\xi_k^2\log k}{n}} + \frac{1}{k}\right) = o(c_{nT}^{-1}), \qquad \textbf{(A.3)}$$

uniformly over $x\in\mathcal{C}$, ensured by the imposed conditions. Similar to Theorem 3.1,

$$\frac{\widehat{\theta}(x)-\theta(x)}{\sigma_\theta(x)} = \sqrt{n}a_T\cdot\frac{\widehat{\theta}(x)-\theta(x)}{\left\|l_\theta(x)'\Omega^{1/2}\right\|} = \frac{l_\theta(x)'\Omega^{1/2}}{\left\|l_\theta(x)'\Omega^{1/2}\right\|}\mathcal{N}_k + o_p\left(c_{nT}^{-1}\right), \text{ in } L_\infty(\mathcal{C}). \qquad \textbf{(A.4)}$$

Combining (A.3) and (A.4), we can further show that

$$t_\theta(x) = \frac{\widehat{\theta}(x)-\theta(x)}{\widehat{\sigma}_\theta(x)} = \frac{l_\theta(x)'\Omega^{1/2}}{\left\|l_\theta(x)'\Omega^{1/2}\right\|}\mathcal{N}_k + o_p\left(c_{nT}^{-1}\right), \text{ in } L_\infty(\mathcal{C}). \qquad \textbf{(A.5)}$$

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press

Next, we proceed to show the validity of our sieve score bootstrap method for attaining uniform size controls. We employ Yurinskii's coupling principle (Lemma A.2 of the Supplementary Material). To apply the strong approximation, we enhance the probability space in which the standard Gaussian variate $\mathcal{N}_k$ is measurable and apply the coupling principle to

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i^*, \quad \xi_i^* := \Sigma^{-1/2}\omega_i \cdot \mathbb{G}_T[\widetilde{p}_{it}\widetilde{\epsilon}_{it}]. \tag{A.6}$$

In (A.6), each component as $(1/\sqrt{n})\omega_i\Sigma^{-1/2}\cdot\mathbb{G}_T[\widetilde{p}_{it}\widetilde{\epsilon}_{it}]$ is centered at zero because

$$\mathbb{E}\left[\omega_i|\{y_{it},x_{it},\epsilon_{it}\}_{i\in[n],t\in[T]}\right]=0, \text{ for any } i\in[n],$$

by the independence between ω_i and other random variates. Then, it follows that $\mathbb{E}[\xi_i^*]=\mathbf{0}_{k\times 1}$ for each $i\in[n]$. Next, we verify the sufficient condition for the Gaussian coupling as

$$\mathbb{E}\left\|\xi_i^*\right\|^3 \lesssim \mathbb{E}\left\|\omega_i\mathbb{G}_T[\widetilde{p}_{it}\widetilde{\epsilon}_{it}]\right\|^3 \lesssim \mathbb{E}\left\|\mathbb{G}_T[\widetilde{p}_{it}]\right\|^2 a_T\tau_k \lesssim k\cdot a_T^3\tau_k,$$

in which the conditions $\sup_{x\in\mathcal{X}}\mathbb{E}[|\epsilon_{it}|^3|x_{it}=x]\lesssim 1$, $\sup_{x\in\mathcal{X}}|r_\theta(x)|\lesssim 1$, and

$$\sup_{i\in[n]}\sup_{t\in[T]}\mathbb{E}\left[\omega_i^3|\{y_{is},x_{is},\epsilon_{is}\}_{i\in[n],s\in[T]}\right]\lesssim 1.$$

Therefore, the coupling of Lemma A.2 of the Supplementary Material yields that for any $\delta>0$,

$$\Pr\left\{\left\|\frac{\sum_{i=1}^n\xi_i^*}{\sqrt{n}}-\mathcal{N}_k\right\|\geq 3\delta c_{nT}^{-1}\right\}\lesssim\frac{nk^2a_T\tau_k}{\left(\delta c_{nT}^{-1}\sqrt{n}\right)^3}\left(1+\frac{|\log\left(k^2\cdot a_T\tau_k\right)|}{k}\right)\to 0,$$

due to $c_{nT}^6k^4a_T^6\tau_k^2\log^2 n/n\to 0$. Then, we have

$$\sup_{x\in\mathcal{C}}\left|t_\theta^b(x)-\frac{l_\theta(x)'\cdot\Omega^{1/2}\mathcal{N}_k}{\sigma_\theta(x)}\right|=o(c_{nT}^{-1}). \tag{A.7}$$

Further, combining (A.6) and (A.7), we have

$$\sup_{x\in\mathcal{C}}\left|t_\theta(x)-t_\theta^b(x)\right|=o(c_{nT}^{-1}), \tag{A.8}$$

in which the t test statistic process is given in (26).

Based on (A.8) and Theorem 5.5 of Belloni et al. (2015), it follows that

$$\sup_{s\in\mathbb{R}}\left|\Pr\left(\sup_{x\in\mathcal{C}}\left|t_\theta^b(x)\right|\leq s\right)-\Pr\left(\sup_{x\in\mathcal{C}}|t_\theta(x)|\leq s\right)\right|=o(1), \tag{A.9}$$

for any $s\in\mathbb{R}$ when $n\to\infty$. In the end, we prove the final result.

https://doi.org/10.1017/S026646662510025X Published online by Cambridge University Press