The Econometrics Journal

Econometrics Journal (2012), volume **15**, pp. 56–100.
doi: 10.1111/j.1368-423X.2011.00361.x

Testing for common trends in semi-parametric panel data models with fixed effects

YONGHUI ZHANG[†], LIANGJUN SU[†] AND PETER C. B. PHILLIPS[†,‡,§,¶]

[†]*School of Economics, Singapore Management University, 90 Stamford Road, Singapore 178903.*
E-mails: yhzhang.2008@phdecons.smu.edu.sg, ljsu@smu.edu.sg

[‡]*Cowles Foundation, P.O. Box 208281, Yale University, New Haven, CT 06520-8281, USA.*
[§]*Department of Economics, Owen G. Glenn Building, University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.*
[¶]*Department of Economics, School of Social Sciences, University of Southampton, Southampton SO17 1BJ, UK.*
E-mail: peter.phillips@yale.edu

First version received: May 2011; final version accepted: October 2011

Summary This paper proposes a non-parametric test for common trends in semi-parametric panel data models with fixed effects based on a measure of non-parametric goodness-of-fit (R^2). We first estimate the model under the null hypothesis of common trends by the method of profile least squares, and obtain the augmented residual which consistently estimates the sum of the fixed effect and the disturbance under the null. Then we run a local linear regression of the augmented residuals on a time trend and calculate the non-parametric R^2 for each cross-section unit. The proposed test statistic is obtained by averaging all cross-sectional non-parametric R^2s, which is close to 0 under the null and deviates from 0 under the alternative. We show that after appropriate standardization the test statistic is asymptotically normally distributed under both the null hypothesis and a sequence of Pitman local alternatives. We prove test consistency and propose a bootstrap procedure to obtain P-values. Monte Carlo simulations indicate that the test performs well in finite samples. Empirical applications are conducted exploring the commonality of spatial trends in UK climate change data and idiosyncratic trends in OECD real GDP growth data. Both applications reveal the fragility of the widely adopted common trends assumption.

Keywords: *Common trends*, *Local polynomial estimation*, *Non-parametric goodness-of-fit*, *Panel data*, *Profile least squares*.

1. INTRODUCTION

Modelling trends in time series has a long history. Phillips (2001, 2005, 2010) provides recent overviews covering the development, challenges, and some future directions of trend modelling in time series. White and Granger (2011) offer working definitions of various kinds of trends and invite more discussion on trends in order to facilitate development of increasingly better methods for prediction, estimation and hypothesis testing for non-stationary time-series data. Due to the wide availability of panel data in recent years, research on trend modelling has spread to panel models. Most of the literature falls into two categories depending on whether the trends are stochastic or deterministic. But there is also work on evaporating trends (Phillips, 2007) and econometric convergence testing (Phillips and Sul, 2007, 2009). For reviews on stochastic trends in panel data models, see Banerjee (1999), and Breitung and Pesaran (2008).

Recently, modelling deterministic time trends in non-parametric and semi-parametric settings has attracted interest. Cai (2007) studies a time-varying coefficient time series model with a time trend function and serially correlated errors to characterize the non-linearity, non-stationarity, and trending phenomenon. Robinson (2010) considers non-parametric trending regression in panel data models with cross-sectional dependence. Atak et al. (2011) propose a semi-parametric panel data model to model climate change in the United Kingdom (UK hereafter), where seasonal dummies enter the model linearly with heterogeneous coefficients and the time trend enters non-parametrically. Li et al. (2011) extend the work of Cai (2007) to panel data time-varying coefficient models. Chen et al. (2011, CGL hereafter) extend Robinson's (2010) non-parametric trending panel data models to semi-parametric partially linear panel data models with cross-sectional dependence where all individual unit share a common time trend that enters the model non-parametrically. They propose a semi-parametric profile likelihood approach to estimate the model.

A conventional feature of work on deterministic trending panel models is the imposition of a common trends assumption, implying that each individual unit follows the same time trend behaviour. Such an assumption greatly simplifies the estimation and inference process, and the proposed estimators can be efficient if there is no heterogeneity in individual time trend functions and some other conditions are met. Nevertheless, if the common trends assumption does not stand, the estimates based on non-parametric or semi-parametric panel data models with common trends will be generally inefficient and statistical inference will be misleading. It is therefore prudent to test for the common trends assumption before imposing it.

Since Stock and Watson (1988) there has been a large literature on testing for common trends. But to our knowledge, most empirical works have focused on testing for common stochastic trends. Tests for common deterministic trends are far and few between. Vogelsang and Franses (2005) propose tests for common deterministic trend slopes by assuming linear trend functions and a stationary variance process and examining whether two or more trend-stationary time series have the same slopes. Xu (2011) considers tests for multivariate deterministic trend coefficients in the case of non-stationary variance process. Sun (2011) develops a novel testing procedure for hypotheses on deterministic trends in a multivariate trend stationary model where the long-run variance is estimated by series method. In all cases, the models are parametric and the asymptotic theory is established by passing the time series dimension T to infinity and keeping the number of cross-sectional units n fixed. Empirical applications include Fomby and Vogelsang (2003), and Bacigál (2005), who apply the Vogelsang–Franses test to temperature data and geodetic data, respectively.

This paper develops a test for common trends in a semi-parametric panel data model of the form

$$Y_{it} = \beta' X_{it} + f_i(t/T) + \alpha_i + \varepsilon_{it},\ i = 1, \ldots, n,\ t = 1, \ldots, T, \tag{1.1}$$

where β is a $d \times 1$ vector of unknown parameters, X_{it} is a $d \times 1$ vector of regressors, f_i is an unknown smooth time trend function for cross-section unit i, the α_is represent fixed effects that can be correlated with X_{it}, and ε_{it}s are idiosyncratic errors. The trend functions $f_i(t/T)$

that appear in (1.1) provide for idiosyncratic trends for each individual i. For simplicity, we will assume that (i) $\{\varepsilon_{it}\}$ satisfies certain martingale difference conditions along the time dimension but may be correlated across individuals, and (ii) $\{\varepsilon_{it}\}$ are independent of $\{X_{it}\}$. Note that f_i and α_i are not identified in (1.1) without further restrictions.

Model (1.1) covers and extends some existing models. First, when $f_i \equiv 0$ for all i, (1.1) becomes the traditional panel data model with fixed effects. Second, if $n = 1$, then model (1.1) reduces to the model discussed in Gao and Hawthorne (2006). Third, when $f_i = f$ for some unknown smooth function f and all i, (1.1) becomes the semi-parametric trending panel data model of CGL (2011).

The main objective of this paper is to construct a non-parametric test for common trends. Under the null hypothesis of common trends: $f_i = f$ for all i in (1.1), we can pool the observations from both cross-section and time dimensions to estimate both the finite dimensional parameter (β) and the infinite dimensional parameter (f) under the single identification restriction $\sum_{i=1}^{n} \alpha_i = 0$ or $f(0) = 0$, whichever is convenient. Let $u_{it} \equiv \alpha_i + \varepsilon_{it}$. Let $\widehat{u}_{it}$ denote the estimate of u_{it} based on the pooled regression. The residuals $\{\widehat{u}_{it}\}$ should not contain any useful trending information in the data. This motivates us to construct a residual-based test for the null hypothesis of common trends. To be concrete, we will propose a test for common trends by averaging the n measures of non-parametric goodness-of-fit (R^2) from the non-parametric time series regressions of $\widehat{u}_{it}$ on the time trend for each cross-sectional unit i. Such non-parametric R^2 should tend to zero under the null hypothesis of common trends and diverge from zero otherwise. We show that after being properly centred and scaled, the average non-parametric R^2 is asymptotically normally distributed under the null hypothesis of common trends and a sequence of Pitman local alternatives. We also establish the consistency of the test and propose a bootstrap method to obtain the bootstrap P-values.[1]

To proceed, it is worth mentioning that (1.1) complements the model of Atak et al. (2011) who allow for heterogeneous slopes but a single non-parametric common trend across cross-sections. As mentioned in the concluding remarks, it is also possible to allow the slope coefficients in (1.1) to vary across individuals and consider a joint test for the homogeneity of the slope coefficients and trend components. But this is beyond the scope of this paper.

The rest of the paper is organized as follows. The hypotheses and the test statistic are given in Section 2. We study the asymptotic distributions of the test under the null and a sequence of local alternatives, establish the consistency of the test, and propose a bootstrap procedure to obtain the bootstrap P-values in Section 3. Section 4 conducts a small simulation experiment to evaluate the finite sample performance of our test and reports empirical applications of the test to UK climate change data and OECD economic growth data. Section 5 concludes. Proofs of the main and subsidiary results are given in the Appendices.

1.1. Notation

Throughout the paper we adopt the following notation. For a matrix A, its transpose is A' and Euclidean norm is $\|A\| \equiv [\mathrm{tr}(AA')]^{1/2}$, where $\equiv$ signifies 'is defined as'. When A is a symmetric

[1] To the best of our knowledge, Su and Ullah (2011) are the first to suggest applying such a measure of non-parametric R^2 to conduct model specification test based on residuals from restricted parametric, non-parametric, or semi-parametric regressions, and apply this idea to test for conditional heteroscedasticity of unknown form. Clearly, the non-parametric R^2 statistic can serve as a useful tool for testing many popular hypotheses in econometrics and statistics by playing a role comparable to the important role that R^2 plays in the parametric setup.

matrix, we use $\lambda_{\max}(A)$ to denote its maximum eigenvalue. For a natural number l, we use i_l and I_l to denote the $l \times 1$ vector of ones and the $l \times l$ identity matrix, respectively. For a function f defined on the real line, we use $f^{(a)}$ to denote its ath derivative whenever it is well defined. The operator $\stackrel{p}{\to}$ denotes convergence in probability, and $\stackrel{d}{\to}$ convergence in distribution. We use $(n, T) \to \infty$ to denote the joint convergence of n and T when n and T pass to the infinity simultaneously.

2. BASIC FRAMEWORK

In this section, we state the null and alternative hypotheses, introduce the estimation of the restricted model under the null, and then propose a test statistic based on the average of non-parametric goodness-of-fit measures.

2.1. Hypotheses

The main objective is to construct a test for common trends in model (1.1). We are interested in the null hypothesis that

$$H_0 : f_i(\tau) = f(\tau) \text{ for } \tau \in [0, 1] \text{ and some smooth function } f, i = 1, \ldots, n, \tag{2.1}$$

i.e. all the n cross-sectional units share the common trends function f. The alternative hypothesis is

$$H_1\text{: the negation of } H_0.$$

As mentioned in the introduction, we will propose a residual-based test for the above null hypothesis. To do so, we need to estimate the model under the null hypothesis and obtain the augmented residual, which estimates $\alpha_i + \varepsilon_{it}$. Then for each i, we run the local linear regression of the augmented residuals on t/T, and calculate the non-parametric R^2. Our test statistic is constructed by averaging these n non-parametric R^2s.

2.2. Estimation under the null

To proceed, we introduce the following notation.

$$Y_i \equiv (Y_{i1}, \ldots, Y_{iT})', \; Y \equiv (Y_1', \ldots, Y_n')', \; X_i \equiv (X_{i1}, \ldots, X_{iT})', \; X \equiv (X_1', \ldots, X_n')',$$

$$\varepsilon_i \equiv (\varepsilon_{i1}, \ldots, \varepsilon_{iT})', \; \varepsilon \equiv (\varepsilon_1', \ldots, \varepsilon_n')', \; \alpha \equiv (\alpha_2, \ldots, \alpha_n)', \; D \equiv (-i_{n-1}, I_{n-1})' \otimes i_T,$$

$$\mathbf{f}_i \equiv (f_i(1/T), \ldots, f_i(T/T))', \; \mathbf{F} \equiv (\mathbf{f}_1, \ldots, \mathbf{f}_n)', \; \mathbf{f} \equiv [f(1/T), \ldots, f(T/T)]'.$$

Note that under H_0, $\mathbf{F} = i_n \otimes \mathbf{f}$, and we can write the model (1.1) as

$$Y_{it} = X_{it}'\beta + f(t/T) + \alpha_i + \varepsilon_{it}, \tag{2.2}$$

or in matrix notation as

$$Y = X\beta + i_n \otimes \mathbf{f} + D\alpha + \varepsilon, \tag{2.3}$$

provided we impose the identification condition $\sum_{i=1}^{n} \alpha_i = 0$.

Following Su and Ullah (2006) and CGL (2011), we estimate the model (2.2) by using the profile least squares method. Let $k(\cdot)$ denote a univariate kernel function and h a bandwidth. Let $k_h(\cdot) \equiv k(\cdot/h)/h$. For any positive integer p, let $z_{h,t}^{[p]}(\tau) \equiv (1, (t/T - \tau)/h, \ldots, [(t/T - \tau)/h]^p)'$,

$$z_h^{[p]}(\tau) \equiv \left(z_{h,1}^{[p]}(\tau), \ldots, z_{h,T}^{[p]}(\tau)\right)', \quad \text{and} \quad Z_h^{[p]}(\tau) \equiv i_n \otimes z_h^{[p]}(\tau).$$

We assume that f is $(p+1)$th order continuously differentiable a.e. Let $D_h^p f(\tau) \equiv (f(\tau), hf^{(1)}(\tau), \ldots, h^p f^{(p)}(\tau)/p!)'$. Then for t/T in the neighbourhood of $\tau \in (0, 1)$, we have by the pth order Taylor expansion that $f(\frac{t}{T}) = D_h^p f(\tau)' z_{h,t}^{[p]}(\tau) + o((\frac{t}{T} - \tau)^p)$. Let $k_{h,t}(\tau) \equiv k_h(t/T - \tau)$, $K_h(\tau) \equiv \operatorname{diag}(k_{h,1}(\tau), \ldots, k_{h,T}(\tau))$, and $\mathbb{K}_h(\tau) \equiv I_n \otimes K_h(\tau)$. Define

$$s(\tau) \equiv \left(z_h^{[p]}(\tau)' K_h(\tau) z_h^{[p]}(\tau)\right)^{-1} z_h^{[p]}(\tau)' K_h(\tau) \text{ and}$$
$$S(\tau) \equiv \left(Z_h^{[p]}(\tau)' \mathbb{K}_h(\tau) Z_h^{[p]}(\tau)\right)^{-1} Z_h^{[p]}(\tau)' \mathbb{K}_h(\tau) = n^{-1} i_n' \otimes s(\tau).$$

The profile least-squares method is composed of the following three steps:

1. Let $\theta \equiv (\alpha', \beta')'$. For given θ and $\tau \in (0, 1)$, we estimate $D_h^p f(\tau)$ by

$$\widehat{D}_{h,\theta}^p f(\tau) \equiv \underset{F \in \mathbb{R}^{p+1}}{\arg\min} \left(Y - X\beta - D\alpha - Z_h^{[p]}(\tau) F\right)' \mathbb{K}_h(\tau) \left(Y - X\beta - D\alpha - Z_h^{[p]}(\tau) F\right).$$

   Noting that $S(\tau) D = 0$ by straightforward calculations, the estimator $\widehat{D}_{h,\theta}^p f(\tau)$ is in fact free of α and its first element is given by

$$\widehat{f}_\beta(\tau) \equiv e_1' S(\tau)(Y - X\beta - D\alpha) = n^{-1} \sum_{i=1}^{n} e_1' s(\tau)(Y_i - X_i \beta), \tag{2.4}$$

   where $e_1 = (1, 0, \ldots, 0)'$ is a $(p+1) \times 1$ vector. Let $\widehat{\mathbf{f}}_\beta \equiv (\widehat{f}_\beta(1/T), \ldots, \widehat{f}_\beta(T/T))'$, $S_T \equiv ([e_1' S(1/T)]', \ldots, [e_1' S(T/T)]')'$, and $S_{nT} \equiv i_n \otimes S_T$. Then we have

$$\widehat{\mathbf{F}}_\beta \equiv i_n \otimes \widehat{\mathbf{f}}_\beta = S_{nT}(Y - X\beta). \tag{2.5}$$

2. We estimate (α, β) by

$$\begin{aligned}(\widehat{\alpha}, \widehat{\beta}) &\equiv \underset{\alpha,\beta}{\arg\min}(Y - X\beta - D\alpha - \widehat{\mathbf{F}}_\beta)'(Y - X\beta - D\alpha - \widehat{\mathbf{F}}_\beta) \\ &= \underset{\alpha,\beta}{\arg\min}(Y^* - X^*\beta - D\alpha)'(Y^* - X^*\beta - D\alpha),\end{aligned}$$

   where $Y^* \equiv (I_{nT} - S_{nT})Y$ and $X^* \equiv (I_{nT} - S_{nT})X$. Let $M_D \equiv I_{nT} - D(D'D)^{-1}D'$. Using the formula for partitioned regression, we obtain

$$\widehat{\beta} = (X^{*\prime} M_D X^*)^{-1} X^{*\prime} M_D Y^*, \text{ and} \tag{2.6a}$$

$$\widehat{\alpha} \equiv (\widehat{\alpha}_2, \ldots, \widehat{\alpha}_n) = (D'D)^{-1} D'(Y^* - X^*\widehat{\beta}). \tag{2.6b}$$

   Then α_1 can be estimated by $\widehat{\alpha}_1 \equiv -\sum_{i=2}^n \widehat{\alpha}_i$.

3. Plugging (2.6a) into (2.4), we obtain the estimator of $f(\tau)$:

$$\widehat{f}(\tau) = e_1' S(\tau)(Y - X\widehat{\beta}). \tag{2.7}$$

Let

$$\widehat{\mathbf{f}} \equiv (\widehat{f}(1/T), \ldots, \widehat{f}(T/T))' \quad \text{and} \quad \widehat{\mathbf{F}} \equiv S_{nT}(Y - X\widehat{\beta}) = i_n \otimes \widehat{\mathbf{f}}. \tag{2.8}$$

After we obtain estimates of β and $f(t/T)$, we can estimate $u_{it} \equiv \alpha_i + \varepsilon_{it}$ by $\widehat{u}_{it} \equiv Y_{it} - \widehat{\beta}' X_{it} - \widehat{f}(t/T)$ under the null. Let $\widehat{u}_i \equiv (\widehat{u}_{i1}, \ldots, \widehat{u}_{iT})'$ and $\widehat{u} \equiv (\widehat{u}_1', \ldots, \widehat{u}_n')'$. Then it is easy to verify that

$$\begin{aligned}
\widehat{u} &= (\varepsilon - S_{nT}\varepsilon) + D\alpha + X^*(\beta - \widehat{\beta}) + \mathbf{F}^*, \\
\widehat{u}_i &= (\varepsilon_i - S_T\varepsilon) + \alpha_i i_T + (X_i - S_T X)(\beta - \widehat{\beta}) + (\mathbf{f}_i - S_T \mathbf{F}), \\
\widehat{u}_{it} &= \alpha_i + [\varepsilon_{it} - e_1' S(t/T)\varepsilon] + [X_{it}' - e_1' S(t/T)X](\beta - \widehat{\beta}) + [f_i(t/T) - e_1' S(t/T)\mathbf{F}],
\end{aligned}$$

where $\mathbf{F}^* \equiv (I_{nT} - S_{nT})\mathbf{F}$.

2.3. A non-parametric R^2-based test for common trends

The idea behind our test is simple. Under H_0, $\widehat{u}_{it}$ is a consistent estimate for $u_{it} = \alpha_i + \varepsilon_{it}$, and there is no time trend in $\{u_{it}\}_{t=1}^T$ for each cross-sectional unit i. Nevertheless, under H_1 $\widehat{u}_{it}$ includes an individual-specific time trend component $f_i(t/T) - f^0(t/T)$, where $f^0(\tau) \equiv p \lim \widehat{f}(\tau)$. This motivates us to consider a residual-based test for common trends.

For each i, we propose to run the non-parametric regression of $\{\widehat{u}_{it}\}_{t=1}^T$ on $\{t/T\}_{t=1}^T$:

$$\widehat{u}_{it} = m_i(t/T) + \eta_{it}, \tag{2.9}$$

where $m_i(\tau) \equiv f_i(\tau) - f^0(\tau)$ and $\eta_{it} = \alpha_i + \varepsilon_{it}^* + (\beta - \widehat{\beta})' X_{it}^* + f^0(t/T) - e_1' S(t/T)\mathbf{F}$ is the new error term in the above regression. Clearly, under H_0 we have $m_i(\tau) = 0$ for $\tau \in [0, 1]$. Given observations $\{\widehat{u}_{it}\}_{t=1}^T$, the local linear regression of $\widehat{u}_{it}$ on t/T is fitted by the weighted least squares (WLS) as follows

$$\min_{(c_{i0}, c_{i1}) \in \mathbb{R}^2} \frac{1}{T} \sum_{t=1}^{T} \left[\widehat{u}_{it} - c_{i0} - c_{i1} \left(\frac{t}{T} - \tau \right) \right]^2 \overline{w}_{b,t}(\tau) \tag{2.10}$$

where $b \equiv b(T)$ is a bandwidth parameter such that $b \to 0$ as $T \to \infty$, $\overline{w}_{b,t}(\tau) \equiv w_b(t/T - \tau) / \int_0^1 w_b(t/T - s) ds$, $w_b(\cdot) \equiv w(\cdot/b)/b$, and $w(\cdot)$ is a probability density function (p.d.f.) that has support $[-1, 1]$. By the proof of Lemma B.1 in the Appendix, $\lambda_{tT} \equiv \int_0^1 w_b(t/T - s) ds = 1$ for $t/T \in [b, 1 - b]$ and is larger than 1/2 otherwise. Therefore, $\overline{w}_{b,t}(\tau)$ plays the role of a boundary kernel to ensure that $\int_0^1 \overline{w}_{b,t}(\tau) d\tau = 1$ for any $t = 1, \ldots, T$.[2]

Let $\widetilde{c}_i \equiv (\widetilde{c}_{i0}, \widetilde{c}_{i1})'$ denote the solution to the above minimization problem. Following Su and Ullah (2011), the normal equations for the above regression imply the following local ANOVA decomposition of the total sum of squares (TSS)

$$TSS_i(\tau) = ESS_i(\tau) + RSS_i(\tau), \tag{2.11}$$

[2] Alternatively, one can use the standard kernel weight $w_b(t/T - \tau)$ in place of $\overline{w}_{b,t}(\tau)$ in (2.10) and decompose $TSS_i(\tau)$ analogously to the decomposition in (2.11). But as $\lambda_{tT} \equiv \int_0^1 w_b(t/T - s) ds$ is not identically 1 for all t, $\int_0^1 TSS_i(\tau) d(\tau)$ in this case does not lead to the simple expression in (2.13).

where

$$TSS_i(\tau) \equiv \sum_{t=1}^{T}(\widehat{u}_{it} - \overline{\widehat{u}}_i)^2\overline{w}_{b,t}(\tau),$$

$$ESS_i(\tau) \equiv \sum_{t=1}^{T}(\widetilde{c}_{i0} + \widetilde{c}_{i1}(t/T - \tau) - \overline{\widehat{u}}_i)^2\overline{w}_{b,t}(\tau),$$

$$RSS_i(\tau) \equiv \sum_{t=1}^{T}(\widehat{u}_{it} - \widetilde{c}_{i0} - \widetilde{c}_{i1}(t/T - \tau))^2\,\overline{w}_{b,t}(\tau),$$

and $\overline{\widehat{u}}_i \equiv T^{-1}\sum_{t=1}^{T}\widehat{u}_{it}$. A global ANOVA decomposition of TSS_i is given by

$$TSS_i = ESS_i + RSS_i, \tag{2.12}$$

where

$$\begin{aligned} TSS_i &\equiv \int_0^1 TSS_i(\tau)d\tau = \sum_{t=1}^{T}(\widehat{u}_{it} - \overline{\widehat{u}}_i)^2,\ ESS_i \equiv \int_0^1 ESS_i(\tau)d\tau,\ \text{and} \\ RSS_i &\equiv \int_0^1 RSS_i(\tau)d\tau. \end{aligned} \tag{2.13}$$

Then one can define the non-parametric goodness-of-fit (R^2) for the above local linear regression as

$$R_i^2 \equiv \frac{ESS_i}{TSS_i}.$$

Under H_0, $\{\widehat{u}_{it}\}$ contains no useful trending information so that the above R_i^2 should be close to 0 for each individual i.

Let $W_b(\tau) \equiv \operatorname{diag}(\overline{w}_{b,1}(\tau), \ldots, \overline{w}_{b,T}(\tau))$, $H(\tau) \equiv W_b(\tau)z_b^{[1]}(\tau)(z_b^{[1]}(\tau)'W_b(\tau)z_b^{[1]}(\tau))^{-1} z_b^{[1]}(\tau)'W_b(\tau)$, and $\bar{H} \equiv \int_0^1 H(\tau)d\tau$. It is easy to show that

$$TSS_i = \widehat{u}_i'M\widehat{u}_i,\ ESS_i = \widehat{u}_i'(\bar{H} - L)\widehat{u}_i, \quad \text{and} \quad RSS_i = \widehat{u}_i'(I_T - \bar{H})\widehat{u}_i,$$

where $M \equiv I_T - L$ and $L \equiv i_T i_T'/T$. Define the average non-parametric R^2 as

$$\overline{R}^2 \equiv \frac{1}{n}\sum_{i=1}^{n} R_i^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{ESS_i}{TSS_i}.$$

Clearly $0 \le \overline{R}^2 \le 1$ by construction. We will show that after being appropriately centred and scaled, $\overline{R}^2$ is asymptotically normally distributed under the null and a sequence of Pitman local alternatives.

Before proceeding further, it is worth mentioning a related test statistic that is commonly used in the literature. Under H_0, the $m_i(\cdot)$ function in (2.9) is also common for all i and thus can be written as $m(\cdot)$. $m(t/T) = 0$ for all $t = 1, \ldots, T$ under H_0 and we can estimate this zero function by pulling all the cross-sectional and time series observations together to obtain the estimate $\widehat{m}(\cdot)$, say. Then we can compare this estimate with the non-parametric trend regression estimate $\hat{m}_i(t/T)$ of $m_i(t/T)$ to obtain the following L_2 type of test statistic

$$D_{nT} \equiv \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}[\hat{m}_i(t/T) - \hat{m}(t/T)]^2.$$

Noting that the estimate $\hat{m}(t/T)$ has a faster convergence rate than $\hat{m}_i(t/T)$ to 0 under the null, it is straightforward to show that under suitable conditions this test statistic is asymptotically equivalent to $\bar{D}_{nT} \equiv \frac{1}{n}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{m}_i(t/T)^2$ under the null. Further noticing that $\sum_{t=1}^{T}\hat{m}_i(t/T)^2/TSS_i$ can be regarded as a version of non-parametric *non-centred* R^2 measure for the cross-sectional unit i, we can simply interpret $\bar{D}_{nT}$ as a weighted non-parametric non-centred R^2 -based test where the weight for cross-sectional unit i is given by TSS_i. In this paper, we focus on the test based on $\overline{R}^2$ because it is scale-free and is asymptotically pivotal under the null after bias-correction. See the remark after Theorem 3.1 for further discussion.

3. ASYMPTOTIC DISTRIBUTIONS

In this section, we first present the assumptions that are used in later analysis and then study the asymptotic distribution of average non-parametric R^2 under both the null hypothesis and a sequence of Pitman local alternatives. We then prove the consistency of the test and propose a bootstrap procedure to obtain bootstrap P-values.

3.1. Assumptions

Let $\mathcal{F}_{n,t}(\xi)$ denote the σ-field generated by $(\xi_1, \ldots, \xi_t)$ for a time series $\{\xi_t\}$. To establish the asymptotic distribution of our test statistic, we make the following assumptions.

ASSUMPTION 3.1. (*a*) *The regressor X_{it} is generated as follows:*

$$X_{it} = g_i\left(\frac{t}{T}\right) + v_{it}. \tag{3.1}$$

(*b*) *Let $v_t \equiv (v_{1t}, \ldots, v_{nt})'$ for $t = 1, \ldots, T$. $\{v_t,\ \mathcal{F}_{n,t}(v)\}$ is a stationary martingale difference sequence (m.d.s.) of $n \times d$ random matrices.* (*iii*) *$E[\|v_{it}\|^2|\mathcal{F}_{n,t-1}(v)] = \sigma_{v,i}^2$ a.s. for each i and* $\max_{1\le i\le n} E\|v_{it}\|^4 < c_v < \infty$. *There exist $d \times d$ positive definite matrices Σ_v and Σ_v^* such that*

$$\frac{1}{n}\sum_{i=1}^{n}E(v_{it}v_{it}') \to \Sigma_v,\ \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}E(v_{it}v_{jt}') \to \Sigma_v^*, \quad and \quad E\left\|\sum_{i=1}^{n}v_{it}\right\|^{\delta} = O(n^{\delta 2}),$$

for some $\delta > 2$.

ASSUMPTION 3.2. (*a*) *Let $\varepsilon_t \equiv (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$ for $t = 1, \ldots, T$. $\{\varepsilon_t, t \ge 1\}$ is a stationary sequence.* (*b*) *$\{\varepsilon_t, \mathcal{F}_{n,t}(\varepsilon)\}$ is an m.d.s. such that $E(\varepsilon_{it}|\mathcal{F}_{n,t-1}(\varepsilon)) = 0$ a.s. for each i.* (*c*) *Let $\xi_{it} \equiv \varepsilon_{it}^2 - \sigma_i^2$. There exists an even number $\lambda \ge 4$ such that*

$$\frac{1}{nT^{\lambda 2}}\sum_{i=1}^{n}\sum_{1\le t_1,t_2,\ldots,t_\lambda\le T}E(\xi_{it_1}\xi_{it_2}\ldots\xi_{it_\lambda}) < \infty.$$

(d) ε_{it} *is independent of* v_{js} *for all* i, j, t, s. *(e) There exists a* $d \times d$ *positive definite matrix* $\Sigma_{v\varepsilon}$ *such that as* $n \to \infty$,

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} E(v_{i1}v_{j1}')E(\varepsilon_{i1}\varepsilon_{j1}) \to \Sigma_{v\varepsilon}.$$

ASSUMPTION 3.3. *The trend functions* $f_i(\cdot)$ *and* $g_i(\cdot)$ *have continuous derivatives up to the* $(p+1)$*th order.*

ASSUMPTION 3.4. *The kernel functions* $k(\cdot)$ *and* $w(\cdot)$ *are continuous and symmetric p.d.f.'s with compact support* $[-1, 1]$.

ASSUMPTION 3.5. *As* $(n, T) \to \infty, b \to 0, h \to 0$, $\sqrt{nb^{-1}h^2}/\log(nT) \to \infty$, $\min(Tb, nh^{1/2}) \to \infty$, $n^{1/2}Th^{2p+2} \to 0$, *and* $n^{1/2+2/\lambda}T^{-1} \to 0$.

REMARK 3.1. Assumption 3.1 is similar to Assumption 3.2 in CGL (2011). Like CGL (2011), we allow for cross-sectional dependence in $\{v_{it}\}$ and the degree of cross-sectional dependence is controlled by the moment conditions in Assumption 3.1. (c). Unlike CGL (2011), we allow $\{X_{it}\}$ to possess heterogeneous time trends $\{g_i\}$ in (3.1), and we relax their i.i.d. assumption of v_t to the m.d.s. condition. Assumption 3.2 specifies conditions on $\{\varepsilon_{it}\}$ and their interaction with $\{v_{it}\}$. Note that we allow for cross-sectional dependence in $\{\varepsilon_{it}\}$ but rule out serial dependence in Assumption 3.2. (b). To facilitate the derivation of the asymptotic variance of our test statistic, we also impose time-invariant conditional correlations among all cross-sectional units in Assumption 3.2. (c). Assumption 3.2. (d) is readily satisfied under suitable mixing conditions together with moment conditions. The independence between $\{\varepsilon_{it}\}$ and $\{v_{it}\}$ in Assumption 3.2. (e) can be relaxed by modifying the proofs in CGL (2011) significantly. Assumption 3.3 is standard for local polynomial regressions. Assumption 3.4. is a mild and commonly-used condition in the non-parametrics literature. Assumption 3.5 specifies conditions on the bandwidths h and b and sample sizes n and T. Note that we allow $n/T \to c \in [0, \infty]$ as $(n, T) \to \infty$. If we use the optimal rate of bandwidths, i.e. $h \propto (nT)^{-1/(2p+3)}$ in the p-th order local polynomial regression and $b \propto T^{-1/5}$ in the local linear regression, then Assumption 3.5 requires

$$\frac{n^{4p+5}}{T} \to \infty,\ \frac{n^{\frac{1}{2}-\frac{1}{2p+3}}T^{\frac{1}{10}-\frac{1}{2p+3}}}{\log(nT)} \to \infty,\ \frac{(nT)^{\frac{1}{2p+3}}}{n^{1/2}} \to 0, \text{ and } \frac{n^{1/2+2/\lambda}}{T} \to 0.$$

More specifically, if we choose $p = 3$, then Assumption 3.5 implies: $n^{7/18}/(T^{1/90}\log(nT)) \to \infty$, $T/n^{3.5} \to 0$, and $n^{1/2+2/\lambda}/T \to 0$. If $n \propto T^a$, Assumption 3.5 requires $a \in (2/7, 1/(0.5 + 2/\lambda))$.

3.2. Asymptotic null distribution

Let $\bar{H}_{ts}$ denote the (t, s)th element of $\bar{H}$, $\alpha_{ts} \equiv T\bar{H}_{ts} - 1$, and $Q \equiv T^{-1}\mathrm{diag}(\alpha_{11}, \ldots, \alpha_{TT})$. Define

$$B_{nT} \equiv \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \frac{\varepsilon_i' Q \varepsilon_i}{T^{-1} TSS_i},$$

$$\Omega_{nT} \equiv \frac{2b}{T^2} \sum_{1 \le t \ne s \le T} \alpha_{ts}^2 \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij}^2 \right), \quad \text{where } \rho_{ij} \equiv \omega_{ij} \sigma_i^{-1} \sigma_j^{-1}$$

$$\Gamma_{nT} \equiv n^{1/2} T b^{1/2} \overline{R}^2 - B_{nT} = \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \frac{ESS_i - \varepsilon_i' Q \varepsilon_i}{T^{-1} TSS_i}.$$

The following theorem gives the asymptotic null distribution of Γ_{nT}.

THEOREM 3.1. *Suppose Assumptions 3.1–3.5 hold. Then under* H_0,

$$\Gamma_{nT} \xrightarrow{d} N(0, \Omega_0)$$

where $\Omega_0 \equiv \lim_{(n,T)\to\infty} \Omega_{nT}$.

REMARK 3.2. The proof of the above theorem is lengthy and involves several subsidiary propositions, which are given in Appendix A. Under the null hypothesis, we first demonstrate that $\Gamma_{nT} = \Gamma_{nT,1} + o_P(1)$, where $\Gamma_{nT,1} \equiv \sum_{i=1}^{n} \varphi_i(\varepsilon_i)$ and $\varphi_i(\varepsilon_i) = n^{-1/2} T^{-1} b^{1/2} \sum_{1 \le t < s \le T} \alpha_{ts} \varepsilon_{it} \varepsilon_{is} / \sigma_i^2$. Then we apply the martingale central limit theorem (CLT) to show that $\Gamma_{nT,1} \xrightarrow{d} N(0, \Omega_0)$. In general, Γ_{nT} is not asymptotically pivotal as cross-sectional dependence enters its asymptotic variance Ω_0. Nevertheless, if cross-sectional dependence is absent, then Γ_{nT} is an *asymptotic pivotal* test because now $\Omega_0 = \lim_{(n,T)\to\infty} \frac{2b}{T^2} \sum_{1 \le t \ne s \le T} \alpha_{ts}^2$, which is free of nuisance parameters. This is one advantage to base a test on the scale-free non-parametric R^2 measure.

To implement the test, we need to estimate both the asymptotic bias and variance terms. Let

$$\widehat{B}_{nT} \equiv \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \frac{\widehat{u}_i' M Q M \widehat{u}_i}{TSS_i / T} \quad \text{and} \quad \widehat{\Omega}_{nT} \equiv \frac{2b}{T^2} \sum_{1 \le t \ne s \le T} \alpha_{ts}^2 \left(\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \widehat{\rho}_{ij}^2 \right),$$

where $\widehat{\rho}_{ij} \equiv \hat{\omega}_{ij} / (\widehat{\sigma}_i \hat{\sigma}_j)$, $\hat{\omega}_{ij} \equiv T^{-1} \sum_{t=1}^{T} (\widehat{u}_{it} - \overline{\widehat{u}}_i)(\widehat{u}_{jt} - \overline{\widehat{u}}_j)$, $\widehat{\sigma}_i^2 = T^{-1} \sum_{t=1}^{T} (\widehat{u}_{it} - \overline{\widehat{u}}_i)^2$ and $\overline{\widehat{u}}_i \equiv T^{-1} \sum_{t=1}^{T} \widehat{u}_{it}$. We show in the proof of Corollary 3.1 below that $\widehat{B}_{nT} = B_{nT} + o_P(1)$ and $\widehat{\Omega}_{nT} = \Omega_0 + o_P(1)$. Then we obtain a feasible test statistic as

$$\overline{\Gamma}_{nT} = \frac{n^{1/2} T b^{1/2} \overline{R}^2 - \widehat{B}_{nT}}{\sqrt{\widehat{\Omega}_{nT}}} = \frac{1}{\sqrt{\widehat{\Omega}_{nT}}} \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \frac{ESS_i - \widehat{u}_i' M Q M \widehat{u}_i}{TSS_i / T}. \tag{3.2}$$

COROLLARY 3.1. *Under Assumptions 3.1–3.5,* $\overline{\Gamma}_{nT} \xrightarrow{d} N(0, 1)$.

We then compare $\overline{\Gamma}_{nT}$ with the one-sided critical value z_α, i.e. the upper αth percentile from the standard normal distribution. We reject the null when $\overline{\Gamma}_{nT} > z_\alpha$ at the α significance level.

3.3. Asymptotic distribution under local alternatives

To examine the asymptotic local power of our test, we consider the following sequence of Pitman local alternatives:

$$H_1(\gamma_{nT}): f_i(\tau) = f(\tau) + \gamma_{nT}\Delta_{ni}(\tau) \text{ for all } \tau \in [0,1] \text{ and } i = 1,\ldots,n, \tag{3.3}$$

where $\gamma_{nT} \to 0$ as $(n,T) \to \infty$ and $\Delta_{ni}(\cdot)$ is a continuous function on [0, 1]. Let $\Delta_{ni} \equiv (\Delta_{ni}(1/T),\ldots,\Delta_{ni}(T/T))'$. Define

$$\Theta_0 \equiv \lim_{(n,T)\to\infty} \frac{1}{nT}\sum_{i=1}^{n} \Delta_{ni}'(\bar{H} - L)\Delta_{ni}/\sigma_i^2.$$

In the Appendix, we show that $\Theta_0 = C_w \lim_{n\to\infty}(n^{-1}\sum_{i=1}^n \int_0^1 \Delta_{ni}^2(\tau)d\tau/\sigma_i^2)$, where $C_w \equiv \int_{-1}^1\{\int_{-1}^1[1+\omega_2^{-1}u(u-v)]\ w(u)w(u-v)du\ [\int_{-1}^1 w(z-v)dz]^{-1} - 1\}dv$ and $\omega_2 \equiv \int_{-1}^1 w(u) u^2 du$.

To derive the asymptotic property of our test under the alternatives, we add the following assumption.

ASSUMPTION 3.6. $n^{-1}\sum_{i=1}^n \int_0^1 |g_i(\tau) - \overline{g}(\tau)|d\tau = o(1)$ *where* $\overline{g}(\cdot) \equiv n^{-1}\sum_{i=1}^n g_i(\cdot)$.

That is, the non-parametric trending functions $\{g_i(\cdot), 1 \le i \le n\}$ that appear in Assumption 3.1 are *asymptotically homogeneous*. This assumption is needed to determine the probability order of $\widehat{\beta} - \beta$ under $H_1(\gamma_{nT})$ and H_1. Without Assumption 3.6, we can only show that $\widehat{\beta} - \beta = O_P(\gamma_{nT})$ under $H_1(\gamma_{nT})$ and that $\widehat{\beta} - \beta = O_P(1)$ under H_1 for γ_{nT} that converges to zero no faster than $n^{-1/2}T^{-1/2}$. With Assumption 3.6, we demonstrate in Lemma B.6 that $\widehat{\beta} - \beta = o_P(\gamma_{nT})$ under $H_1(\gamma_{nT})$ and that $\widehat{\beta} - \beta = o_P(1)$ under H_1, which are sufficient for us to establish the local power property and the global consistency of our test respectively in Theorems 3.2 and 3.3.

The following theorem establishes the local power property of our test.

THEOREM 3.2. *Suppose Assumptions 3.1–3.6 hold. Suppose that* $\Delta_{ni}(\cdot)$ *is a continuous function such that* $\sum_{i=1}^n \Delta_{ni}(\tau) = 0$ *for* $\tau \in [0,1]$ *and* $\sup_{n\ge1}\max_{1\le i\le n}\int_0^1 \Delta_{ni}^2(\tau)d\tau < \infty$. *Then with* $\gamma_{nT} = n^{-1/4}T^{-1/2}b^{-1/4}$ *in* (3.3) *the local power of our test satisfies*

$$P(\overline{\Gamma}_{nT} > z_\alpha | H_1(\gamma_{nT})) \to 1 - \Phi(z_\alpha - \Theta_0/\sqrt{\Omega_0}),$$

where $\Phi(\cdot)$ *is the cumulative distribution function (CDF) of the standard normal distribution.*

REMARK 3.3. Theorem 3.2 implies that our test has non-trivial asymptotic power against alternatives that diverge from the null at the rate $n^{-1/4}T^{-1/2}b^{-1/4}$. The power increases with the magnitude of Θ_0. Clearly, as either n or T increases, the power of our test will increase but it increases faster as $T \to \infty$ than as $n \to \infty$ for the same choice of b.

3.4. Consistency of the test

To study the consistency of our test, we take $\gamma_{nT} = 1$ and $\Delta_{ni}(\tau) = \Delta_i(\tau)$ in (3.3), where $\Delta_i(\cdot)$ is a continuous function on [0, 1] such that $\underline{c}_\Delta \le n^{-1}\sum_{i=1}^n \int_0^1 \Delta_i(\tau)^2 d\tau \le \overline{c}_\Delta$ for some

$0 < \underline{c}_\Delta < \overline{c}_\Delta < \infty$. Let $\Delta_i \equiv (\Delta_i(1/T), \ldots, \Delta_i(T/T))'$. Define

$$\Theta_A \equiv \lim_{(n,T)\to\infty} \frac{1}{nT} \sum_{i=1}^{n} \Delta_i'(\bar{H} - L)\Delta_i/\overline{\sigma}_i^2,$$

where $\overline{\sigma}_i^2 \equiv \sigma_i^2 + \int_0^1 \Delta_i(\tau)^2 d\tau - (\int_0^1 \Delta_i(\tau) d\tau)^2$. The following theorem establishes the consistency of the test.

THEOREM 3.3. *Suppose Assumptions 3.1–3.6 hold. Under H_1,*

$$n^{-1/2}T^{-1}b^{-1/2}\overline{\Gamma}_{nT} = \Theta_A + o_P(1).$$

Theorem 3.3 implies that under H_1, $P(\overline{\Gamma}_{nT} > d_{nT}) \to 1$ as $(n, T) \to \infty$ for any sequence $d_{nT} = o(n^{1/2}Tb^{1/2})$ provided $\Theta_A > 0$, thus establishing the global consistency of the test.

3.5. *A bootstrap version of the test*

It is well known that asymptotic normal distribution of many non-parametric tests may not approximate their finite sample distributions well in practice. Therefore, we now propose a fixed-regressor bootstrap method (e.g. Hansen, 2000) to obtain the bootstrap approximation to the finite sample distribution of our test statistic under the null.

We propose to generate the bootstrap version of our test statistic $\overline{\Gamma}_{nT}$ as follows:

1. Obtain the augmented residuals $\widehat{u}_{it} = Y_{it} - \widehat{f}(t/T) - X_{it}'\widehat{\beta}$, where $\widehat{f}$ and $\widehat{\beta}$ are obtained by the profile least squares estimation of the restricted model. Calculate the test statistic $\overline{\Gamma}_{nT}$.
2. Let $\overline{\widehat{u}}_i \equiv T^{-1}\sum_{t=1}^{T}\widehat{u}_{it}$ and $\widehat{u}_t \equiv (\widehat{u}_{1t} - \overline{\widehat{u}}_1, \ldots, \widehat{u}_{nt} - \overline{\widehat{u}}_n)'$. Obtain the bootstrap error u_t^* by random sampling with replacement from $\{\widehat{u}_s, s = 1, 2, \ldots, T\}$. Generate the bootstrap analogue of Y_{it} by holding X_{it} as fixed: $Y_{it}^* = \widehat{f}(t/T) + X_{it}'\widehat{\beta} + \overline{\widehat{u}}_i + u_{it}^*$ for $i = 1, \ldots, n$ and $t = 1, \ldots, T$, where u_{it}^* is the ith element in the n-vector u_t^*.
3. Based on the bootstrap resample $\{Y_{it}^*, X_{it}\}$, run the profile least squares estimation of the restricted model to obtain the bootstrap augmented residuals $\{\widehat{u}_{it}^*\}$.
4. Based on $\{\widehat{u}_{it}^*\}$, compute the bootstrap test statistic $\overline{\Gamma}_{nT}^* \equiv (Tn^{1/2}b^{1/2}\overline{R}^{2*} - \widehat{B}_{nT}^*)/\sqrt{\widehat{\Omega}_{nT}^*}$, where $\overline{R}^{2*}$, $\widehat{B}_{nT}^*$ and $\widehat{\Omega}_{nT}^*$ are defined analogously to $\overline{R}^2$, $\widehat{B}_{nT}$ and $\widehat{\Omega}_{nT}$, respectively, but with $\widehat{u}_{it}$ being replaced by $\widehat{u}_{it}^*$.
5. Repeat Step 2–4 for B times and index the bootstrap statistics as $\{\overline{\Gamma}_{nT,l}^*\}_{l=1}^B$. The bootstrap P-value is calculated by $p^* \equiv B^{-1}\sum_{l=1}^{B} 1\{\overline{\Gamma}_{nT,l}^* > \overline{\Gamma}_{nT}\}$, where $1\{\cdot\}$ is the usual indicator function.

Some facts are worth mentioning: (a) Conditionally on the original sample $\mathcal{W} \equiv \{(Y_{it}, X_{it}), i = 1, \ldots, n, t = 1, \ldots, T\}$, the bootstrap replicates u_{it}^* are dependent among cross-sectional units, and i.i.d. across time for fixed i; (b) the regressor X_{it} is held fixed during the bootstrap procedure and (c) the null hypothesis of common trends is imposed in Step 2.

4. SIMULATIONS AND APPLICATIONS

This section conducts a small set of simulations to assess the finite sample performance of the test. We then report empirical applications of the common trends test to UK climate change data and OECD real GDP growth data.

4.1. Simulation study

4.1.1. Data generating processes. We generate data according to six data generating processes (DGPs), among which DGPs 1–2 are used for the level study, and DGPs 3–6 are for the power study.

DGP 1:

$$y_{it} = x_{it}\beta + \left[\left(\frac{t}{T}\right)^3 + \frac{t}{T}\right] + \alpha_i + \varepsilon_{it},$$

where $i = 1, \ldots, n, t = 1, \ldots, T$, $\beta = 2$, for each i we generate x_{it} as i.i.d. $U(a_i - 3, a_i + 3)$ across t with a_i being i.i.d. $N(0, 1)$, $\alpha_i = T^{-1}\sum_{t=1}^{T} x_{it}$ for $i = 2, \ldots, n$, and $\alpha_1 = -\sum_{i=2}^{n}\alpha_i$.

DGP 2:

$$y_{it} = x_{it,1}\beta_1 + x_{it,2}\beta_2 + \left[2\left(\frac{t}{T}\right)^2 + \frac{t}{T}\right] + \alpha_i + \varepsilon_{it},$$

where $i = 1, \ldots, n, t = 1, \ldots, T$, $\beta_1 = 1$, $\beta_2 = 1/2$, $x_{it,1} = 1 + \sin(\pi t/T) + v_{it,1}$, $x_{it,2} = 0.5t/T + v_{i2,t}$, $v_{it,1}$ and $v_{it,2}$ are each i.i.d. $N(0, 1)$ and independent of each other, $\alpha_i = \max(T^{-1}\sum_{t=1}^{T} x_{it,1}, T^{-1}\sum_{t=1}^{T} x_{it,2})$ for $i = 2, \ldots, n$, and $\alpha_1 = -\sum_{i=2}^{n}\alpha_i$.

DGP 3:

$$y_{it} = x_{it}\beta + \left[(1 + \delta_{i1})\left(\frac{t}{T}\right)^3 + (1 + \delta_{i2})\frac{t}{T}\right] + \alpha_i + \varepsilon_{it},$$

where $i = 1, \ldots, n$, $t = 1, \ldots, T$, β, x_{it}, and α_i are generated as in DGP 1, and δ_{i1} and δ_{i2} are each i.i.d. $U(-1/2, 1/2)$, mutually independent and independent of x_{it} and α_i.

DGP 4:

$$y_{it} = x_{it,1}\beta_1 + x_{it,2}\beta_2 + \left[(2 + \delta_{i1})\left(\frac{t}{T}\right)^2 + (1 + \delta_{i2})\frac{t}{T}\right] + \alpha_i + \varepsilon_{it},$$

where $i = 1, \ldots, n, t = 1, \ldots, T$, β_1, β_2, $x_{it,1}$, $x_{it,2}$, and α_i are generated as in DGP 2, and δ_{i1} and δ_{i2} are each i.i.d. $U(-1/2, 1/2)$, mutually independent and independent of $(x_{it,1}, x_{it,2}, \alpha_i)$.

DGP 5:

$$y_{it} = x_{it}\beta + \left[\left(1 + \delta_{nT,i1}\right)\left(\frac{t}{T}\right)^3 + \left(1 + \delta_{nT,i2}\right)\frac{t}{T}\right] + \alpha_i + \varepsilon_{it},$$

where $i = 1, \ldots, n$, $t = 1, \ldots, T$, β, x_{it}, and α_i are generated as in DGP 1, and $\delta_{nT,i1}$ and $\delta_{nT,i2}$ are each i.i.d. $U(-7\gamma_{nT}, 7\gamma_{nT})$, mutually independent, and independent of x_{it} and α_i.

DGP 6:

$$y_{it} = x_{it,1}\beta_1 + x_{it,2}\beta_2 + \left[\left(1 + \delta_{nT,i1}\right)\left(\frac{t}{T}\right)^2 + \left(1 + \delta_{nT,i2}\right)\frac{t}{T}\right] + \alpha_i + \varepsilon_{it},$$

where $i = 1, \ldots, n$, $t = 1, \ldots, T$, β_1, β_2, $x_{it,1}$, $x_{it,2}$, and α_i are generated as in DGP 2, and $\delta_{nT,i1}$ and $\delta_{nT,i2}$ are each i.i.d. $U(-7\gamma_{nT}, 7\gamma_{nT})$, mutually independent and independent of $(x_{it,1}, x_{it,2}, \alpha_i)$.

Note that DGPs 5–6 are used to examine the finite sample behaviour of our test under the sequence of Pitman local alternatives. For both DGPs, we set $\gamma_{nT} = n^{-1/4}T^{-1/2}T^{1/20}$ by choosing $b = T^{-1/5}$, and keep $\{\delta_{nT,i1}\}$and $\{\delta_{nT,i2}\}$fixed through the simulations. Similarly, $\{\delta_{i1}\}$ and $\{\delta_{i2}\}$ are kept fixed through the simulations for DGPs 3–4.

In all of the above DGPs, we generate $\{\varepsilon_{it}\}$ analogously to that in CGL (2011) and independently of all other variables on the right hand side of each DGP. Specifically, we generate ε_t as i.i.d. n-dimensional vector of Gaussian variables with zero mean and covariance matrix $(\omega_{ij})_{n\times n}$. We consider two configurations for $(\omega_{ij})_{n\times n}$:

$$\text{CD (I)} : \omega_{ij} = 0.5^{|j-i|}\sigma_i\sigma_j \quad \text{and} \quad \text{CD (II)}: \omega_{ij} = 0.8^{|j-i|}\sigma_i\sigma_j,$$

where $i, j = 1, \ldots, n$, and σ_i are i.i.d. $U(0, 1)$. By construction, $\{\varepsilon_{it}\}$ are independent across t and cross-sectionally dependent across i.

4.1.2. Test results. To implement our test, we need to choose two kernel functions and two bandwidth sequences. We choose both k and w to be the Epanechnikov kernel: $k(v) = w(v) = 0.75(1 - v^2)1\{|v| \leq 1\}$. To estimate the restricted semi-parametric model, we use the third order local polynomial regression and adopt the 'leave-one-out' cross-validation method to select the bandwidth h. To run the local linear regression of $\widehat{u}_{it}$ on t/T for each cross-sectional unit i, we set $b = c\sqrt{1/12}T^{-1/5}$ for $c = 0.5$, 1 and 1.5 to examine the sensitivity of our test to the choice of bandwidth.[3]

We consider $n, T = 25, 50, 100$. For each combination of n and T, we use 500 replications for both level and power study and 200 bootstrap resamples in each replication.

Table 1 reports the finite sample level of our test when the nominal level is 5%. From Table 1, we see that the levels of our test behave reasonably well except when n/T is big (e.g. $(n, T) = (50, 25)$ or $(100, 25)$). In the latter case, our test is undersized. For fixed n, as T increases, the level of our test approaches the nominal level fairly fast. We also note that the size of our test is robust to different choices of bandwidth.

Tables 2 reports the finite sample power of our test against global alternatives at the 5% nominal level. There is no time trend in the regressor x_{it} in DGP 3 whereas both regressors $x_{it,1}$ and $x_{it,2}$ contain a time trend component in DGP 4. We summarize some important findings from Table 2. First, as either n or T increases, the power of our test generally increases and finally reaches 1, but it increases faster as T increases than as n increases. This is compatible with our asymptotic theory. Secondly, comparing the power behaviour of our test under CD (I) and

[3] Here, the time trend regressor $\{t/T, t = 1, 2, \ldots, T\}$ can be regarded as uniformly distributed on the interval (0, 1) and thus has variance 1/12.

Table 1. Finite sample rejection frequency for DGPs 1–2.

			CD (I)			CD (II)		
DGP	n	T	$c = 0.5$	$c = 1$	$c = 1.5$	$c = 0.5$	$c = 1$	$c = 1.5$
1	25	25	0.036	0.038	0.038	0.034	0.028	0.032
		50	0.038	0.044	0.036	0.032	0.038	0.030
		100	0.046	0.054	0.052	0.042	0.042	0.056
	50	25	0.014	0.028	0.042	0.030	0.028	0.030
		50	0.034	0.056	0.054	0.038	0.044	0.044
		100	0.056	0.048	0.046	0.042	0.038	0.054
	100	25	0.018	0.024	0.022	0.018	0.028	0.028
		50	0.038	0.030	0.024	0.048	0.052	0.048
		100	0.052	0.038	0.054	0.042	0.050	0.048
2	25	25	0.048	0.050	0.050	0.036	0.022	0.038
		50	0.046	0.040	0.054	0.034	0.026	0.038
		100	0.056	0.064	0.072	0.030	0.038	0.062
	50	25	0.026	0.024	0.036	0.018	0.026	0.042
		50	0.056	0.056	0.062	0.040	0.036	0.046
		100	0.056	0.066	0.054	0.044	0.044	0.058
	100	25	0.014	0.016	0.016	0.020	0.022	0.036
		50	0.044	0.032	0.028	0.022	0.034	0.042
		100	0.042	0.046	0.058	0.032	0.040	0.040

CD (II) indicates that the degree of cross-sectional dependence in the error terms has negative impact on the power of our test. This is as expected, as stronger cross-sectional dependence implies less information in each additional cross-sectional observation. Third, the choice of the bandwidth b has some effect on the power of our test. Surprisingly, a larger value of b is associated with a larger testing power.

Table 3 reports the finite sample power of our test against Pitman local alternatives at the 5% nominal level. From the table, we see that our test has non-trivial power to detect the local alternatives at the rate $n^{-1/4}T^{-1/2}b^{-1/4}$, which confirms the asymptotic result in Theorem 3.2. As either n or T increases, we observe the alteration of the local power, which, unlike the case of global alternatives, does not necessarily increase.

4.2. Applications to real data

In this subsection, we apply our test to two real data sets to illustrate its power to detect deviations from common trends, one is to UK climate change data and the other is to OECD economic growth data.

4.2.1. UK climate change data. The issue of global warming has received a lot of attention recently. Atak et al. (2011) develop a semi-parametric model to describe the trend in UK regional temperatures and other weather outcomes over the last century, where a single common trend is

Table 2. Finite sample rejection frequency for DGPs 3–4.

			CD (I)			CD (II)		
DGP	n	T	$c = 0.5$	$c = 1$	$c = 1.5$	$c = 0.5$	$c = 1$	$c = 1.5$
3	25	25	0.294	0.486	0.650	0.128	0.184	0.336
		50	0.502	0.710	0.840	0.182	0.326	0.454
		100	0.938	0.996	0.998	0.580	0.888	0.980
	50	25	0.196	0.424	0.606	0.072	0.136	0.224
		50	0.700	0.936	0.982	0.268	0.496	0.654
		100	1.000	1.000	1.000	0.924	0.996	1.000
	100	25	0.456	0.806	0.938	0.162	0.336	0.494
		50	0.912	1.000	1.000	0.462	0.756	0.898
		100	1.000	1.000	1.000	0.910	0.998	1.000
4	25	25	0.288	0.530	0.730	0.124	0.206	0.344
		50	0.432	0.674	0.788	0.156	0.308	0.434
		100	0.790	0.948	0.988	0.348	0.656	0.816
	50	25	0.352	0.732	0.900	0.142	0.282	0.424
		50	0.802	0.962	0.988	0.336	0.586	0.776
		100	1.000	1.000	1.000	0.926	0.996	0.998
	100	25	0.334	0.712	0.884	0.126	0.234	0.384
		50	0.972	0.996	1.000	0.500	0.824	0.946
		100	1.000	1.000	1.000	0.926	0.996	1.000

assumed across all locations.[4] It is interesting to check whether such a common trend restriction is satisfied. To conserve space, in this application we investigate the pattern of climate change in the UK over the last 32 years. The data set contains monthly mean maximum temperature (in Celsius degrees, *Tmax* for short), mean minimum temperature (in Celsius degrees, *Tmin* for short), total rainfall (in millimeters, *Rain* for short) from 37 stations covering the UK (available from the UK Met Office at: www.metoce.gov.uk/climate/uk/stationdata). According to data availability we adopt a *balanced* panel data set that spans from October 1978 to July 2010 for 26 selected stations ($n = 26$, $T = 382$) to see if there exists a single common trend among these selected stations in *Tmax*, *Tmin*, and *Rain*, respectively. Note that the time span for our data set is much shorter than that in Atak et al. (2011).

For each series we consider a model of the following form

$$y_{it} = D_t'\beta + f_i\left(\frac{t}{T}\right) + \alpha_i + \varepsilon_{it},\ i = 1, \ldots, 26,\ T = 1, \ldots, 382,$$

where y_{it} is *Tmax*, *Tmin*, or *Rain* for station i at time t, $D_t \in \mathbb{R}^{11}$ is a 11-dimensional vector of monthly dummy variables, α_i is the fixed effect for station i, and the time trend function $f_i(\cdot)$ is unknown. We are interested in testing for $f_i = f$ for all $i = 1, 2, \ldots, n$.

[4] Atak et al. (2011) study a model that allows for heterogeneous effects of seasonal dummy variables and use different data sets than ours. Consequently, our result is not directly comparable with theirs.

Table 3. Finite sample rejection frequency for DGPs 5–6.

DGP	n	T	γ_{nT}	CD (I)			CD (II)		
				$c = 0.5$	$c = 1$	$c = 1.5$	$c = 0.5$	$c = 1$	$c = 1.5$
5	25	25	0.1051	0.550	0.862	0.954	0.280	0.532	0.758
		50	0.0769	0.574	0.796	0.876	0.218	0.390	0.542
		100	0.0563	0.884	0.978	0.994	0.532	0.800	0.916
	50	25	0.0883	0.436	0.774	0.928	0.200	0.344	0.530
		50	0.0647	0.662	0.890	0.952	0.234	0.422	0.554
		100	0.0473	0.878	0.976	0.998	0.336	0.556	0.708
	100	25	0.0743	0.410	0.770	0.926	0.146	0.272	0.416
		50	0.0544	0.612	0.884	0.954	0.198	0.332	0.474
		100	0.0398	0.664	0.892	0.960	0.212	0.346	0.516
6	25	25	0.1051	0.570	0.896	0.956	0.288	0.574	0.796
		50	0.0769	0.494	0.764	0.876	0.192	0.354	0.538
		100	0.0563	0.878	0.976	0.994	0.386	0.408	0.770
	50	25	0.0883	0.488	0.836	0.936	0.178	0.366	0.544
		50	0.0647	0.702	0.914	0.980	0.232	0.416	0.580
		100	0.0473	0.886	0.976	0.996	0.352	0.622	0.796
	100	25	0.0743	0.350	0.702	0.902	0.130	0.276	0.422
		50	0.0544	0.640	0.924	0.976	0.282	0.468	0.624
		100	0.0398	0.722	0.918	0.962	0.290	0.472	0.662

Table 4. Bootstrap P-values for application to the U.K. climate data.

Series \ c	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5
Tmax	0.006	0.010	0.007	0.008	0.006	0.007	0.009	0.011	0.015	0.024
Tmin	0.014	0.016	0.015	0.013	0.010	0.005	0.004	0.003	0.002	0.001
Rain	0.873	0.816	0.737	0.659	0.592	0.567	0.573	0.589	0.627	0.679

Note: bandwidth $b = c\sqrt{1/12}T^{-1/5}$ and bootstrap number B = 10 000.

To implement our test, the Epanechnikov kernel is used in both stages. We choose bandwidth h by the "leave-one-out" cross-validation method and consider 10 different bandwidths of the form: $b = c\sqrt{1/12}T^{-1/5}$, where $c = 0.6, 0.7, \ldots, 1.5$. 10 000 bootstrap resamples are used to construct the bootstrap distribution.

The results are reported in Table 4. From the table, we see that the p -values are smaller than 0.05 for *Tmax* and *Tmin* and larger than 0.1 for *Rain* for all choices of b. We can reject the null hypothesis of common trends at 5% level for both *Tmax* and *Tmin* but not for *Rain* even at 10% level.

4.2.2. OECD economic growth data. Economic growth has been a key issue in marcoeconomics over many decades. It is interesting to model the source of economic growth which incorporates a time trend. In this application we consider a model for the OECD economic growth data which incorporates a time trend. The data set consists of four economic

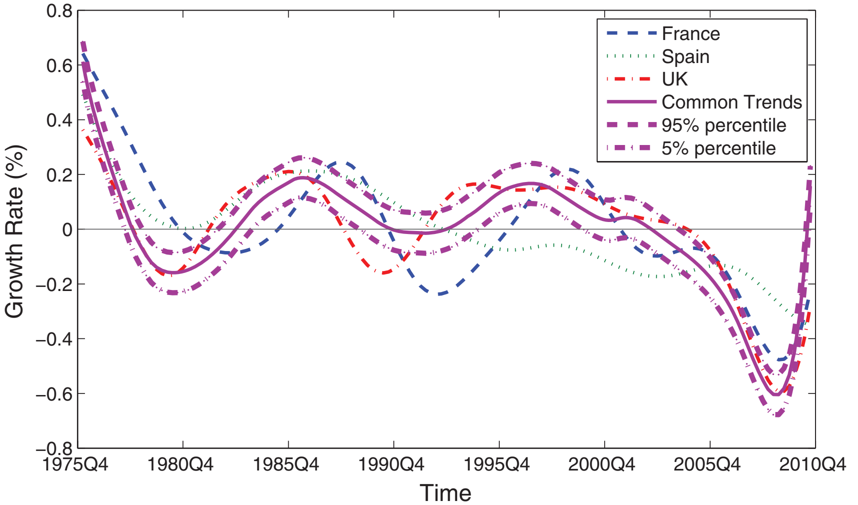

Figure 1. Trends in OECD real GDP growth rates: 1975Q4–2010Q3.

variables from 16 OECD countries ($n = 16$) : Gross domestic product (GDP), Capital Stock (K), Labour input (L), and Human capital (H). We download GDP (at 2005 US$), Capital stock (at 2005 US$), and Labour input (Employment, at thousand persons) from http://www.datastream.com, and Human capital (Educational Attainment for Population Aged 25 and Over) from http://www.barrolee.com. The first three variables are seasonally adjusted quarterly data and span from 1975Q4 to 2010Q3 ($T = 140$). For Human capital, we have only five-years census data from the Barro-Lee data set so that we have to use linear interpolation to obtain the quarterly observations.

We consider the following model for growth rates

$$\Delta \ln GDP_{it} = \beta_1 \Delta \ln L_{it} + \beta_2 \Delta \ln K_{it} + \beta_3 \Delta \ln H_{it} + f_i(t/T) + \alpha_i + \varepsilon_{it},$$

where $i = 1, \ldots, 16$, $T = 1, \ldots, 140$, and α_i is the fixed effect, $f_i(\cdot)$ is unknown smooth time trends function for country i, and $\Delta \ln Z_{it} = \ln Z_{it} - \ln Z_{i,t-1}$ for $Z = GDP$, L, K, and H. We are interested in testing for common time trends for the 16 OECD countries.

The kernels, bandwidths, and number of bootstrap resamples are chosen as in the previous application. In Figure 1 we plot the estimated common trends (where we use the recentred trend: $\widehat{f}(\tau) - \int_0^1 \widehat{f}(\tau) d\tau$ for comparison) from the restricted semi-parametric regression model together with its 90% pointwise confidence bands. Also plotted in Figure 1 are three representative individual trend functions for France, Spain, and the UK, which are estimated from the unrestricted semi-parametric regression models. For the purpose of comparison, for the unconstrained model we impose the identification condition that the integral of each individual trend function over $(0, 1)$ equals zero and use the Silverman rule-of-thumb to choose the bandwidth. Clearly, Figure 1 suggests that the estimated common trends function is significantly different from zero over a wide range its support. In addition, the trend functions for the three representative individual countries are obviously different from the estimated common trends, which implies that the widely used common trends assumption may not be plausible at all.

Table 5. Bootstrap P-values for application to OECD real GDP growth rate data.

Series\ c	0.6	0.7	0.8	0.9	1.0
$\Delta \ln GDP$	0.0001	0.0005	0.0020	0.0063	0.0141
Series\ c	1.1	1.2	1.3	1.4	1.5
$\Delta \ln GDP$	0.0281	0.0336	0.0536	0.0645	0.0820

Note: bandwidth $b = c\sqrt{1/12}T^{-1/5}$ and bootstrap number B $= 10000$.

Table 5 reports the bootstrap P-values for our test of common trends. From the table, we can see that the P-values for all bandwidths are smaller than 0.1 for all bandwidths under investigation. Then we can reject the null hypothesis of common trends at the 10% level.

5. CONCLUDING REMARKS

In this paper, we propose a non-parametric test for common trends in semi-parametric panel data models with fixed effects. We first estimate the restricted semi-parametric model to obtain the augmented residuals and then run a local linear regression of the augmented residuals on the time trend for each cross-sectional unit to obtain n non-parametric R^2 measures. We construct our test statistic by averaging these individual non-parametric R^2's, and show that after being appropriately centred and scaled, the statistic is asymptotically normally distributed under both the null hypothesis of common trends and a sequence of Pitman local alternatives. We also prove the consistency of the test and propose a bootstrap procedure to obtain the bootstrap P-values. Monte Carlo simulations and applications to both the UK climate change data and the OECD economic growth data are reported, both of which point to the empirical fragility of a common trend assumption.

Some extensions are possible. First, our semi-parametric model in (1.1) only complements that in Atak et al. (2011), and it is possible to allow the slope coefficients also to be heterogeneous when we test for the null hypothesis of common trends for the non-parametric component. In this case, the profile least-squares estimation of Su and Ullah (2006) and Chen et al. (2011) and the non-parametric-R^2-based test lose much of their advantage and the heterogeneous slope coefficients can only be estimated at a slower convergence rate. It seems straightforward to estimate the unrestricted model for each cross-sectional unit to obtain the individual trend function estimates $\widehat{f}_i(\tau)$ and propose an L_2-distance-based test by averaging the squared L_2 -distance between $\widehat{f}_i(\tau)$ and $\widehat{f}_j(\tau)$ for all $i \neq j$. It is also possible to test for the homogeneity of the slope coefficients and trend components jointly. Second, to derive the distribution theory of our test statistic, we allow for cross-sectional dependence but rule out serial dependence. It is possible to allow the presence of both as in Bai (2009) by imposing some high-level assumptions. Nevertheless, the asymptotic variance of the non-normalized version of test statistic will become complicated and there seems no obvious way to estimate it consistently in order to implement our test in practice.

ACKNOWLEDGMENTS

We thank the co-editor, Oliver Linton, and two anonymous referees for their valuable suggestions and comments on the original version of this paper. Su acknowledges support from SMU under grant number #10-C244-SMU-009. Phillips acknowledges partial support from the NSF under Grant SES 09-56687.

REFERENCES

Atak, A., O. Linton, and Z. Xiao (2011). A semiparametric panel data model for unbalanced data with application to climate change in the United Kingdom. *Journal of Econometrics 164*, 92–115.

Bacigál, T. (2005). Testing for common deterministic trends in geodetic data. *Journal of Electrical Engineering 12*, 1–5.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica 77*, 1229–79.

Banerjee, A., (1999). Panel data unit roots and cointegration: an overview. *Oxford Bulletin of Economics and Statistics 61*, 607–29.

Bernstein, D. S. (2005). *Matrix Mathematics: Theory, Facts, and Formulas with Application to Linear Systems Theory*. Princeton: Princeton University Press.

Breitung, J. and M. H. Pesaran (2008). Unit roots and cointegration in panels. In L. Matyas and P. Sevestre (Eds.), *The Econometrics of Panel Data* (3rd ed.), 279–322, Berlin: Springer.

Chen, J., J. Gao and D. Li (2011). Semiparametric trending panel data models with cross-sectional dependence. Working paper, University of Adelaide.

Cai, Z. (2007). Trending time-varying coefficients time series models with serially correlated errors. *Journal of Econometrics 136*, 163–88.

Fomby, T. B. and T. J. Vogelsang (2003). Tests of common deterministic trend slopes applied to quarterly temperature data. *Advances in Econometrics 17*, 29–43.

Gao, J. and K. Hawthorne (2006). Semiparametric estimation and testing of the trend of temperature series. *Econometrics Journal 9*, 332–55.

Hansen, B. E. (2000). Testing for structural change in conditional model. *Journal of Econometrics 97*. 93–115.

Li, D., J. Chen and J. Gao (2010). Nonparametric time-varying coefficient panel data models with fixed effects. *Econometrics Journal 14*, 387–408.

Phillips, P. C. B. (2001). Trending time series and macroeconomic activity: some present and future challenges. *Journal of Econometrics 100*, 21–7.

Phillips, P. C. B. (2005). Challenges of trending time series econometrics. *Mathematics and Computers in Simulation 68*, 401–16.

Phillips, P. C. B. (2007). Regression with slowly varying regressors and nonlinear trends. *Econometric Theory 23*, 557–614.

Phillips, P. C. B. (2010). The mysteries of trend. *Macroeconomic Review IX*, 82–9.

Phillips, P. C. B. and D. Sul (2007). Transition modeling and econometric convergence tests. *Econometrica 75*, 1771–855.

Phillips, P. C. B. and D. Sul (2009). Economic transition and growth. *Journal of Applied Econometrics 24*, 1153–85.

Robinson, P. M. (2010). Nonparametric trending regression with cross-sectional dependence. Working paper, London School of Economics.

Stock, J. H. and M. W. Watson (1988). Testing for common trends. *Journal of the American Statistical Association 83*, 1097–107.

Su, L., A. Ullah (2006). Profile likelihood estimation of partially linear panel data models with fixed effects. *Economics Letters 92*, 75–81.

Su, L. and A. Ullah (2011). A nonparametric goodness-of-fit-based test for conditional heteroskedasticity. Forthcoming in *Econometric Theory*.

Sun, Y. (2011). Robust trend inference with series variance estimator and testing-optimal smoothing parameter. *Journal of Econometrics 164*, 345–66.
Vogelsang, T. J. and P. H. Franses (2005). Testing for common deterministic trend slopes. *Journal of Econometrics 126*, 1–24.
White, H. (2001). *Asymptotic Theory for Econometricians* (2nd ed.). San Diego: Academic Press.
White, H. and C. Granger (2011). Consideration of trends in time series. *Journal of Time Series Econometrics 3*, 1–40.
Xu, K.-L. (2011). Robustifying multivariate trend tests to nonstationary volatility. Forthcoming in *Journal of Econometrics*.

APPENDIX A: PROOFS OF MAIN RESULTS

Proof of Theorem 3.1: Noting that

$$\begin{aligned}\Gamma_{nT} &= \sqrt{\frac{b}{n}}\sum_{i=1}^{n}\frac{(ESS_i - \varepsilon_i' Q\varepsilon_i)}{\sigma_i^2} + \sqrt{\frac{b}{n}}\sum_{i=1}^{n}(ESS_i - \varepsilon_i' Q\varepsilon_i)\left(\frac{1}{TSS_i/T} - \frac{1}{\sigma_i^2}\right)\\ &\equiv \Gamma_{nT,1} + \Gamma_{nT,2},\ \text{say,}\end{aligned}$$

we complete the proof by showing that (i) $\Gamma_{nT,1} \xrightarrow{d} N(0, \Omega_0)$, and (ii) $\Gamma_{nT,2} = o_P(1)$. These results are established in Propositions A.1 and A.2, respectively. □

PROPOSITION A.1. $\Gamma_{nT,1} \xrightarrow{d} N(0, \Omega_0)$.

Proof: Decompose

$$\Gamma_{nT,1} = \sqrt{\frac{b}{n}}\sum_{i=1}^{n}\frac{\widehat{u}_i'(\bar{H} - L)\widehat{u}_i}{\sigma_i^2} - \sqrt{\frac{b}{n}}\sum_{i=1}^{n}\frac{\varepsilon_i' Q\varepsilon_i}{\sigma_i^2} \equiv \Gamma_{nT,11} - \Gamma_{nT,12}. \tag{A.1}$$

Let $X_i^* \equiv X_i - S_T X$ and $\varepsilon_i^* \equiv \varepsilon_i - S_T\varepsilon$. Define

$$\bar{\mathbf{f}} \equiv (\overline{f}(1/T), \ldots, \overline{f}(T/T))' \text{ and } \bar{\mathbf{f}}^* \equiv \bar{\mathbf{f}} - S_T\mathbf{F}, \tag{A.2}$$

where $\overline{f}(\tau) \equiv n^{-1}\sum_{i=1}^{n} f_i(\tau)$. Noting that

$$\widehat{u}_i = \varepsilon_i^* - X_i^*(\widehat{\beta} - \beta) + \bar{\mathbf{f}}^* + (\mathbf{f}_i - \bar{\mathbf{f}}) + \alpha_i i_T \tag{A.3}$$

and $Mi_T = 0$, we have

$$\Gamma_{nT,11} = \sqrt{\frac{b}{n}}\sum_{i=1}^{n}\frac{\widehat{u}_i'(\bar{H} - L)\widehat{u}_i}{\sigma_i^2} = \sum_{l=1}^{10} D_{nTl}, \tag{A.4}$$

where

$$D_{nT1} \equiv \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \varepsilon_i^{*\prime}(\bar{H} - L)\varepsilon_i^{*}/\sigma_i^2,$$

$$D_{nT2} \equiv \sqrt{\frac{b}{n}} \sum_{i=1}^{n} (\mathbf{f}_i - \bar{\mathbf{f}})'(\bar{H} - L)(\mathbf{f}_i - \bar{\mathbf{f}})/\sigma_i^2,$$

$$D_{nT3} \equiv \sqrt{\frac{b}{n}} \sum_{i=1}^{n} (\widehat{\beta} - \beta)' X_i^{*\prime}(\bar{H} - L) X_i^{*}(\widehat{\beta} - \beta)/\sigma_i^2,$$

$$D_{nT4} \equiv \upsilon \sum_{i=1}^{n} \bar{\mathbf{f}}^{*\prime}(\bar{H} - L)\bar{\mathbf{f}}^{*}/\sigma_i^2,$$

$$D_{nT5} \equiv -2\sqrt{\frac{b}{n}} \sum_{i=1}^{n} \varepsilon_i^{*\prime}(\bar{H} - L) X_i^{*}(\widehat{\beta} - \beta)/\sigma_i^2,$$

$$D_{nT6} \equiv 2\sqrt{\frac{b}{n}} \sum_{i=1}^{n} \varepsilon_i^{*\prime}(\bar{H} - L)\bar{\mathbf{f}}^{*}/\sigma_i^2,$$

$$D_{nT7} \equiv -2\sqrt{\frac{b}{n}} \sum_{i=1}^{n} (\widehat{\beta} - \beta)' X_i^{*\prime}(\bar{H} - L)\bar{\mathbf{f}}^{*}/\sigma_i^2,$$

$$D_{nT8} \equiv 2\sqrt{\frac{b}{n}} \sum_{i=1}^{n} \varepsilon_i^{*\prime}(\bar{H} - L)(\mathbf{f}_i - \bar{\mathbf{f}})/\sigma_i^2,$$

$$D_{nT9} \equiv -2\sqrt{\frac{b}{n}} \sum_{i=1}^{n} (\widehat{\beta} - \beta)' X_i^{*\prime}(\bar{H} - L)(\mathbf{f}_i - \bar{\mathbf{f}})/\sigma_i^2,$$

$$D_{nT10} \equiv 2\sqrt{\frac{b}{n}} \sum_{i=1}^{n} \bar{\mathbf{f}}^{*\prime}(\bar{H} - L)(\mathbf{f}_i - \bar{\mathbf{f}})/\sigma_i^2.$$

Under H_0, $D_{nTs} = 0$ for $s = 2, 8, 9, 10$. We complete the proof of the proposition by showing that:

$$\mathcal{D}_{nT1} \equiv D_{nT1} - \Gamma_{nT,12} \xrightarrow{d} N(0, \Omega_0), \text{ and} \tag{A.5}$$

$$D_{nTs} = o_P(1), \; s = 3, \ldots, 7. \tag{A.6}$$

Step 1. We first prove (A.5). Noting that $\varepsilon_i^{*} \equiv \varepsilon_i - S_T\varepsilon$, we can decompose $\mathcal{D}_{nT1}$ as:

$$\begin{aligned}
\mathcal{D}_{nT1} &= \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \frac{\varepsilon_i^{*\prime}(\bar{H} - L)\varepsilon_i^{*}}{\sigma_i^2} - \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \frac{\varepsilon_i' Q \varepsilon_i}{\sigma_i^2} \\
&= \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \frac{\varepsilon_i'(\bar{H} - L - Q)\varepsilon_i}{\sigma_i^2} + \sqrt{\frac{b}{n}} \varepsilon' S_T'(\bar{H} - L) S_T \varepsilon \sum_{i=1}^{n} \frac{1}{\sigma_i^2}
\end{aligned}$$

$$
- 2\sqrt{\frac{b}{n}} \sum_{i=1}^{n} \frac{\varepsilon_i'(\bar{H} - L) S_T \varepsilon}{\sigma_i^2}
$$

$$
\equiv \mathcal{D}_{nT11} + \mathcal{D}_{nT12} - 2\mathcal{D}_{nT13}.
$$

We prove (A.5) by showing that $\mathcal{D}_{nT11} \xrightarrow{d} N(0, \Omega_0)$ and $\mathcal{D}_{nT1s} = o_P(1)$ for $s = 2, 3$. The former claim follows from Lemma A.1 below. We now prove the latter claim. Let $\overline{\mathcal{D}}_{nT12} \equiv \sqrt{nb}\varepsilon' S_T'(\bar{H} - L) S_T \varepsilon$. By Lemmata B.2 and B.5, we have

$$
\begin{aligned}
\overline{\mathcal{D}}_{nT12} &= \sqrt{nb} \sum_{t=1}^{T} \sum_{s=1}^{T} (e_1' S(t/T)\varepsilon)(\bar{H}_{ts} - T^{-1})(e_1' S(s/T)\varepsilon) \\
&\leq \sqrt{nb} \max_{1 \leq t \leq T} |e_1' S(t/T)\varepsilon|^2 \sum_{t=1}^{T} \sum_{s=1}^{T} |\bar{H}_{ts} - T^{-1}| \\
&= \sqrt{nb} O_P\left(\frac{\log(nT)}{nTh}\right) O(T) = O_P\left(\frac{\log(nT)}{\sqrt{nb^{-1}h^2}}\right) = o_P(1).
\end{aligned}
$$

Then $\mathcal{D}_{nT12} = o_P(1)$ by Assumption 3.2 (c). For $\mathcal{D}_{nT13}$, we have $\mathcal{D}_{nT13} = n^{-1/2} b^{1/2} \sum_{i=1}^{n} \varepsilon_i'(\bar{H} - L) S_T \varepsilon / \sigma_i^2 = \mathcal{D}_{nT131} + \mathcal{D}_{nT132}$, where

$$
\mathcal{D}_{nT131} \equiv \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \sum_{t=1}^{T} a_{tt} \varepsilon_{it} e_1' S(t/T) \varepsilon \sigma_i^{-2}, \quad \text{and}
$$

$$
\mathcal{D}_{nT132} \equiv \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \sum_{1 \leq s \neq t \leq T} a_{ts} \varepsilon_{it} e_1' S(s/T) \varepsilon \sigma_i^{-2},
$$

and $a_{ts} \equiv \bar{H}_{ts} - T^{-1}$. For $\mathcal{D}_{nT131}$, write

$$
\begin{aligned}
\mathcal{D}_{nT131} &= \frac{b^{1/2}}{n^{3/2}} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{t=1}^{T} a_{tt} \varepsilon_{it} e_1' s(t/T) \varepsilon_j \sigma_i^{-2} \\
&= \frac{b^{1/2}}{Tn^{3/2}} \sum_{1 \leq i,j \leq n} \sum_{1 \leq t,s \leq T} a_{tt} c_{ts} k_{h,ts} \varepsilon_{it} \varepsilon_{js} \sigma_i^{-2} \\
&= \frac{b^{1/2}}{Tn^{3/2}} \sum_{i=1}^{n} \sum_{t=1}^{T} a_{tt} c_{tt} k_{h,tt} \varepsilon_{it}^2 \sigma_i^{-2} \\
&\quad + \frac{b^{1/2}}{Tn^{3/2}} \sum_{i=1}^{n} \sum_{1 \leq t < s \leq T} (a_{tt} c_{ts} + a_{ss} c_{st}) k_{h,ts} \varepsilon_{it} \varepsilon_{is} \sigma_i^{-2} \\
&\quad + \frac{b^{1/2}}{Tn^{3/2}} \sum_{1 \leq i \neq j \leq n} \sum_{t=1}^{T} a_{tt} c_{tt} k_{h,tt} \varepsilon_{it} \varepsilon_{jt} \sigma_i^{-2}
\end{aligned}
$$

$$+\frac{b^{1/2}}{Tn^{3/2}}\sum_{1\le i\ne j\le n}\sum_{1\le t<s\le T}(a_{tt}c_{ts}+a_{ss}c_{st})k_{h,ts}\varepsilon_{it}\varepsilon_{js}\sigma_i^{-2}$$
$$\equiv \mathcal{D}_{nT131a}+\mathcal{D}_{nT131b}+\mathcal{D}_{nT131c}+\mathcal{D}_{nT131d},$$

where $c_{ts}\equiv e_1'[T^{-1}z_h^{[p]}(t/T)'K_h(t/T)z_h^{[p]}(t/T)]^{-1}z_{h,s}^{[p]}(t/T)$. By Lemmata B.2 and B.4(c) and Assumption 3.5, we have

$$E|\mathcal{D}_{nT131a}|\le \frac{k(0)b^{1/2}}{n^{1/2}h}\max_{1\le t\le n}|a_{tt}|\left(\frac{1}{T}\sum_{t=1}^{T}|c_{tt}|\right)$$
$$= n^{-1/2}b^{1/2}h^{-1}O(T^{-1}b^{-1})O(1)=o(1).$$

So $\mathcal{D}_{nT131a}=o_P(1)$ by the Markov inequality. For $\mathcal{D}_{nT131b}$, we have by Lemmata B.2 and B.4 (b)

$$\begin{aligned}
E\left(\mathcal{D}^2_{nT131b}\right) &= \frac{b}{T^2n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{1\le t_1<t_2\le T}\sum_{1\le t_3<t_4\le T}e_{t_1t_2}k_{h,t_1t_2}e_{t_3t_4}k_{h,t_3t_4}E(\varepsilon_{it_1}\varepsilon_{it_2}\varepsilon_{jt_3}\varepsilon_{jt_4})\sigma_i^{-2}\sigma_j^{-2}\\
&= \frac{b}{T^2n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{1\le t_1<t_2\le T}(e_{t_1t_2}k_{h,t_1t_2})^2E(\varepsilon_{it_1}\varepsilon_{it_2}\varepsilon_{jt_1}\varepsilon_{jt_2})\sigma_i^{-2}\sigma_j^{-2}\\
&\le \frac{2b}{T^2n^3}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{1\le t_1<t_2\le T}\left(a_{t_1t_1}^2c_{t_1t_2}^2+a_{t_2t_2}^2c_{t_2t_1}^2\right)k_{h,t_1t_2}^2|E(\varepsilon_{it_1}\varepsilon_{it_2}\varepsilon_{jt_1}\varepsilon_{jt_2})|\sigma_i^{-2}\sigma_j^{-2}\\
&\le \frac{2b}{T^2n^2}\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}^2\right)\sum_{1\le t_1<t_2\le T}\left(a_{t_1t_1}^2c_{t_1t_2}^2+a_{t_2t_2}^2c_{t_2t_1}^2\right)k_{h,t_1t_2}^2\\
&\le \frac{2b}{n^2h}\left(\max_{1\le t\le T}a_{tt}^2\right)\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}^2\right)\left(\frac{h}{T^2}\sum_{1\le t_1\ne t_2\le T}c_{t_1t_2}^2k_{h,t_1t_2}^2\right)\\
&= \frac{2b}{n^2h}O(T^{-2}b^{-2})O(1)=O(n^{-2}T^{-2}b^{-1}h^{-1})=o(1),
\end{aligned}$$

where $e_{ts}\equiv a_{tt}c_{ts}+a_{ss}c_{st}$, $\rho_{ij}\equiv\omega_{ij}\sigma_i^{-1}\sigma_j^{-1}$, and the second equality follows from the fact that $E(\varepsilon_{it_1}\varepsilon_{it_2}\varepsilon_{jt_1}\varepsilon_{jt_3})=0$ and $E(\varepsilon_{it_1}\varepsilon_{it_2}\varepsilon_{jt_3}\varepsilon_{jt_4})=0$ when t_1, t_2, t_3, and t_4 are all distinct by Assumptions 3.2(b)–(c). It follows that $\mathcal{D}_{nT131b}=o_P(1)$ by the Chebyshev inequality. For $\mathcal{D}_{nT131c}$, we have by Lemma B.2 and Assumptions 3.2 and 3.5

$$\begin{aligned}
E\left[\mathcal{D}^2_{nT131c}\right] &= \frac{b}{T^2n^3}\sum_{1\le i_1\ne i_2\le n}\sum_{1\le i_3\ne i_4\le n}\sum_{t=1}^{T}\sum_{s=1}^{T}a_{tt}c_{tt}k_{h,tt}a_{ss}c_{ss}k_{h,ss}E(\varepsilon_{i_1t}\varepsilon_{i_2t}\varepsilon_{i_3s}\varepsilon_{i_4s})\sigma_{i_1}^{-2}\sigma_{i_3}^{-2}\\
&= \frac{bk^2(0)}{T^2n^3h^2}\sum_{1\le i_1\ne i_2\le n}\sum_{1\le i_3\ne i_4\le n}\sum_{1\le t\ne s\le T}a_{tt}c_{tt}a_{ss}c_{ss}\omega_{i_1i_2}\omega_{i_3i_4}\sigma_{i_1}^{-2}\sigma_{i_3}^{-2}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{bk^2(0)}{T^2n^3h^2}\sum_{t=1}^{T}\left[a_{tt}^2c_{tt}^2\sum_{1\le i_1\ne i_2\le n}\sum_{1\le i_3\ne i_4\le n}E(\varepsilon_{i_1t}\varepsilon_{i_2t}\varepsilon_{i_3t}\varepsilon_{i_4t})\sigma_{i_1}^{-2}\sigma_{i_3}^{-2}\right]\\
&\le\frac{b}{nh^2}\Big(\max_{1\le t\le T}a_{tt}^2\Big)\left(\frac{1}{n}\sum_{1\le i_1\ne i_2\le n}\omega_{i_1i_2}\sigma_{i_1}^{-2}\right)^2\left(\frac{1}{T}\sum_{t=1}^{T}|c_{tt}|\right)^2\\
&+\frac{b}{Tnh^2}\Big(\max_{1\le t\le T}a_{tt}^2\Big)\left|\frac{1}{n^2}\sum_{1\le i_1\ne i_2\le n}\sum_{1\le i_3\ne i_4\le n}E(\varepsilon_{i_1t}\varepsilon_{i_2t}\varepsilon_{i_3t}\varepsilon_{i_4t})\sigma_{i_1}^{-2}\sigma_{i_3}^{-2}\right|\left(\frac{1}{T}\sum_{t=1}^{T}c_{tt}^2\right)\\
&=\frac{b}{nh^2}O(T^{-2}b^{-2})O(1)\,O(1)+\frac{b}{Tnh^2}O(T^{-2}b^{-2})O(1)\,O(1)\\
&=O(n^{-1}T^{-2}h^{-2}b^{-1}+n^{-1}T^{-3}b^{-1}h^{-2})=o(1).
\end{aligned}
$$

It follows that $\mathcal{D}_{nT131c}=o_P(1)$ by the Chebyshev inequality. Similarly, $\mathcal{D}_{nT131d}=o_P(1)$ because

$$
\begin{aligned}
E(\mathcal{D}_{nT131d})^2&=\frac{4b}{T^2n^3}\sum_{1\le i_1\ne i_2\le n}\sum_{1\le i_3\ne i_4\le n}\sum_{1\le t_1<t_2\le T}a_{t_1t_1}^2c_{t_1t_2}^2k_{h,t_1t_2}^2E(\varepsilon_{i_1t_1}\varepsilon_{i_2t_2}\varepsilon_{i_3t_1}\varepsilon_{i_4t_2})\sigma_{i_1}^{-2}\sigma_{i_3}^{-2}\\
&=\frac{4b}{T^2n^3}\sum_{1\le i_1\ne i_2\le n}\sum_{1\le i_3\ne i_4\le n}\sum_{1\le t_1<t_2\le T}a_{t_1t_1}^2c_{t_1t_2}^2k_{h,t_1t_2}^2\omega_{i_1i_3}\omega_{i_2i_4}\sigma_{i_1}^{-2}\sigma_{i_3}^{-2}\\
&\le\frac{4\underline{c}^{-2}b}{nh}\Big(\max_{1\le t\le T}a_{tt}^2\Big)\left(\frac{h}{T^2}\sum_{1\le t_1<t_2\le T}c_{t_1t_2}^2k_{h,t_1t_2}^2\right)\left(\frac{1}{n}\sum_{1\le i_1,i_2\le n}|\omega_{i_1i_2}|\right)^2\\
&=\frac{b}{nh}O(T^{-2}b^{-2})O(1)\,O(1)=O(n^{-1}T^{-2}h^{-1}b^{-1})=o(1).
\end{aligned}
$$

In sum, we have shown that $\mathcal{D}_{nT131}=o_P(1)$. For $\mathcal{D}_{nT132}$, we have

$$
\begin{aligned}
\mathcal{D}_{nT132}&=\frac{b^{1/2}}{n^{3/2}}\sum_{1\le i,j\le n}\sum_{1\le s\ne t\le T}a_{ts}\varepsilon_{it}e_1's(s/T)\,\varepsilon_j\sigma_i^{-2}\\
&=\frac{b^{1/2}}{Tn^{3/2}}\sum_{1\le i,j\le n}\sum_{1\le s\ne t\le T}\sum_{r=1}^{T}a_{ts}c_{sr}k_{h,sr}\varepsilon_{it}\varepsilon_{jr}\sigma_i^{-2}\\
&=\frac{b^{1/2}}{Tn^{3/2}}\sum_{1\le i\ne j\le n}\sum_{1\le s\ne t\ne r\le T}a_{ts}c_{sr}k_{h,sr}\varepsilon_{it}\varepsilon_{jr}\sigma_i^{-2}+o_P(1)\\
&\equiv\mathcal{D}_{nT132a}+o_P(1).
\end{aligned}
$$

Following the same arguments as used in the proof of $\mathcal{D}_{nT131a}=o_P(1)$, we can show that $E(\mathcal{D}_{nT132a})^2=o(1)$. It follows that $\mathcal{D}_{nT132a}=o_P(1)$ and $\mathcal{D}_{nT132}=o_P(1)$.

Step 2. We now prove (A.6). For D_{nT3}, by Assumption 3.2(c), and Lemmata B.3, B.6(a) and B.7, we have

$$\begin{aligned}|D_{nT3}| &\leq \underline{c}^{-1}n^{-1/2}b^{1/2}\|\bar{H}-L\|\|\widehat{\beta}-\beta\|^2\sum_{i=1}^{n}\|X_i-S_TX\|^2\\&= \underline{c}^{-1}n^{-1/2}(b^{1/2}\|\bar{H}-L\|)\|\widehat{\beta}-\beta\|^2\|X-S_{nT}X\|^2\\&= n^{-1/2}O(1)O_P(n^{-1}T^{-1})O_P(nT)=O_P(n^{-1/2})=o_P(1).\end{aligned}$$

For D_{nT4}, noting that $\max_{1\leq t\leq T}|\overline{f}(t)-e_1'S(t/T)\mathbf{F}|=O(h^{p+1})$ by analysis analogous to CGL (2011), by Lemma B.3 and Assumption 3.5 we have

$$|D_{nT4}|\leq \underline{c}^{-1}n^{1/2}(b^{1/2}\|\bar{H}-L\|)\|\overline{\mathbf{f}}^*\|^2=O(n^{1/2}Th^{2p+2})=o(1).$$

Now decompose D_{nT5} as follows

$$\begin{aligned}D_{nT5} &= -2\left[\sqrt{\frac{b}{n}}\sum_{i=1}^{n}\varepsilon_i'(\bar{H}-L)X_i^*\sigma_i^{-2}-\sqrt{\frac{b}{n}}\sum_{i=1}^{n}(S_T\varepsilon)'(\bar{H}-L)X_i^*\sigma_i^{-2}\right](\widehat{\beta}-\beta)\\&\equiv -2\,(D_{nT51}-D_{nT52})\,(\widehat{\beta}-\beta),\ \text{say.}\end{aligned}$$

Noting that $D_{nT51}=\sqrt{\frac{b}{n}}\sum_{i=1}^{n}\sigma_i^{-2}\varepsilon_i'(\bar{H}-L)(X_i-S_TX)=\sqrt{\frac{b}{n}}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{s=1}^{T}\sigma_i^{-2}\times\varepsilon_{it}a_{ts}[X_{is}-e_1'S(s/T)X]$, by Assumption 3.2, the Cauchy inequality, and Lemma B.3(c),

$$\begin{aligned}E\|D_{nT51}\|^2 &= \frac{b}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\{E\{\mathrm{tr}[(X_{is}-e_1'S(s/T)X)(X_{jr}-e_1'S(r/T)X)']\}\\&\quad\times a_{ts}a_{tr}\omega_{ij}\sigma_i{}^{-2}\sigma_j{}^{-2}\}\\&\leq Tb\max_{1\leq i\leq n}\max_{1\leq s\leq T}(E\|X_{is}-e_1'S(s/T)X\|^2)\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}|\rho_{ij}|\right)\\&\quad\times\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}|a_{ts}a_{tr}|\right)\\&= TbO\,(1)\,O\,(1)\,O\,(1)=O\,(Tb)\,.\end{aligned}$$

For D_{nT52} we have

$$\begin{aligned}\|D_{nT52}\|^2 &= \frac{b}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{tr}[(\bar{H}-L)X_i^*X_j^{*\prime}(\bar{H}-L)S_T\varepsilon\varepsilon'S_T']\sigma_i^{-2}\sigma_j^{-2}\\&= \frac{b}{n}\mathrm{tr}\left[\left(\sum_{i=1}^{n}\sum_{j=1}^{n}X_i^*X_j^{*\prime}\sigma_i^{-2}\sigma_j^{-2}\right)(\bar{H}-L)S_T\varepsilon\varepsilon'S_T'(\bar{H}-L)\right]\end{aligned}$$

$$\leq \frac{\underline{c}^{-2}}{n}\left(\sum_{i=1}^{n}\|X_i^*\|\right)^2 (b\|\bar{H}-L\|^2)\|S_T\varepsilon\|^2$$

$$= \frac{1}{n}O_P(Tn^2)O(1)O_P(1/(nh)) = O(T/h).$$

It follows that $D_{nT5} = O_P(T^{1/2}b^{1/2} + T^{1/2}h^{-1/2})O_P((nT)^{-1/2}) = O_P(n^{-1/2}(b^{1/2} + h^{-1/2})) = o_P(1)$.

For D_{nT6}, we write

$$\begin{aligned} D_{nT6} =& \ 2\sqrt{\frac{b}{n}}\sum_{i=1}^{n}\sigma_i^{-2}\varepsilon_i'(\bar{H}-L)(\bar{\mathbf{f}}-S_T\overline{\mathbf{F}}) - 2\sqrt{\frac{b}{n}}\sum_{i=1}^{n}\sigma_i^{-2}(S_T\varepsilon)'(\bar{H}-L)(\bar{\mathbf{f}}-S_T\overline{\mathbf{F}}) \\ \equiv& \ 2D_{nT61} - 2D_{nT62}, \end{aligned}$$

where $\overline{\mathbf{F}} \equiv i_n \otimes \bar{\mathbf{f}} = i_n \otimes \mathbf{f}$ under H_0. Noting that $D_{nT61} = n^{-1/2}b^{1/2}\sum_{i=1}^n\sum_{t=1}^T\sum_{s=1}^T\sigma_i^{-2} \times \varepsilon_{it}a_{ts}[\overline{f}(s/T) - e_1'S(s/T)\overline{\mathbf{F}}]$, by Assumptions 3.2 and 3.5 and Lemma B.3(b), we have

$$\begin{aligned} &E\left(D_{nT61}^2\right) \\ &= \frac{b}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\omega_{ij}a_{ts}a_{tr}[\overline{f}(s/T)-e_1'S(s/T)\overline{\mathbf{F}}][\overline{f}(r/T)-e_1'S(r/T)\overline{\mathbf{F}}]\sigma_i^{-2}\sigma_j^{-2} \\ &\leq \underline{c}^{-2}Tb\max_{1\leq s\leq T}\left|\overline{f}\left(\frac{s}{T}\right)-e_1'S\left(\frac{s}{T}\right)\overline{\mathbf{F}}\right|^2\left(\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}|\omega_{ij}|\right)\left(\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}|a_{ts}a_{tr}|\right) \\ &= TbO(h^{2p+2})O(1)O(1) = O(Tbh^{2p+2}) = o(1). \end{aligned}$$

It follows that $D_{nT61} = o_P(1)$ by the Chebyshev inequality. For D_{nT62}, we can follow the proof of D_{nT52} and show that $D_{nT62} = o_P(1)$. Consequently, $D_{nT6} = o_P(1)$. Now write $D_{nT7} \equiv -2\sqrt{b/n}\sum_{i=1}^n \sigma_i^{-2}(\widehat{\beta}-\beta)'X_i^{*\prime}\bar{H}\bar{\mathbf{f}}^* + 2(b/n)^{1/2}\sum_{i=1}^n\sigma_i^{-2}(\widehat{\beta}-\beta)'X_i^{*\prime}L\bar{\mathbf{f}}^* \equiv -2D_{nT71} + 2D_{nT72}$. By the Cauchy-Schwarz inequality, we have

$$\begin{aligned} D_{nT71} \leq& \left(\sqrt{\frac{b}{n}}\|\widehat{\beta}-\beta\|^2\sum_{i=1}^{n}\sigma_i^{-4}\|X_i^{*\prime}\bar{H}X_i^*\|\right)^{1/2}(\sqrt{nb}\bar{\mathbf{f}}^{*\prime}\bar{H}\bar{\mathbf{f}}^*)^{1/2} \\ =& \ [O_P(n^{-1/2})O(Tn^{1/2}h^{2(p+1)})]^{1/2} \\ =& \ O_P(T^{1/2}h^{p+1}) = o_P(1). \end{aligned}$$

Similarly, we have $D_{nT72} = o_P(1)$. Thus $D_{nT7} = o_P(1)$. □

LEMMA A.1. $\mathcal{D}_{nT11} = \frac{b^{1/2}}{\sqrt{n}}\sum_{i=1}^n\varepsilon_i'(\bar{H}-L-Q)\varepsilon_i/\sigma_i^2 \xrightarrow{d} N(0,\Omega_0)$.

Proof: Write $\mathcal{D}_{nT11} = \frac{1}{\sqrt{T}}\sum_{t=2}^T Z_{nT,t}$, where $Z_{nT,t} \equiv \frac{2b^{1/2}}{\sqrt{nT}}\sum_{s=1}^{t-1}\sum_{i=1}^n\alpha_{ts}\sigma_i^{-2}\varepsilon_{it}\varepsilon_{is}$ and $\alpha_{ts} \equiv T\bar{H}_{ts} - 1 = Ta_{ts}$. Noting that $\{Z_{nT,t},\ \mathcal{F}_{n,t}(\varepsilon)\}$ is an m.d.s., we prove the lemma by applying the martingale CLT. By Corollary 5.26 of White (2001) it suffices to show that: (a) $E(Z_{nT,t}^4) < C$ for all t and (n,T) for some $C < \infty$, and (b) $T^{-1}\sum_{t=2}^T Z_{nT,t}^2 - \Omega_0 = o_P(1)$.

We first prove (a). For $2 \leq t \leq T$, decompose

$$
\begin{aligned}
Z_{nT,t}^2 &= \frac{4b}{nT}\sum_{s_1=1}^{t-1}\sum_{s_2=1}^{t-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\alpha_{ts_1}\alpha_{ts_2}\sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\varepsilon_{i_1t}\varepsilon_{i_1s_1}\varepsilon_{i_2t}\varepsilon_{i_2s_2} \\
&= \frac{4b}{nT}\sum_{s=1}^{t-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\alpha_{ts}^2\sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\varepsilon_{i_1t}\varepsilon_{i_1s}\varepsilon_{i_2t}\varepsilon_{i_2s} \\
&\quad + \frac{4b}{nT}\sum_{1\leq s_1<s_2\leq t-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\alpha_{ts_1}\alpha_{ts_2}\sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\varepsilon_{i_1t}\varepsilon_{i_1s_1}\varepsilon_{i_2t}\varepsilon_{i_2s_2} \\
&\quad + \frac{4b}{nT}\sum_{1\leq s_2<s_1\leq t-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\alpha_{ts_1}\alpha_{ts_2}\sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\varepsilon_{i_1t}\varepsilon_{i_1s_1}\varepsilon_{i_2t}\varepsilon_{i_2s_2} \\
&\equiv z_{1t} + z_{2t} + z_{3t}, \text{ say.}
\end{aligned} \tag{A.7}
$$

Then $E(Z_{nT,t}^4) = E(z_{1t}+z_{2t}+z_{3t})^2 \leq 3\{E(z_{1t}^2)+E(z_{2t}^2)+E(z_{3t}^2)\} \equiv 3\{\mathcal{Z}_{1t}+\mathcal{Z}_{2t}+\mathcal{Z}_{3t}\}$, say.

$$
\begin{aligned}
\mathcal{Z}_{1t} &= \frac{16b^2}{n^2T^2}\sum_{s_1=1}^{t-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{s_2=1}^{t-1}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n}\Big\{\alpha_{ts_1}^2\alpha_{ts_2}^2\sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\sigma_{i_3}^{-2}\sigma_{i_4}^{-2} \\
&\quad \times E(\varepsilon_{i_1t}\varepsilon_{i_2t}\varepsilon_{i_3t}\varepsilon_{i_4t}\varepsilon_{i_1s_1}\varepsilon_{i_2s_1}\varepsilon_{i_3s_2}\varepsilon_{i_4s_2})\Big\} \\
&= \frac{16b^2}{n^2T^2}\sum_{s_1=1}^{t-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{s_2=1}^{t-1}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n}\Big\{\alpha_{ts_1}^2\alpha_{ts_2}^2\sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\sigma_{i_3}^{-2}\sigma_{i_4}^{-2} \\
&\quad \times \kappa_{i_1i_2i_3i_4}E(\varepsilon_{i_1s_1}\varepsilon_{i_2s_1}\varepsilon_{i_3s_2}\varepsilon_{i_4s_2})\Big\} \\
&= \frac{16b^2}{n^2T^2}\sum_{s=1}^{t-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n}\alpha_{ts}^4\sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\sigma_{i_3}^{-2}\sigma_{i_4}^{-2}\kappa_{i_1i_2i_3i_4}^2 \\
&\quad + \frac{16b^2}{n^2T^2}\sum_{1\leq s_1\neq s_2\leq t-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n}\alpha_{ts_1}^2\alpha_{ts_2}^2\sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\sigma_{i_3}^{-2}\sigma_{i_4}^{-2}\kappa_{i_1i_2i_3i_4}\omega_{i_1i_2}\omega_{i_3i_4} \\
&\quad \leq \frac{Cb^2}{T^2}\sum_{s=1}^{t-1}\alpha_{ts}^4 + C\left(\frac{b}{T}\sum_{s=1}^{t-1}\alpha_{ts}^2\right)^2 \leq \frac{C}{Tb} + C \leq 2C.
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
\mathcal{Z}_{2t} &= \frac{16b^2}{n^2T^2}\sum_{1\leq s_1<s_2\leq t-1}\sum_{1\leq s_3<s_4\leq t-1}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n}\alpha_{ts_1}\alpha_{ts_2}\alpha_{ts_3}\alpha_{ts_4}\sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\sigma_{i_3}^{-2}\sigma_{i_4}^{-2} \\
&\quad \times E(\varepsilon_{i_1t}\varepsilon_{i_2t}\varepsilon_{i_3t}\varepsilon_{i_4t}\varepsilon_{i_1s_1}\varepsilon_{i_2s_2}\varepsilon_{i_3s_3}\varepsilon_{i_4s_4})
\end{aligned}
$$

$$= \frac{16b^2}{n^2T^2} \sum_{1 \le s_1 < s_2 \le t-1} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} \alpha_{ts_1}^2 \alpha_{ts_2}^2 \sigma_{i_1}^{-2} \sigma_{i_2}^{-2} \sigma_{i_3}^{-2} \sigma_{i_4}^{-2} \kappa_{i_1 i_2 i_3 i_4} \omega_{i_1 i_2} \omega_{i_3 i_4}$$
$$\le \frac{Cb^2}{T^2} \sum_{1 \le s_1 < s_2 \le t-1} \alpha_{ts_1}^2 \alpha_{ts_2}^2 \le C,$$

where we have used the fact that $T^{-1} b \sum_{s=1}^{t} \alpha_{ts}^2 \le C$ uniformly in t and C may vary across lines. By the same token $\mathcal{Z}_{3t} \le C$ for all t. Consequently, $E(Z_{nT,t}^4) < C$ for all t and some large enough constant C.

Now we prove (b) by the Chebyshev inequality. First, by Assumption 3.2(b)–(c),

$$E\left(\frac{1}{T}\sum_{t=2}^{T} Z_{nT,t}^2\right) = \frac{4b}{nT^2} \sum_{t=2}^{T} \sum_{s=1}^{t-1} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ts}^2 \sigma_i^{-2} \sigma_j^{-2} \omega_{ij}^2 = \frac{2b}{nT^2} \sum_{1 \le t \ne s \le T} \alpha_{ts}^2 \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij}^2,$$

where $\rho_{ij} = \omega_{ij}/(\sigma_i \sigma_j)$ by Assumption 3.2. Second, decompose

$$E\left[\left(\frac{1}{T}\sum_{t=2}^{T} Z_{nT,t}^2\right)^2\right] = \frac{1}{T^2}\sum_{t=2}^{T} E\left(Z_{nT,t}^4\right) + \frac{2}{T^2} \sum_{2 \le t < s \le T} E\left(Z_{nT,t}^2 Z_{nT,s}^2\right) \equiv \mathbb{Z}_{1nT} + \mathbb{Z}_{2nT}.$$

By the proof of (a), $\mathbb{Z}_{1nT} = T^{-2} \sum_{t=2}^{T} E(Z_{nT,t}^4) = O(1/T) = o(1)$. For $\mathbb{Z}_{2nT}$, by (A.7) we have $\mathbb{Z}_{2nT} = 2T^{-2} \sum_{2 \le t < s \le T} E(z_{1t}z_{1s} + z_{1t}z_{2s} + z_{1t}z_{3s} + z_{2t}z_{1s} + z_{2t}z_{2s} + z_{2t}z_{3s} + z_{3t}z_{1s} + z_{3t}z_{2s} + z_{3t}z_{3s}) \equiv \sum_{j=1}^{9} \mathbb{Z}_{2nTj}$, where, e.g. $\mathbb{Z}_{2nT1} = 2T^{-2} \sum_{2 \le t < s \le T} E(z_{1t}z_{1s})$. For $\mathbb{Z}_{2nT1}$, we have

$$\mathbb{Z}_{2nT1} = \frac{32b^2}{n^2T^4} \sum_{2 \le t_1 < t_2 \le T} \sum_{s_1=1}^{t_1-1} \sum_{s_2=1}^{t_2-1} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} \alpha_{t_1 s_1}^2 \alpha_{t_2 s_2}^2 \sigma_{i_1}^{-2} \sigma_{i_2}^{-2} \sigma_{i_3}^{-2} \sigma_{i_4}^{-2}$$
$$\times \omega_{i_3 i_4} E(\varepsilon_{i_1 t_1} \varepsilon_{i_2 t_1} \varepsilon_{i_1 s_1} \varepsilon_{i_2 s_1} \varepsilon_{i_3 s_2} \varepsilon_{i_4 s_2})$$
$$= \frac{32b^2}{n^2T^4} \sum_{2 \le t_1 < t_2 \le T} \sum_{s_1=1}^{t_1-1} \sum_{s_2=1}^{t_2-1} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} \alpha_{t_1 s_1}^2 \alpha_{t_2 s_2}^2 \sigma_{i_1}^{-2} \sigma_{i_2}^{-2} \sigma_{i_3}^{-2} \sigma_{i_4}^{-2} \omega_{i_1 i_2}^2 \omega_{i_3 i_4}^2$$
$$+ O(1/T)$$
$$= \frac{16b^2}{n^2T^4} \sum_{t_1=1}^{T} \sum_{t_2=1}^{T} \sum_{s_1=1}^{t_1-1} \sum_{s_2=1}^{t_2-1} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \sum_{i_3=1}^{n} \sum_{i_4=1}^{n} \alpha_{t_1 s_1}^2 \alpha_{t_2 s_2}^2 \rho_{i_1 i_2}^2 \rho_{i_3 i_4}^2 + O(1/T)$$
$$= \left(\frac{2b}{nT^2} \sum_{1 \le t \ne s \le T} \alpha_{ts}^2 \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij}^2\right)^2 + O(1/T).$$

Similarly, by Assumption 3.2 and Lemmata B.2 and B.3(b)

$$
\begin{aligned}
\mathbb{Z}_{2nT2} = & \frac{32b^2}{n^2T^4} \sum_{2\le t_1<t_2\le T} \sum_{s=1}^{t_1-1} \sum_{1\le s_1<s_2\le t_2-1} \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n} \alpha_{t_1 s}^2 \alpha_{t_2 s_1}\alpha_{t_2 s_2} \sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\sigma_{i_3}^{-2}\sigma_{i_4}^{-2} \\
& \times \varsigma_{i_2 i_3 i_4} E(\varepsilon_{i_1 t_1}\varepsilon_{i_1 s}\varepsilon_{i_2 s}\varepsilon_{i_3 s_1}\varepsilon_{i_4 s_2}) \\
= & \frac{32b^2}{n^2T^4} \sum_{2\le t_1<t_2\le T} \sum_{s=1}^{t_1-1} \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n} \alpha_{t_1 s}^2 \alpha_{t_2 s}\alpha_{t_2 t_1} \sigma_{i_1}^{-2}\sigma_{i_2}^{-2}\sigma_{i_3}^{-2}\sigma_{i_4}^{-2} \varsigma_{i_2 i_3 i_4}\omega_{i_1 i_4}\varsigma_{i_1 i_2 i_3} \\
& \le C\left(b^2 \max_{1\le t\neq s\le T} a_{ts}^2\right)\left(\sum_{2\le t_1<t_2\le T}\sum_{s=1}^{t_1-1} |a_{t_2 s}a_{t_2 t_1}|\right) \\
& \times \left(\frac{1}{n^2}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n}\sum_{i_3=1}^{n}\sum_{i_4=1}^{n} |\varsigma_{i_2 i_3 i_4}\varsigma_{i_1 i_2 i_3}|\right) \\
= & O(T^{-2})O(T)O(1) = o(1),
\end{aligned}
$$

where recall $\varsigma_{ijk} \equiv E(\varepsilon_{it}\varepsilon_{jt}\varepsilon_{kt})$. Analogously we can show that $\mathbb{Z}_{2nTl} = o(1)$ for $l = 3, 4, \dots, 9$. It follows that

$$
E\left[\left(\frac{1}{T}\sum_{t=2}^{T} Z_{nT,t}^2\right)^2\right] = \left(\frac{2b}{nT^2}\sum_{1\le t\neq s\le T} \alpha_{ts}^2 \sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}^2\right)^2 + o(1),
$$

and

$$
\mathrm{Var}\left(\frac{1}{T}\sum_{t=2}^{T} Z_{nT,t}^2\right) = E\left[\left(\frac{1}{T}\sum_{t=2}^{T} Z_{nT,t}^2\right)^2\right] - \left[E\left(\frac{1}{T}\sum_{t=2}^{T} Z_{nT,t}^2\right)\right]^2 = o(1).
$$

Consequently, $\frac{1}{T}\sum_{t=2}^{T} Z_{nT,t}^2 - \frac{2b}{nT^2}\sum_{1\le t\neq s\le T}\alpha_{ts}^2 \sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}^2 = o_P(1)$ and (b) follows by the definition of Ω_0. □

PROPOSITION A.2. $\Gamma_{nT,2} = o_P(1)$.

Proof: Let $\widehat{\sigma}_i^2 \equiv TSS_i/T$. By a geometric expansion, $1/\widehat{\sigma}_i^2 - 1/\sigma_i^2 = -(\widehat{\sigma}_i^2 - \sigma_i^2)/\sigma_i^4 + (\widehat{\sigma}_i^2 - \sigma_i^2)^2/(\sigma_i^4\widehat{\sigma}_i^2)$. It follows that

$$
\begin{aligned}
\Gamma_{nT,2} = & -\sqrt{\frac{b}{n}}\sum_{i=1}^{n}(ESS_i - \varepsilon_i' Q\varepsilon_i)\frac{\widehat{\sigma}_i^2 - \sigma_i^2}{\sigma_i^4} + \sqrt{\frac{b}{n}}\sum_{i=1}^{n}(ESS_i - \varepsilon_i' Q\varepsilon_i)\frac{\left(\widehat{\sigma}_i^2 - \sigma_i^2\right)^2}{\sigma_i^4\widehat{\sigma}_i^2} \\
\equiv & -\Gamma_{nT,21} + \Gamma_{nT,22}, \text{ say.}
\end{aligned}
$$

Noting that $\widehat{u}_i = \varepsilon_i^* - X_i^*(\widehat{\beta} - \beta) + \bar{\mathbf{f}}^* + (\mathbf{f}_i - \bar{\mathbf{f}}) + \alpha_i i_T$ and $M i_T = 0$ where $\bar{\mathbf{f}}$ and $\bar{\mathbf{f}}^*$ are defined in (A.2), we have

$$\widehat{\sigma}_i^2 = TSS_i/T = \widehat{u}_i' M \widehat{u}_i/T = \sum_{l=1}^{10} TSS_{il}/T, \tag{A.8}$$

where

$$\begin{aligned}
TSS_{i1} &\equiv \varepsilon_i^{*\prime} M \varepsilon_i^*, & TSS_{i2} &\equiv (\widehat{\beta} - \beta)' X_i^{*\prime} M X_i^*(\widehat{\beta} - \beta), \\
TSS_{i3} &\equiv \bar{\mathbf{f}}^{*\prime} M \bar{\mathbf{f}}^*, & TSS_{i4} &\equiv -2\varepsilon_i^{*\prime} M X_i^*(\widehat{\beta} - \beta), \\
TSS_{i5} &\equiv 2\varepsilon_i^{*\prime} M \bar{\mathbf{f}}^*, & TSS_{i6} &\equiv -2\bar{\mathbf{f}}^{*\prime} M X_i^*(\widehat{\beta} - \beta), \\
TSS_{i7} &\equiv 2\varepsilon_i^{*\prime} M(\mathbf{f}_i - \bar{\mathbf{f}}), & TSS_{i8} &\equiv (\mathbf{f}_i - \bar{\mathbf{f}})' M(\mathbf{f}_i - \bar{\mathbf{f}}), \\
TSS_{i9} &\equiv 2\bar{\mathbf{f}}^{*\prime} M(\mathbf{f}_i - \bar{\mathbf{f}}), & TSS_{i10} &\equiv -2(\widehat{\beta} - \beta)' X_i^{*\prime} M(\mathbf{f}_i - \bar{\mathbf{f}}).
\end{aligned}$$

Under H_0, we have $\mathbf{f}_i - \bar{\mathbf{f}} = 0$. Thus $TSS_{il} = 0$ for $l = 7, \ldots, 10$. We want to show that

$$\max_{1 \le i \le n} \left| T^{-1} TSS_{i1} - \sigma_i^2 \right| = O_P(\upsilon_{nT}), \text{ and } \max_{1 \le i \le n} T^{-1} TSS_{il} = o_P(\upsilon_{nT}) \text{ for } l = 2, \ldots, 6, \tag{A.9}$$

where $\upsilon_{nT} \equiv n^{1/\lambda} T^{-1/2}$.

For TSS_{i1}, we have

$$T^{-1} TSS_{i1} - \sigma_i^2 = \left(T^{-1} \varepsilon_i' M \varepsilon_i - \sigma_i^2 \right) - 2T^{-1} \varepsilon_i' M S_T \varepsilon + T^{-1} (S_T \varepsilon)' M S_T \varepsilon. \tag{A.10}$$

We first bound the last term in (A.10). By the idempotence of M and the Markov inequality, $T^{-1}(S_T\varepsilon)' \; MS_T\varepsilon \le T^{-1}\|S_T\varepsilon\|^2 = O_P(n^{-1}T^{-1}h^{-1})$. For the first term in (A.10), we want to show that $\max_{1\le i\le n} |\varepsilon_i' M \varepsilon_i/T - \sigma_i^2| = O_P(\upsilon_{nT})$. Write $\varepsilon_i' M \varepsilon_i/T = T^{-1}\sum_{t=1}^T (\varepsilon_{it} - \bar{\varepsilon}_i)^2 = T^{-1}\sum_{t=1}^T \varepsilon_{it}^2 - \bar{\varepsilon}_i^2$. Let $\xi_{it} \equiv \varepsilon_{it}^2 - \sigma_i^2$. Then by Assumption 3.2(d) and the Chebyshev inequality, for any $\epsilon > 0$

$$P\left(\max_{1\le i\le n} \frac{1}{T} \sum_{t=1}^T \xi_{it} \ge \epsilon \upsilon_{nT} \right) \le \epsilon^{-\lambda} \upsilon_{nT}^{-\lambda} \sum_{i=1}^n E\left(\frac{1}{T} \sum_{t=1}^T \xi_{it} \right)^\lambda = O\left(nT^{-\lambda/2} \upsilon_{nT}^{-\lambda} \right) = O(1).$$

It follows that $\max_{1\le i\le n} |T^{-1}\sum_{t=1}^T \varepsilon_{it}^2 - \sigma_i^2| = O_P(\upsilon_{nT})$. Similarly, $\max_{1\le i\le n} |\bar{\varepsilon}_i| = O_P(\upsilon_{nT}^2) = o_P(\upsilon_{nT})$. It follows that $\varepsilon_i' M \varepsilon_i/T = \sigma_i^2 + O_P(\upsilon_{nT})$ uniformly in i. Then by the Cauchy–Schwarz inequality, we can readily show that the second term in (A.10) is $O_P(n^{-1/2}T^{-1/2}h^{-1/2}) = o_P(\upsilon_{nT})$. Consequently, the first result in (A.9) follows and $\max_{1\le i\le n} T^{-1} TSS_{i1} = O_P(1)$.

For TSS_{i2}, we have

$$\begin{aligned}
\max_{1\le i\le n} \left\{ T^{-1} TSS_{i2} \right\} &\le C\|\widehat{\beta} - \beta\|^2 \max_{1\le i\le n} \{T^{-1}\|X_i - S_T X\|^2\} \\
&= O_P(n^{-1}T^{-1}) O_P(\sqrt{n/T} + 1),
\end{aligned}$$

where we use the fact that $\max_{1\le i\le n} T^{-1}\|X_i - S_T X\|^2 = O_P(\sqrt{n/T} + 1)$ under our moment conditions. For TSS_{i3}, noting that $\|\bar{\mathbf{f}}^*\| = \|\bar{\mathbf{f}} - S_T \bar{\mathbf{F}}\| = O(T^{1/2}h^{p+1})$, we have

$T^{-1}TSS_{i3} \leq T^{-1}\|\bar{\mathbf{f}} - S_T\overline{\mathbf{F}}\|^2 = O(h^{2p+2})$. By the Cauchy-Schwarz inequality, we have

$$\max_{1\leq i\leq n} T^{-1}|TSS_{i4}| \leq \max_{1\leq i\leq n} \left(T^{-1}TSS_{i1}\right)^{1/2}\left(T^{-1}TSS_{i2}\right)^{1/2}$$
$$= O_P(n^{-1/4}T^{-3/4}+n^{-1/2}T^{-1/2}) = o_P(\upsilon_{nT}),$$
$$\max_{1\leq i\leq n} T^{-1}|TSS_{i5}| \leq \max_{1\leq i\leq n} \left(T^{-1}TSS_{i1}\right)^{1/2}\left(T^{-1}TSS_{i3}\right)^{1/2} = O_P(h^{p+1}) = o_P(\upsilon_{nT}),$$
$$\text{and} \max_{1\leq i\leq n} T^{-1}|TSS_{i6}| \leq \max_{1\leq i\leq n} \left(T^{-1}TSS_{i2}\right)^{1/2}\left(T^{-1}TSS_{i3}\right)^{1/2} = o_P(\upsilon_{nT}).$$

Consequently, we have $\max_{1\leq i\leq n}|\widehat{\sigma}_i^2 - \sigma_i^2| = O_P(\upsilon_{nT})$. Then by Assumption 3.5

$$\begin{aligned}
\Gamma_{nT,22} &\leq \frac{\max_{1\leq i\leq n}|\widehat{\sigma}_i^2-\sigma_i^2|^2}{\min_{1\leq i\leq n}\sigma_i^4\widehat{\sigma}_i^2}\frac{b^{1/2}}{\sqrt{n}}\sum_{i=1}^n |ESS_i - \varepsilon_i' Q\varepsilon_i| \\
&\leq \frac{\sqrt{n}\max_{1\leq i\leq n}|\widehat{\sigma}_i^2-\sigma_i^2|^2}{\min_{1\leq i\leq n}\sigma_i^4\widehat{\sigma}_i^2}\left(\frac{b}{n}\sum_{i=1}^n (ESS_i - \varepsilon_i' Q\varepsilon_i)^2\right)^{1/2} \\
&= \sqrt{n}O_P\left(\upsilon_{nT}^2\right)O_P(1) = O_P(n^{1/2+2/\lambda}T^{-1}) = o(1),
\end{aligned}$$

because one can easily show that $\frac{b}{n}\sum_{i=1}^n(ESS_i - \varepsilon_i' Q\varepsilon_i)^2 = O_P(1)$. For $\Gamma_{nT,21}$, we have $\Gamma_{nT,21} = \sum_{l=1}^6 \Gamma_{nT,21l}$, where

$$\Gamma_{nT,211} \equiv \sqrt{\frac{b}{n}}\sum_{i=1}^n \sigma_i^{-4}(ESS_i - \varepsilon_i' Q\varepsilon_i)\left(T^{-1}TSS_{i1} - \sigma_i^2\right), \text{ and}$$
$$\Gamma_{nT,21l} \equiv \sqrt{\frac{b}{n}}\sum_{i=1}^n \sigma_i^{-4}(ESS_i - \varepsilon_i' Q\varepsilon_i)\left(T^{-1}TSS_{il}\right) \text{ for } l = 2,\ldots,6.$$

Following the proof of Proposition A.1 and the above analysis for TSS_{il}, we can show that $\Gamma_{nT,21l} = o_P(1)$ for $l = 1,\ldots,6$. □

Proof of Corollary 3.1: Given Theorem 3.1, it suffices to show that: (a) $\widehat{B}_{nT} = B_{nT} + o_P(1)$, and (b) $\widehat{\Omega}_{nT} = \Omega_0 + o_P(1)$. We first prove (a). By (A.3) and the fact that $Mi_T = 0$, we have

$$\widehat{u}_i'\bar{Q}\widehat{u}_i = \sum_{l=1}^{10} B_{nT,il}, \tag{A.11}$$

where

$$\begin{aligned}
B_{nT,i1} &\equiv \varepsilon_i^{*\prime}\bar{Q}\varepsilon_i^*, & B_{nT,i2} &\equiv (\widehat{\beta}-\beta)'X_i^{*\prime}\bar{Q}X_i^*(\widehat{\beta}-\beta), \\
B_{nT,i3} &\equiv \bar{\mathbf{f}}^{*\prime}\bar{Q}\bar{\mathbf{f}}^*, & B_{nT,i4} &\equiv -2\varepsilon_i^{*\prime}\bar{Q}X_i^*(\widehat{\beta}-\beta), \\
B_{nT,i5} &\equiv 2\varepsilon_i^{*\prime}\bar{Q}\bar{\mathbf{f}}^*, & B_{nT,i6} &\equiv -2\bar{\mathbf{f}}^{*\prime}\bar{Q}X_i^*(\widehat{\beta}-\beta),
\end{aligned}$$

$$B_{nT,i7} \equiv 2\bar{\mathbf{f}}^{*\prime}\bar{Q}(\mathbf{f}_i - \bar{\mathbf{f}}) \qquad B_{nT,i8} \equiv -2(\widehat{\beta} - \beta)' X_i^{*\prime}\bar{Q}(\mathbf{f}_i - \bar{\mathbf{f}}),$$

$$B_{nT,i9} \equiv 2\varepsilon_i^{*\prime}\bar{Q}(\mathbf{f}_i - \bar{\mathbf{f}}), \quad B_{nT,i10} \equiv (\mathbf{f}_i - \bar{\mathbf{f}})'\bar{Q}(\mathbf{f}_i - \bar{\mathbf{f}}),$$

$\bar{Q} \equiv MQM$, and $\bar{\mathbf{f}}$ and $\bar{\mathbf{f}}^*$ are defined in (A.2). Under H_0, we have $\mathbf{f}_i - \bar{\mathbf{f}} = 0$. Thus $B_{nT,il} = 0$ for $l = 7, \ldots, 10$. By (3.2) and (A.11), it suffices to show that

$$\mathcal{B}_{nT,1} \equiv \sqrt{\frac{b}{n}}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}(B_{nT,i1} - B_{nT}) = \sqrt{\frac{b}{n}}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}[\varepsilon_i^{*\prime}\bar{Q}\varepsilon_i^* - \varepsilon_i' Q\varepsilon_i] = o_P(1)$$

$$B_{nT,l} \equiv n^{-1/2}b^{1/2}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}B_{nT,il} = o_P(1) \text{ for } l = 2, \ldots, 6.$$

Recalling $\varepsilon_i^* \equiv \varepsilon_i - S_T\varepsilon$, we decompose $\mathcal{B}_{nT,1}$ as follows

$$\begin{aligned}
\mathcal{B}_{nT,1} &= n^{-1/2}b^{1/2}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}[(\varepsilon_i - S_T\varepsilon)'\bar{Q}(\varepsilon_i - S_T\varepsilon) - \varepsilon_i' Q\varepsilon_i] \\
&= n^{-1/2}b^{1/2}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}[\varepsilon_i'\bar{Q}\varepsilon_i - \varepsilon_i' Q\varepsilon_i] - 2n^{-1/2}b^{1/2}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}\varepsilon_i'\bar{Q}S_T\varepsilon \\
&\quad + n^{-1/2}b^{1/2}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}(S_T\varepsilon)'\bar{Q}S_T\varepsilon \\
&\equiv \mathcal{B}_{nT,11} - 2\mathcal{B}_{nT,12} + \mathcal{B}_{nT,13}.
\end{aligned}$$

Noting that $\bar{Q} - Q = (I_T - L)Q(I_T - L) - Q = LQL - QL - LQ$ and both Q and L are symmetric, we have

$$\begin{aligned}
\mathcal{B}_{nT,11} &= n^{-1/2}b^{1/2}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}\varepsilon_i' LQL\varepsilon_i - 2n^{-1/2}b^{1/2}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}\varepsilon_i' QL\varepsilon_i \\
&\equiv \mathcal{B}_{nT,11a} - 2\mathcal{B}_{nT,11b}.
\end{aligned}$$

Following the proof of Proposition A.2, we can show that $\mathcal{B}_{nT,11a} = B_{nT,11a} + o_P(1)$, where $B_{nT,11a} = n^{-1/2}b^{1/2}\sum_{i=1}^{n}\sigma_i^{-2}\varepsilon_i' LQL\varepsilon_i$. Even though Q is not positive semi-definite (p.s.d.), it can be written as the difference between two p.s.d. matrices: $Q = Q^* - T^{-1}I_T$, where $Q^* = \text{diag}(\bar{H}_{11}, \ldots, \bar{H}_{TT})$. So we can write $B_{nT,11a} = n^{-1/2}b^{1/2}\sum_{i=1}^{n}\sigma_i^{-2}\varepsilon_i' LQ^*L\varepsilon_i - n^{-1/2}T^{-1}b^{1/2}\sum_{i=1}^{n}\sigma_i^{-2}\varepsilon_i' LL\varepsilon_i = B_{nT,11a1} - B_{nT,11a2}$. Noting that

$$\begin{aligned}
E|B_{nT,11a1}| &= n^{-1/2}b^{1/2}\sum_{i=1}^{n}\sigma_i^{-2}E(\varepsilon_i' LQ^*L\varepsilon_i) = T^{-2}n^{-1/2}b^{1/2}\sum_{i=1}^{n}\sum_{t=1}^{T}i_T' Q^* i_T \\
&= O(T^{-1}n^{1/2}b^{1/2})\text{tr}(Q^*) = O(T^{-1}n^{1/2}b^{1/2})O(b^{-1}) = o(1),
\end{aligned}$$

and similarly $E|B_{nT,11a2}| = O(T^{-1}n^{1/2}b^{1/2}) = o(1)$, we have $\mathcal{B}_{nT,11a} = o_P(1)$ by the Markov inequality. Similarly, $\mathcal{B}_{nT,11b} = o_P(1)$. Consequently $\mathcal{B}_{nT,11} = o_P(1)$. Analogously, we can show that $\mathcal{B}_{nT,1l} = o_P(1)$ for $l = 2, 3$. It follows that $\mathcal{B}_{nT,1} = o_P(1)$.

Using the fact that $|\mathrm{tr}(AB)| \leq \lambda_{\max}(A)$ tr (B) for any conformable p.s.d. matrix B and symmetric matrix A (see, e.g. Bernstein, 2005, p. 309) and that $\lambda_{\max}(M) = 1$, we can show that $\|X_i^{*\prime}\bar{Q}X_i^*\|^2 = \mathrm{tr}(MQMX_i^*X_i^{*\prime}MQMX_i^*X_i^{*\prime}) \leq \|X_i^{*\prime}QX_i^*\|^2$. It follows that

$$\begin{aligned} B_{nT,2} &= n^{-1/2}b^{1/2}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}(\widehat{\beta}-\beta)X_i^{*\prime}\bar{Q}X_i^*(\widehat{\beta}-\beta) \\ &\leq n^{-1/2}b^{1/2}\|\widehat{\beta}-\beta\|^2\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}\|X_i^{*\prime}QX_i^*\| \\ &= n^{-1/2}b^{1/2}O_P((nT)^{-1})O_P(nb^{-1}) = O_P(n^{-1/2}T^{-1}b^{-1/2}) = o_P(1) \end{aligned}$$

where we use the fact that $\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}\|X_i^{*\prime}QX_i^*\| = O_P(nb^{-1})$. Similarly, we have

$$\begin{aligned} B_{nT,3} &= n^{-1/2}b^{1/2}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}\overline{\mathbf{f}}^{*\prime}\bar{Q}\overline{\mathbf{f}}^* \leq n^{-1/2}b^{1/2}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}\|\overline{\mathbf{f}}^{*\prime}Q\overline{\mathbf{f}}^*\| \\ &= n^{-1/2}b^{1/2}\left\|\sum_{t=1}^{T}(\bar{H}_{tt}-T^{-1})[\overline{f}(t/T)-e_1'S(t/T)\mathbf{F}]^2\right\|^2\sum_{i=1}^{n}\widehat{\sigma}_i^{-2} \\ &= n^{-1/2}b^{1/2}O_P(b^{-1}h^{2p+2})O_P(n) = O_P(n^{1/2}h^{2p+2}b^{-1/2}) = o_P(1). \end{aligned}$$

By the repeated use of the Cauchy-Schwarz inequality, we can show that $B_{nT,il} = o_P(1)$ for $l = 4, 5$, and 6.

To show (b), it suffices to show that $DV_{nT} \equiv n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}(\widehat{\rho}_{ij}^2 - \rho_{ij}^2) = o_P(1)$. Noting that $x^2 - y^2 = (x-y)^2 + 2(x-y)y$, we can decompose DV_{nT} as follows

$$DV_{nT} = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(\widehat{\rho}_{ij}-\rho_{ij})^2 + \frac{2}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}(\widehat{\rho}_{ij}-\rho_{ij})\rho_{ij} \equiv DV_{nT1} + 2DV_{nT2}.$$

Following the argument in the proof of Proposition A.2, we can show that

$$\begin{aligned} DV_{nT1} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\widehat{u}_i'M\widehat{u}_j}{\widehat{\sigma}_i\widehat{\sigma}_j} - \frac{\omega_{ij}}{\sigma_i\sigma_j}\right)^2 = \overline{DV}_{nT1} + o_P(1), \text{ and} \\ DV_{nT2} &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{\widehat{u}_i'M\widehat{u}_j}{\widehat{\sigma}_i\widehat{\sigma}_j} - \frac{\omega_{ij}}{\sigma_i\sigma_j}\right)\rho_{ij} = \overline{DV}_{nT2} + o_P(1). \end{aligned}$$

where $\overline{DV}_{nT1} \equiv n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i^{-2}\sigma_j^{-2}(\widehat{u}_i'M\widehat{u}_j - \omega_{ij})^2$ and $\overline{DV}_{nT2} \equiv n^{-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\rho_{ij}\sigma_i^{-1}\sigma_j^{-1}(\widehat{u}_i'M\widehat{u}_j - \omega_{ij})$.

By (A.3) and the fact that $Mi_T = 0$, we have that under H_0, $\widehat{u}_i'M\widehat{u}_j = \varepsilon_i^{*\prime}M\varepsilon_j^* + (\widehat{\beta}-\beta)'X_i^{*\prime}MX_j^*(\widehat{\beta}-\beta) + \overline{\mathbf{f}}^{*\prime}M\overline{\mathbf{f}}^* - (\varepsilon_i^{*\prime}MX_j^* + \varepsilon_j^{*\prime}MX_i^*)(\widehat{\beta}-\beta) + (\varepsilon_i^* + \varepsilon_j^*)'M\overline{\mathbf{f}}^* - \overline{\mathbf{f}}^{*\prime}M(X_i^* + X_j^*)$

$(\widehat{\beta}-\beta) \equiv \sum_{l=1}^{6} DV_{nT,ijl}$. We can prove that $\overline{DV}_{nT1} = o_P(1)$ by showing that

$$\overline{DV}_{nT1,1} \equiv \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i^{-2}\sigma_j^{-2}(DV_{nT,ij1}-\omega_{ij})^2 = o_P(1), \text{ and}$$

$$\overline{DV}_{nT1,l} \equiv \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_i^{-2}\sigma_j^{-2}(DV_{nT,ijl})^2 = o_P(1) \text{ for } l = 2,\ldots,6.$$

Similarly we can prove $\overline{DV}_{nT2} = o_P(1)$ by using the above decomposition for $\widehat{u}_i' M\widehat{u}_j$. The details are omitted for brevity. □

Proof of Theorem 3.2: By (3.2) we have

$$\begin{aligned}
\sqrt{\widehat{\Omega}_{nT}}\,\overline{\Gamma}_{nT} &= \frac{b^{1/2}}{n^{1/2}}\sum_{i=1}^{n}\widehat{\sigma}_i^{-2}(ESS_i - \widehat{u}_i'\bar{Q}\widehat{u}_i)\\
&= \sqrt{\frac{b}{n}}\sum_{i=1}^{n}\sigma_i^{-2}(ESS_i - \varepsilon_i' Q\varepsilon_i) - \sqrt{\frac{b}{n}}\sum_{i=1}^{n}(ESS_i - \varepsilon_i' Q\varepsilon_i)\left(\frac{1}{\widehat{\sigma}_i^2} - \frac{1}{\sigma_i^2}\right)\\
&\quad - \sqrt{\frac{b}{n}}\sum_{i=1}^{n}\sigma_i^{-2}(\widehat{u}_i'\bar{Q}\widehat{u}_i - \varepsilon_i' Q\varepsilon_i) + \sqrt{\frac{b}{n}}\sum_{i=1}^{n}(\widehat{u}_i'\bar{Q}\widehat{u}_i - \varepsilon_i' Q\varepsilon_i)\left(\frac{1}{\widehat{\sigma}_i^2} - \frac{1}{\sigma_i^2}\right)\\
&\equiv \Gamma_{nT,1} - \Gamma_{nT,2} - \Gamma_{nT,3} + \Gamma_{nT,4}, \text{ say,}
\end{aligned} \tag{A.12}$$

where $\Gamma_{nT,1}$ and $\Gamma_{nT,2}$ are as defined in the proof of Theorem 3.1, and $\widehat{\sigma}_i^2 \equiv TSS_i/T$. It is easy to show that $\widehat{\Omega}_{nT} = \Omega_0 + o_P(1)$ under $H_1(\gamma_{nT})$ with $\gamma_{nT} = n^{-1/4}T^{-1/2}b^{-1/4}$. It suffices to show that: (a) $\Gamma_{nT,1} \xrightarrow{d} N(\Theta_0, \Omega_0)$, (b) $\Gamma_{nT,2} = o_P(1)$, (c) $\Gamma_{nT,3} = o_P(1)$, and (d) $\Gamma_{nT,4} = o_P(1)$. We complete the proof by Propositions A.3–A.6. □

PROPOSITION A.3. *$\Gamma_{nT,1} \xrightarrow{d} N(\Theta_0, \Omega_0)$ under $H_1(\gamma_{nT})$.*

Proof: Decompose $\Gamma_{nT,1} = \Gamma_{nT,11} - \Gamma_{nT,12}$ where $\Gamma_{nT,11}$ and $\Gamma_{nT,12}$ are defined in (A.1). Using the notation defined in the proof of Proposition A.1, it suffices to show: (a) $\mathcal{D}_{nT1} \equiv D_{nT1} - \Gamma_{nT,12} \xrightarrow{d} N(0, \Omega_0)$, (b) $D_{nT2} = \Theta_0 + o_P(1)$, and (c) $D_{nTs} = o_P(1)$ for $s = 3,\ldots,10$, where $\Theta_0 = \lim_{(n,T)\to\infty}\Theta_{nT}$ and $\Theta_{nT} \equiv n^{-1/2}b^{1/2}\gamma_{nT}^2\sum_{i=1}^{n}\sigma_i^{-2}\Delta_{ni}'(\bar{H} - L)\Delta_{ni} = n^{-1}T^{-1}\sum_{i=1}^{n}\sigma_i^{-2}\Delta_{ni}'(\bar{H} - L)\Delta_{ni}$. (a) follows the proof of Proposition A.1. We are left to prove (b) and (c).

For (b), letting ω_2 and $\mathbb{S}$ be as defined in the proof of Lemma B.2, by (B.1) we have

$$
\begin{aligned}
D_{nT2} &= \gamma_{nT}^2 \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \sigma_i^{-2} \Delta_{ni}' (\bar{H} - L) \Delta_{ni} \\
&= \frac{1}{nT} \sum_{i=1}^{n} \sigma_i^{-2} \sum_{t=1}^{T} \sum_{s=1}^{T} (\bar{H}_{ts} - T^{-1}) \Delta_{ni} \left(\frac{t}{T}\right) \Delta_{ni} \left(\frac{s}{T}\right) \\
&= \frac{1}{nT^2} \sum_{i=1}^{n} \sigma_i^{-2} \sum_{t=1}^{T} \sum_{s=1}^{T} \left\{ \int_0^1 w_{b,t}(\tau) z_{b,t}^{[1]}(\tau)' \mathbb{S}^{-1} z_{b,s}^{[1]}(\tau) w_{b,s}(\tau) d\tau \right. \\
&\quad \times \left. \left[\int_0^1 w_{b,t}(\tau) d\tau \int_0^1 w_{b,s}(\tau) d\tau \right]^{-1} - 1 \right\} \Delta_{ni} \left(\frac{t}{T}\right) \Delta_{ni} \left(\frac{s}{T}\right) + o(1) \\
&= \frac{1}{nT} \sum_{i=1}^{n} \sigma_i^{-2} \sum_{t=\lfloor Tb \rfloor + 1}^{\lfloor T(1-b) \rfloor - 1} \int_{-t/(Tb)}^{(T-t)/(Tb)} \left\{ \int_{-1}^{1} \left[1 + \omega_2^{-1} u (u - v)\right] w(u) w(u - v) du \right. \\
&\quad \times \left. \left[\left(\int_{-t/(Tb)}^{(T-t)/(Tb)} w(z) dz \int_{-t/(Tb)}^{(T-t)/(Tb)} w(z' - v) dz' \right)^{-1} - 1 \right] \right\} \\
&\quad \times \Delta_{ni} \left(\frac{t}{T}\right) \Delta_{ni} \left(\frac{t}{T} + vb\right) dv + o(1) = \frac{1}{n} \sum_{i=1}^{n} \sigma_i^{-2} \int_0^1 \Delta_{ni}(\tau)^2 d\tau \, C_w + o(1),
\end{aligned}
$$

where $C_w \equiv \int_{-1}^{1} \{ \int_{-1}^{1} [1 + \omega_2^{-1} u(u - v)] w(u) w(u - v) du [\int_{-1}^{1} w(z - v) dz]^{-1} - 1\} dv$. That is, $D_{nT2} = \Theta_{nT} = \Theta_0 + o(1)$. For (iii), following the proof of Proposition A.1, we can show that $D_{nTl} = o_P(1)$ under $H_1(\gamma_{nT})$ for $l = 3, \ldots, 7$. It suffices to prove (c) by showing that $D_{nTl} = o_P(1)$ under $H_1(\gamma_{nT})$ for $l = 8, \ldots, 10$. For D_{nT8}, write

$$
\begin{aligned}
D_{nT8} &\equiv 2\sqrt{\frac{b}{n}} \sum_{i=1}^{n} \varepsilon_i' (\bar{H} - L)(\mathbf{f}_i - \bar{\mathbf{f}}) / \sigma_i^2 - 2\sqrt{\frac{b}{n}} \sum_{i=1}^{n} (S_T \varepsilon)' (\bar{H} - L)(\mathbf{f}_i - \bar{\mathbf{f}}) / \sigma_i^2 \\
&\equiv 2D_{nT8,1} - 2D_{nT8,2}.
\end{aligned}
$$

We can show that $D_{nT8,1} = \sqrt{b/n} O_P(\gamma_{nT}(n^{1/2}T^{1/2} + n^{1/2}T^{-1/2}b^{-1})) = O_P(n^{-1/4}b^{1/4} + n^{-1/4}T^{-1}b^{-3/4}) = o_P(1)$, and $D_{nT8,2} = O_P(n^{-1/4}b^{1/4}\sqrt{\log(nT)}) = o_P(1)$. It follows that $D_{nT8} = o_P(1)$. By the Cauchy-Schwarz inequality, $D_{nTl} = o_P(1)$ for $l = 9, 10$. □

PROPOSITION A.4. $\Gamma_{nT,2} = o_P(1)$ *under* $H_1(\gamma_{nT})$.

Proof: Analogously to the proof of Proposition A.2, we can write

$$
\begin{aligned}
\Gamma_{nT,2} &= -\sqrt{\frac{b}{n}} \sum_{i=1}^{n} (ESS_i - \varepsilon_i' Q \varepsilon_i) \frac{\widehat{\sigma}_i^2 - \sigma_i^2}{\sigma_i^4} + \sqrt{\frac{b}{n}} \sum_{i=1}^{n} (ESS_i - \varepsilon_i' Q \varepsilon_i) \frac{\left(\widehat{\sigma}_i^2 - \sigma_i^2\right)^2}{\sigma_i^4 \widehat{\sigma}_i^2} \\
&\equiv -\Gamma_{nT,21} + \Gamma_{nT,22}, \text{ say.}
\end{aligned}
$$

Note that $\widehat{\sigma}_i^2 = \sum_{l=1}^{10} TSS_{il}/T$ by (A.8). First, we want to show that

$$\max_{1\le i\le n}\left|T^{-1}TSS_{i1}-\sigma_i^2\right| = O_P\left(\upsilon_{nT}\right) \text{ and}$$
$$\max_{1\le i\le n} T^{-1}TSS_{il} = o_P\left(\upsilon_{nT}\right) \text{ for } l=2,\ldots,10, \tag{A.13}$$

where $\upsilon_{nT} \equiv n^{1/\lambda}T^{-1/2}$. By (A.9), it suffices to show that $\max_{1\le i\le n} T^{-1}TSS_{il} = o_P(\upsilon_{nT})$, for $l = 7,\ldots,10$. In the sequel, we will frequently use the fact that $\max_{1\le i\le n}\sup_{\tau\in[0,1]}|f_i(\tau)-\overline{f}(\tau)| = O(\gamma_{nT})$ and $\widehat{\beta}-\beta = o_P(\gamma_{nT})$ under $H_1(\gamma_{nT})$ by Lemma B.6(b). Following the study of TSS_{i2} in Proposition A.2, we can show that $\max_{1\le i\le n} T^{-1}TSS_{i7} = o_P(\upsilon_{nT})$. For TSS_{i8} we have

$$T^{-1}TSS_{i8} = T^{-1}\gamma_{nT}^2\Delta_{ni}'M\Delta_{ni} \le T^{-1}\gamma_{nT}^2\left\|\Delta_{ni}\right\|^2$$
$$= n^{-1/2}T^{-2}b^{-1/2}\sum_{t=1}^{T}\Delta_{ni}^2\left(\frac{t}{T}\right) = O(n^{-1/2}T^{-1}b^{-1/2}) = o(\upsilon_{nT})$$

uniformly in i. By the Cauchy-Schwarz inequality, $\max_{1\le i\le n} T^{-1}TSS_{il} = o_P(\upsilon_{nT})$ for $l = 9, 10$. Consequently, we have $\max_{1\le i\le n}|\widehat{\sigma}_i^2-\sigma_i^2| = O_P(\upsilon_{nT})$. By the proof of Proposition A.2, $\frac{b}{n}\sum_{i=1}^{n}(ESS_i-\varepsilon_i'Q\varepsilon_i)^2 = O_P(1)$. It follows that

$$\Gamma_{nT,22} \le \frac{n^{1/2}\max_{1\le i\le n}\left|\widehat{\sigma}_i^2-\sigma_i^2\right|^2}{\min_{1\le i\le n}\sigma_i^4\widehat{\sigma}_i^2}\left[\frac{b}{n}\sum_{i=1}^{n}(ESS_i-\varepsilon_i'Q\varepsilon_i)^2\right]^{1/2}$$
$$= n^{1/2}O_P\left(\upsilon_{nT}^2\right) = o_P\left(1\right).$$

To analyze $\Gamma_{nT,21}$, using (A.8) we can write

$$\Gamma_{nT,21} = \sqrt{\frac{b}{n}}\sum_{i=1}^{n}(ESS_i-\varepsilon_i'Q\varepsilon_i)\frac{\widehat{\sigma}_i^2-\sigma_i^2}{\sigma_i^4} = \sum_{l=1}^{10}\Gamma_{nT,21l},$$

where $\Gamma_{nT,211} \equiv (b/n)^{1/2}\sum_{i=1}^{n}\sigma_i^{-4}(ESS_i-\varepsilon_i'Q\varepsilon_i)(T^{-1}TSS_{i1}-\sigma_i^2)$, and $\Gamma_{nT,21l} \equiv (b/n)^{1/2}\sum_{i=1}^{n}\sigma_i^{-4}\ (ESS_i-\varepsilon_i'Q\varepsilon_i)T^{-1}TSS_{il}$ for $l=2,\ldots,10$. Following the proof of Proposition A.1 and the analysis for TSS_{il} in the proof of Corollary 3.1, we can show that $\Gamma_{nT,21l} = o_P(1)$ for $l=1,\ldots,10$. It follows that $\Gamma_{nT,21} = o_P(1)$. □

PROPOSITION A.5. *$\Gamma_{nT,3} = o_P(1)$ under $H_1(\gamma_{nT})$.*

Proof: By the proof of Corollary 3.1, we can write

$$\Gamma_{nT,3} = \sqrt{\frac{b}{n}}\sum_{i=1}^{n}\sigma_i^{-2}(\widehat{u}_i'\bar{Q}\widehat{u}_i-\varepsilon_i'Q\varepsilon_i) = \sum_{l=1}^{10}\overline{B}_{nT,l}$$

where $\overline{B}_{nT1} = (b/n)^{1/2}\sum_{i=1}^{n}\sigma_i^{-2}(B_{nT,i1}-\varepsilon_i'Q\varepsilon_i)$, and $\overline{B}_{nTl} = (b/n)^{1/2}\sum_{i=1}^{n}\sigma_i^{-2}B_{nT,il}$ for $l = 2,\ldots,10$. Following the argument in the proof of Corollary 3.1, we can readily show that $\overline{B}_{nTl} = o_P(1)$ for $l=1,2,\ldots,6$ as in the case when H_0 holds. It remains to prove that $\overline{B}_{nTl} = o_P(1)$ for

$l = 7, \ldots, 10$ under $H_1(\gamma_{nT})$. Noting that $\lambda_{\max}(M) = 1$, we have

$$\overline{B}_{nT10} = \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \sigma_i^{-2} (\mathbf{f}_i - \bar{\mathbf{f}})' \bar{Q} (\mathbf{f}_i - \bar{\mathbf{f}}) \leq \frac{b^{1/2}\gamma_{nT}^2}{\sqrt{n}} \sum_{i=1}^{n} \sigma_i^{-2} \Delta_{ni}' Q \Delta_{ni}$$

$$= n^{-1}T^{-1} \sum_{i=1}^{n} \sigma_i^{-2} \sum_{t=1}^{T} \Delta_{ni}^2(t/T)(\bar{H}_{tt} - T^{-1}) = O(T^{-1}b^{-1}) = o(1).$$

By the Cauchy-Schwarz inequality, we have $\overline{B}_{nT7} = o(1)$ and $\overline{B}_{nT8} = o_P(1)$. Decompose $\overline{B}_{nT9} = 2n^{-1/2}b^{1/2} \sum_{i=1}^{n} \sigma_i^{-2} \varepsilon_i' \bar{Q}\ \bar{\mathbf{f}}_i^{*} - 2n^{-1/2}b^{1/2}\ \sum_{i=1}^{n} \sigma_i^{-2} (S_T\varepsilon)' \bar{Q}\ \bar{\mathbf{f}}_i^{*}\ \equiv 2\overline{B}_{nT9,1} - 2\overline{B}_{nT9,2}$. By moments calculation and the Chebyshev inequality, we can show that $\overline{B}_{nT9,1} = O_P(T^{1/2}h^{p+1}b^{1/2}) = o_P(1)$, and $\overline{B}_{nT9,2} = O_P(T^{1/2}h^{p+1}b^{1/2}) = o_P(1)$. Consequently $\overline{B}_{nT9} = o_P(1)$. □

PROPOSITION A.6. $\Gamma_{nT,4} = o_P(1)$ *under* $H_1(\gamma_{nT})$.

Proof: Analogously to the proof of Proposition A.2, we can write

$$\Gamma_{nT,4} = -\sqrt{\frac{b}{n}} \sum_{i=1}^{n} (\widehat{u}_i' \bar{Q} \widehat{u}_i - \varepsilon_i' Q \varepsilon_i) \frac{\widehat{\sigma}_i^2 - \sigma_i^2}{\sigma_i^4} + \sqrt{\frac{b}{n}} \sum_{i=1}^{n} (\widehat{u}_i' \bar{Q} \widehat{u}_i - \varepsilon_i' Q \varepsilon_i) \frac{\left(\widehat{\sigma}_i^2 - \sigma_i^2\right)^2}{\sigma_i^4 \widehat{\sigma}_i^2}$$

$$\equiv -\Gamma_{nT,41} + \Gamma_{nT,42}, \text{ say.}$$

We prove the proposition by showing that $\Gamma_{nT,4l} = o_P(1)$ for $l = 1, 2$. For $\Gamma_{nT,41}$, write $\Gamma_{nT,41} = \sum_{l=1}^{10} \Gamma_{nT,41}(l)$, where

$$\Gamma_{nT,41}(1) = \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \sigma_i^{-4} (B_{nT,i1} - \varepsilon_i' Q \varepsilon_i) \left(\widehat{\sigma}_i^2 - \sigma_i^2\right),$$

$$\Gamma_{nT,41}\,(l) = \sqrt{\frac{b}{n}} \sum_{i=1}^{n} \sigma_i^{-4} B_{nT,il} \left(\widehat{\sigma}_i^2 - \sigma_i^2\right) \text{ for } l = 2, \ldots, 10,$$

and $B_{nT,il}$ are defined after (A.11). Further decompose $\Gamma_{nT,41}(1) = \sum_{m=1}^{10} \Gamma_{nT,41}(1, m)$ by using the decomposition $\widehat{\sigma}_i^2 = \sum_{l=1}^{10} TSS_{il}/T$ in (A.8), where $\Gamma_{nT,41}(1, 1) = (b/n)^{1/2} \sum_{i=1}^{n} \sigma_i^{-4} (B_{nT,i1} - \varepsilon_i' Q \varepsilon_i)\ (T^{-1}TSS_{i1} - \sigma_i^2)$ and $\Gamma_{nT,41}(1, m) = (b/n)^{1/2} \sum_{i=1}^{n} \sigma_i^{-4} \times (B_{nT,i1} - \varepsilon_i' Q \varepsilon_i) T^{-1} TSS_{im}$ for $m = 2, \ldots, 10$. It is easy to show that $\Gamma_{nT,41}(1, m) = o_P(1)$ for $m = 1, \ldots, 10$. Consequently $\Gamma_{nT,41}(1) = o_P(1)$. Similarly, we can show $\Gamma_{nT,41}(l) = (b/n)^{1/2} \sum_{i=1}^{n} \sigma_i^{-4} B_{nT,il} (\widehat{\sigma}_i^2 - \sigma_i^2)$ for $l = 2, \ldots, 10$ by using the decomposition of $\widehat{\sigma}_i^2$ in (A.8). It follows that $\Gamma_{nT,41} = o_P(1)$.

For $\Gamma_{nT,42}$, we can apply the decomposition of $\widehat{u}_i' \bar{Q} \widehat{u}_i$ in (A.11) to demonstrate that $(b/n)^{1/2} \sum_{i=1}^{n} |\widehat{u}_i' \bar{Q} \widehat{u}_i - \varepsilon_i' Q \varepsilon_i| = o_P(n^{1/2})$. Then $\Gamma_{nT,42} = o_P(n^{1/2} \upsilon_{nT}^2) = o_P(n/T) = o_P(1)$ by (A.13). □

Proof of Theorem 3.3: As in the proof of Theorem 3.2, we have the decomposition

$$\sqrt{\widehat{\Omega}_{nT}}\,\overline{\Gamma}_{nT} = \overline{\Gamma}_{nT1} - \overline{\Gamma}_{nT2} - \overline{\Gamma}_{nT3} + \overline{\Gamma}_{nT4}, \tag{A.14}$$

where $\overline{\Gamma}_{nTl}$, $l = 1, 2, 3, 4$, are defined analogously to Γ_{nTl} in (A.12) with σ_i^2 being replaced by $\overline{\sigma}_i^2 \equiv \sigma_i^2 + \Upsilon_{i0}$, $\Upsilon_{i0} \equiv \int_0^1 \Delta_i^2(\tau)d\tau - [\int_0^1 \Delta_i(\tau)d\tau]^2$, and recall $\Delta_i(\tau) \equiv f_i(\tau) - f(\tau)$ under H_1. By (A.8), $\widehat{\sigma}_i^2 = T^{-1}\sum_{l=1}^{10} TSS_{il}$. Under H_1, by Lemma B.6(c) the results in (A.9) become

$$\max_{1\le i\le n} \left|T^{-1}TSS_{i1} - \sigma_i^2\right| = o_P(1) \text{ and } \max_{1\le i\le n} T^{-1}TSS_{il} = o_P(1) \text{ for } l = 2, \ldots, 6.$$

We can also show that $T^{-1}TSS_{il} = o_P(1)$ uniformly in i for $l = 7$, 9, and 10. For TSS_{i8}, we have uniformly in i,

$$\begin{aligned} T^{-1}TSS_{i8} &= T^{-1}\sum_{t=1}^{T}[\Delta_i(t/T) - \overline{\Delta}_i]^2 \\ &= \int_0^1 \Delta_i^2(\tau)d\tau - \left(\int_0^1 \Delta_i(\tau)d\tau\right)^2 + o(1) = \Upsilon_{i0} + o(1), \end{aligned}$$

where $\overline{\Delta}_i \equiv T^{-1}\sum_{t=1}^{T} \Delta_i(t/T)$. It follows that uniformly in i

$$\widehat{\sigma}_i^2 = \sigma_i^2 + \Upsilon_{i0} + o_P(1) = \overline{\sigma}_i^2 + o_P(1). \tag{A.15}$$

That is, $\overline{\sigma}_i^2$ is the probability limit of $\widehat{\sigma}_i^2$ under H_1. We prove the theorem by showing that (a) $\Lambda_{nT1} \equiv (n^{1/2}Tb^{1/2})^{-1}\overline{\Gamma}_{nT1} = \Xi_A + o_P(1)$, and (b) $\Lambda_{nTl} \equiv (n^{1/2}Tb^{1/2})^{-1}\overline{\Gamma}_{nTl} = o_P(1)$ for $l = 2, 3, 4$.

Following the proofs of Propositions A.1 and A.3, we can get $\Lambda_{nT1} = (n^{1/2}Tb^{1/2})^{-1} \times \overline{\Gamma}_{nT1} = \overline{\Lambda}_{nT1} + o_P(1)$, where $\overline{\Lambda}_{nT1} \equiv (n^{1/2}Tb^{1/2})^{-1}D_{nT2}$. Following the analysis of D_{nT2} in the proof of Proposition A.3, we have

$$\overline{\Lambda}_{nT1} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\sum_{s=1}^{T}(\bar{H}_{ts} - T^{-1})\Delta_i(t/T)\Delta_i(s/T)/\overline{\sigma}_i^2 = \Theta_A + o(1),$$

where Θ_A is defined analogously to Θ_0 with $(\sigma_i^2, \Delta_{ni})$ being replaced by $(\overline{\sigma}_i^2, \Delta_i)$. This proves (a). Following the proof of Propositions A.2 and A.4–A.6, we can show that $\Lambda_{nTl} = o_P(1)$ for $l = 2, 3, 4$.

APPENDIX B: SOME USEFUL LEMMATA

In this Appendix, we present some technical lemmata that are used in the proofs of the main results in the paper.

LEMMA B.1. *Let* $\lambda_{tT} \equiv \int_0^1 w_b(\frac{t}{T} - \tau)d\tau$. *Then* $\frac{1}{2} \le \min_{1\le t\le T}\lambda_{tT} \le \max_{1\le t\le T}\lambda_{tT} = 1$.

Proof: Write $\lambda_{tT} = \int_0^1 w(\frac{\tau}{b} - \frac{t}{Tb})d(\frac{\tau}{b}) = \int_0^{1/b} w(u - \frac{t}{Tb})du = \int_{-t/(Tb)}^{1/b-t/(Tb)} w(u)du$. Clearly, $\max_{1\le t\le T}\lambda_{tT} = 1$. If $Tb \le t \le T(1-b)$, then $\lambda_{tT} = \int_{-1}^{1} w(u)du = 1$. If $1 \le t = T\epsilon < Tb$ for some $\epsilon \in (0, b)$, then

$$\lambda_{tT} = \int_{-t/(Tb)}^{1/b-t/(Tb)} w(s)\,ds = \int_{-\epsilon}^{1} w(u)\,du \geq \int_{0}^{1} w(u)\,du = \frac{1}{2}$$

where the last equality follows from the symmetry of w and the fact that $\int_{-1}^{1} w(u)du = 1$. Similarly, if $T(1-b) < t = T\epsilon \leq T$ for some $\epsilon \in (1-b, 1)$, then we have $\int_0^1 w_b(\frac{t}{T} - \tau)d\tau = \int_{-t/(Tb)}^{1/b-t/(Tb)} w(u)du = \int_{-1}^{\epsilon} w(u)du \geq \int_{-1}^{0} w(u)du = \frac{1}{2}$. This proves the lemma. □

LEMMA B.2. $\max_{1\leq t,s\leq T} |\bar{H}_{ts}| \leq C_1(Tb)^{-1}$ *for some* $C_1 < \infty$ *where* $\bar{H}_{ts}$ *is the* (t,s)*th element of* $\bar{H}$, $\bar{H} \equiv \int_0^1 H(\tau)d\tau$, *and* $H(\tau) \equiv W_b(\tau)z_b^{[1]}(\tau)[(z_b^{[1]}(\tau)' W_b(\tau) z_b^{[1]}(\tau)]^{-1}\, z_b^{[1]}(\tau)' W_b(\tau)$.

Proof: Let $S_b(\tau) \equiv T^{-1} z_b^{[1]}(\tau)' W_b(\tau) z_b^{[1]}(\tau)$. Then

$$S_b(\tau) = \mathbb{S} + o(1) \text{ uniformly in } \tau \in (0,1), \tag{B.1}$$

where $\mathbb{S} \equiv \left(\begin{smallmatrix} 1 & 0 \\ 0 & \omega_2 \end{smallmatrix}\right)$ and $\omega_2 = \int_{-1}^{1} w(u)u^2 du$. By (B.1), Lemma B.1, and Assumption A4, we have

$$\begin{aligned}
|\bar{H}_{ts}| &= \left| T^{-1} \int_0^1 z_{b,t}^{[1]}(\tau)' \,[S_b(\tau)]^{-1}\, z_{b,s}^{[1]}(\tau) \overline{w}_{b,t}(\tau) \overline{w}_{b,s}(\tau) d\tau \right| \\
&\approx \left| T^{-1} \int_0^1 z_{b,t}^{[1]}(\tau)' \mathbb{S}^{-1} z_{b,s}^{[1]}(\tau) \overline{w}_{b,t}(\tau) \overline{w}_{b,s}(\tau) d\tau \right| \\
&\leq \left| T^{-1} \int_0^1 w_b\left(\frac{t}{T} - \tau\right) w_b\left(\frac{s}{T} - \tau\right) d\tau (\lambda_{tT}\lambda_{sT})^{-1} \right| \\
&\quad + \left| \omega_2^{-1} T^{-1} \int_0^1 \left(\frac{t - \tau T}{Tb}\right)\left(\frac{s - \tau T}{Tb}\right) w_b\left(\frac{t}{T} - \tau\right) w_b\left(\frac{s}{T} - \tau\right) d\tau (\lambda_{tT}\lambda_{sT})^{-1} \right| \\
&\leq C\,(Tb)^{-1} \int_0^1 w_b\left(\frac{t}{T} - \tau\right) d\tau / \lambda_{tT} + C\,(Tb)^{-1} \int \frac{|t - \tau T|}{Tb} w_b\left(\frac{t}{T} - \tau\right) d\tau \\
&\leq C\,(Tb)^{-1} \left(1 + \int_{-1}^{1} |u|\, w(u)\, d\tau\right) \leq C_1\,(Tb)^{-1},
\end{aligned}$$

where $A \approx B$ denotes $A = B(1 + o(1))$. □

LEMMA B.3 (a) $A_{T1} \equiv b \sum_{1\leq t\neq s\leq T} a_{ts}^2 = O(1)$, (b) $A_{T2} \equiv T^{-1} \sum_{1\leq t,s,r\leq 1} |a_{ts} a_{tr}| = O(1)$, *and* (c) $A_{T3} \equiv \|\bar{H} - L\| = O(b^{-1/2})$, *where recall* $a_{ts} \equiv \bar{H}_{ts} - T^{-1}$ *denotes the* (t,s)*th element of* $\bar{H} - L$, *and* $L \equiv T^{-1} i_T i_T'$.

Proof: For (a) it is easy to show that $A_{T1} = \overline{A}_{T1} + O(b)$, where $\overline{A}_{T1} \equiv b \sum_{1\le t\ne s\le T} \bar{H}_{ts}^2$. By (B.1),

$$
\begin{aligned}
\overline{A}_{T1} &\approx \frac{b}{T^2} \sum_{1\le t\ne s\le T} \left\{ \int_0^1 z_{b,t}^{[1]}(\tau)\mathbb{S}^{-1} z_{b,s}^{[1]}(\tau)\overline{w}_{b,t}(\tau)\overline{w}_{b,s}(\tau)d\tau \right\}^2 \\
&= \frac{b}{T^2} \sum_{1\le t\ne s\le T} \left\{ \int_0^1 \left[1 + \omega_2^{-1} \left(\frac{\tau}{b} - \frac{t}{Tb} \right) \left(\frac{\tau}{b} - \frac{s}{Tb} \right) \right] \right\}^2 \\
&\quad \times \frac{1}{b^2} w\left(\frac{\tau}{b} - \frac{s}{Tb} bigg) w\left(\frac{\tau}{b} - \frac{t}{Tb} \right) d\tau \right\}^2 (\lambda_{tT}\lambda_{sT})^{-2} \\
&= \frac{b}{T^2} \sum_{t=\lfloor Tb\rfloor+1}^{\lfloor T(1-b)\rfloor-1} \sum_{s=1}^{T} \left\{ \int_{-1}^1 \left[1 + \omega_2^{-1} u \left(u + \frac{t-s}{Tb} \right) \right] \frac{1}{b} w(u)\, w\left(u + \frac{t-s}{Tb} \right) du \right\}^2 \\
&\quad \times \left\{ \int_0^{1/b} w\left(z - \frac{t}{Tb} \right) dz \int_0^{1/b} w\left(\frac{s-t}{Tb} - \left(z' - \frac{t}{Tb} \right) \right) dz' \right\}^{-2} + O(b) \\
&= \frac{1}{T} \sum_{t=\lfloor Tb\rfloor+1}^{\lfloor T(1-b)\rfloor-1} \int_{-t/(Tb)}^{(T-t)/(Tb)} \left(\int_{-1}^1 \left[1 + \omega_2^{-1} u(u-v) \right] w(u)\, w(u-v)\, du \right)^2 \\
&\quad \times \left(\int_{-t/(Tb)}^{1/b-t/(Tb)} w(z)\, dz \int_{-t/(Tb)}^{1/b-t/(Tb)} w(z'-v) dz' \right)^{-2} dv + o(1) \\
&= \int_b^{1-b} \int_{-1}^1 \left\{ \left(\int_{-1}^1 \left[1 + \omega_2^{-1} u(u-v) \right] w(u)\, w(u-v)\, du \right)^2 \right. \\
&\quad \left. \times \left(\int_{-1}^1 w(z)\, dz \int_{-1}^1 w(z'-v) \right) dz')^{-2} dv dv' \right\} + o(1) \\
&= \int_{-1}^1 \left(\int_{-1}^1 \left[1 + \omega_2^{-1} u(u-v) \right] w(u)\, w(u-v)\, du \right)^2 \left(\int_{-1}^1 w(z-v)\, dz \right)^{-2} dv + o(1) \\
&= O(1).
\end{aligned}
$$

By the same token, we can show (b). For (c), noting that $\|\bar{H} - L\|^2 = \sum_{1\le t\ne s\le T} a_{ts}^2 + \sum_{t=1}^T a_{tt}^2 = O(b^{-1}) + O(T^{-1}b^{-2})$, $\|\bar{H} - L\| = O(b^{-1/2})$ as $T^{-1}b^{-1} = o(1)$. □

LEMMA B.4. *Let* $c_{ts} \equiv e_1'[T^{-1} z_h^{[p]}(t/T)' K_h(t/T) z_h^{[p]}(t/T)]^{-1} z_{h,s}^{[p]}(t/T)$. *Then* (a) $C_{T1} \equiv T^{-2} \sum_{1\le t\ne s\le T} |c_{ts}| k_{h,ts} = O(1)$; (b) $C_{T2} \equiv T^{-2} h \sum_{1\le t\ne s\le T} c_{ts}^2 k_{h,ts}^2 = O(1)$, (c) $C_{T3} \equiv T^{-1} \sum_{t=1}^T |c_{tt}| = O(1)$; (d) $C_{T4} \equiv T^{-1} \sum_{t=1}^T c_{tt}^2 = O(1)$.

Proof: (a) Let $S_{p,h}(\tau) \equiv T^{-1} z_h^{[p]}(t/T)' K_h(t/T) z_h^{[p]}(t/T)$. The (j,l)th element of $S_{p,h}(\tau)$ is $s_{jl}(\tau) = \frac{1}{Th} \sum_{s=1}^T (\frac{s-\tau T}{Th})^{j+l-2} k(\frac{s-\tau T}{Th})$. For any $\tau \in (0,1)$, we have by the definition of Riemann integral that

$$
\begin{aligned}
s_{jl}(\tau) &= \frac{1}{Th}\sum_{r=1}^{T}\left(\frac{r}{Th}-\frac{\tau}{h}\right)^{j+l-2} k\left(\frac{r}{Th}-\frac{\tau}{h}\right) \\
&= \int_{-\tau/(Th)}^{1/h-\tau/(Th)} u^{j+l-2}k(u)du + o(1) = \int_{-1}^{1} u^{j+l-2}k\,(u)\,du + o\,(1)\,.
\end{aligned}
$$

That is, $S_{p,h}(\tau) = \mathbb{S}_p + o(1)$ for any $\tau \in (0, 1)$, where

$$
\mathbb{S}_p = \begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_p \\ \mu_1 & \mu_2 & \cdots & \mu_{p+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_p & \mu_{p+1} & \cdots & \mu_{2p} \end{pmatrix},
$$

and $\mu_j \equiv \int_{-1}^{1} v^j k(v)dv$ for $j = 0, 1, \ldots, 2p$. It follows that

$$
\begin{aligned}
C_{T1} &= \frac{1}{T^2h}\sum_{t=1}^{T}\sum_{s=1}^{T}\left|e_1'\mathbb{S}_p^{-1}\left[1, \frac{s-t}{Th}, \ldots, \left(\frac{s-t}{Th}\right)^p\right]\right| k\left(\frac{s-t}{Th}\right) + o\,(1) \\
&= \frac{1}{T}\sum_{t=1}^{T}\int_{-t/(Th)}^{(T-t)/(Th)} \left|e_1'\mathbb{S}_p^{-1}[1, v, \ldots, v^p]\right| k\,(v)\,dv + o\,(1) \\
&= \frac{1}{T}\sum_{t=\lfloor Th\rfloor+1}^{\lfloor T(1-h)\rfloor-1}\int_{-t/(Th)}^{(T-t)/(Th)} \left|e_1'\mathbb{S}_p^{-1}[1, v, \ldots, v^p]\right| k\,(v)\,dv + o\,(1) \\
&= \int_{-1}^{1} \left|e_1'\mathbb{S}_p^{-1}[1, v, \ldots, v^p]\right| k\,(v)\,dv + o\,(1) = O\,(1)\,.
\end{aligned}
$$

This proves (a). By the same token,

$$
\begin{aligned}
C_{T2} &= \frac{1}{T^2h}\sum_{t=1}^{T}\sum_{s=1}^{T}\left|e_1'\mathbb{S}_p^{-1}\left[1, \frac{s-t}{Th}, \ldots, \left(\frac{s-t}{Th}\right)^p\right]\right|^2 k\left(\frac{s-t}{Th}\right)^2 + o\,(1) \\
&= \int_{-1}^{1} \left|e_1'\mathbb{S}_p^{-1}[1, v, \ldots, v^p]\right|^2 k\,(v)^2\,dv + o\,(1) = O\,(1)\,.
\end{aligned}
$$

Similarly, we can prove (c)–(d). □

LEMMA B.5. $\sup_{\tau\in(0,1)} e_1' S(\tau)\varepsilon = O_P(\sqrt{\log(nT)/(nTh)})$.

Proof: The proof is analogous to that of (A.11) in Chen, Gao, and Li (2011, pp. 27–30). □

LEMMA B.6. *Suppose Assumptions A1–A5 hold. Recall that* $\gamma_{nT} = n^{-1/4}T^{-1/2}b^{-1/2}$ *in* $H_1(\gamma_{nT})$*. Then as* $(n, T) \to \infty$, (a) $\widehat{\beta} - \beta = O_P(n^{-1/2}T^{-1/2})$ *under* H_0; (b) $\widehat{\beta} - \beta = o_P(\gamma_{nT})$

under $H_1(\gamma_{nT})$ provided that A6 also holds; (c) $\widehat{\beta} - \beta = o_P(1)$ *under H_1 provided that A6 also holds.*

Proof: (a) This can be done by following the proof of Theorem 3.1 in CGL (2011). Note that CGL (2011) also proves the asymptotic normality under the independence of $\{(\varepsilon_{it}, v_{it})\}$ across t and the assumption that g_i in Assumption 3.1 is the same for all i ($g_i = g$, say). One can verify that the above probability order can be attained even if we relax their independence condition to our m.d.s. condition and their homogenous trending assumption on g to our heterogeneous case.

(b) Recalling that $\overline{\mathbf{F}} \equiv i_n \otimes \bar{\mathbf{f}}$ and $S_{nT}\mathbf{F} = S_{nT}\overline{\mathbf{F}}$, we have

$$\begin{aligned}\widehat{\beta} - \beta &= (X^{*\prime}M_D X^*)^{-1}X^{*\prime}M_D(\varepsilon^* + \overline{\mathbf{F}}^*) + (X^{*\prime}M_D X^*)^{-1}X^{*\prime}M_D(\mathbf{F} - \overline{\mathbf{F}}) \\ &\equiv d_1 + d_2, \text{ say.}\end{aligned} \tag{B.2}$$

The first term also appears under H_0 and thus $d_1 = O_P(n^{-1/2}T^{-1/2})$. The second term vanishes under H_0 and plays asymptotically non-negligible role under $H_1(\gamma_{nT})$. Let $\overline{d}_2 \equiv X^{*\prime}M_D(\mathbf{F} - \overline{\mathbf{F}})$. Note that

$$\overline{d}_2 = X^{*\prime}(\mathbf{F} - \overline{\mathbf{F}}) - X^{*\prime}D(D'D)^{-1}D(\mathbf{F} - \overline{\mathbf{F}}). \tag{B.3}$$

Similarly to the proof in CGL (2011), we can show that the leading term on the right hand side of the above equation is $X^{*\prime}(\mathbf{F} - \overline{\mathbf{F}})$. Noting that $X_{it} = g_i(t/T) + v_{it}$ and $X^* = (I - S_{nT})X$, we have

$$\begin{aligned}X^{*\prime}(\mathbf{F} - \overline{\mathbf{F}}) &= \sum_{i=1}^{n}\sum_{t=1}^{T}[X_{it} - e_1'S(t/T)X][f_i(t/T) - \overline{f}(t/T)] \\ &= \sum_{i=1}^{n}\sum_{t=1}^{T} v_{it}[f_i(t/T) - \overline{f}(t/T)] - \sum_{i=1}^{n}\sum_{t=1}^{T}\{e_1'S(t/T)V\}[f_i(t/T) - \overline{f}(t/T)] \\ &\quad + \sum_{i=1}^{n}\sum_{t=1}^{T}[g_i(t/T) - \overline{g}(t/T)][f_i(t/T) - \overline{f}(t/T)] \\ &\quad + \sum_{i=1}^{n}\sum_{t=1}^{T}[\overline{g}(t/T) - e_1'S(t/T)\mathbf{G}][f_i(t/T) - \overline{f}(t/T)] \\ &\equiv \Psi_{nT1} - \Psi_{nT2} + \Psi_{nT3} + \Psi_{nT4},\end{aligned} \tag{B.4}$$

where $V \equiv (v_{11}', \ldots, v_{1T}', \ldots, v_{n1}', \ldots, v_{nT}')'$, $\overline{g}(t/T) \equiv n^{-1}\sum_{i=1}^{n} g_i(t/T)$, $\mathbf{g}_i \equiv (g_i(1/T)', \ldots, g_i(T/T)')'$ and $\mathbf{G} \equiv (\mathbf{g}_1', \ldots, \mathbf{g}_n')'$. Clearly $\Psi_{nTl} = 0$ for $l = 2, 4$ by the definition of $\overline{f}$.

Noting that $\max_{1\le i\le n}\sup_{0\le\tau\le 1}|f_i(\tau)-\overline{f}(\tau)| = O(\gamma_{nT})$, we have

$$\begin{aligned}
E\|\Psi_{nT1}\|^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{t=1}^{T} E(v_{it}'v_{jt})[f_i(t/T)-\overline{f}(t/T)][f_j(t/T)-\overline{f}(t/T)] \\
&\le \Big(\max_{1\le i\le n}\sup_{0\le\tau\le 1}|f_i(\tau)-\overline{f}(\tau)|\Big)^2 \left(T\sum_{i=1}^{n}\sum_{j=1}^{n}|E(v_{i1}'v_{j1})|\right) \\
&= O\left(\gamma_{nT}^2\right)O\left(nT\right) = o\left(nT\right),
\end{aligned}$$

implying that $\Psi_{nT1} = o_P(\sqrt{n}T)$. For Ψ_{nT3}, we have

$$\begin{aligned}
|\Psi_{nT3}| &\le \max_{1\le i\le n}\sup_{0\le\tau\le 1}|f_i(\tau)-\overline{f}(\tau)|\sum_{i=1}^{n}\sum_{t=1}^{T}|g_i(t/T)-\overline{g}(t/T)| \\
&= O\left(\gamma_{nT}\right)T\sum_{i=1}^{n}\left(\int_0^1|g_i(\tau)-\overline{g}(\tau)|\,d\tau + O\left(1/T\right)\right) \\
&= O\left(\gamma_{nT}\right)o\left(nT\right) = o\left(\gamma_{nT}nT\right).
\end{aligned}$$

Consequently, we have shown that $X^{*\prime}(\mathbf{F}-\overline{\mathbf{F}}) = O_P(\sqrt{n}T) + o(\gamma_{nT}nT)$. It follows $X^{*\prime}M_D(\mathbf{F}-\overline{\mathbf{F}}) = O_P(\sqrt{n}T)$. Since $(nT)^{-1}X^{*\prime}M_DX^* = O_P(1)$, we have $(X^{*\prime}M_DX^*)^{-1}X^{*\prime}M_D(\mathbf{F}-\overline{\mathbf{F}}) = o_P(\gamma_{nT})$. Thus $\widehat{\beta}-\beta = o_P(\gamma_{nT})$ under $H_1(\gamma_{nT})$.

(c) Using the notation above, we continue to have $d_1 = O_P(n^{-1/2}T^{-1/2})$ and $(nT)^{-1}X^{*\prime}M_DX^* = O_P(1)$ under H_1. For d_2, we analyze the dominant term $X^{*\prime}(\mathbf{F}-\overline{\mathbf{F}})$ by using the same decomposition in (B.4). Clearly, we still have $\Psi_{nT2} = 0$, $\Psi_{nT3} = o_P(nT)$ and $\Psi_{nT4} = 0$. For Ψ_{nT1}, noting that $\max_{1\le i\le n}\sup_{0\le\tau\le 1}|f_i(\tau)-\overline{f}(\tau)| = O(1)$ under H_1, we have $E(\|\Psi_{nT1}\|^2) = O(nT)$, which implies that $\Psi_{nT1} = O_P(\sqrt{n}T)$. Thus $X^{*\prime}(\mathbf{F}-\overline{\mathbf{F}}) = o_P(nT)$ and $\widehat{\beta}-\beta = o_P(1)$ under H_1. □

REMARK. If $g_i(\tau)-\overline{g}(\tau) = 0$ for all $\tau \in [0, 1]$, then from the proof of (ii) and (iii) we can see that $\widehat{\beta}-\beta = O_P(n^{-1/2}T^{-1/2})$ also holds under $H_1(\gamma_{nT})$ and $H_1(1)$ as $\Psi_{nT3} = 0$ in this case.

LEMMA B.7. $\|X - S_{nT}X\|^2 = O_P(nT)$.

Proof: Recall $\mathbf{g}_i \equiv (g_i(1/T), \ldots, g_i(T/T))'$ and $\mathbf{G} \equiv (\mathbf{g}_1', \ldots, \mathbf{g}_n')'$. Noting that $X_{it} = g_i(t/T) + v_{it}$, we have

$$\begin{aligned}
\|X - S_{nT}X\|^2 &= \sum_{i=1}^{n}\sum_{t=1}^{T}\|X_{it} - e_1S(t/T)X\|^2 \\
&= \sum_{i=1}^{n}\sum_{t=1}^{T}\|v_{it} - e_1S(t/T)V + [g_i(t/T)-\overline{g}(t/T)] + [\overline{g}(t/T) - e_1S(t/T)\mathbf{G}]\|^2
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{n}\sum_{t=1}^{T} v_{it}' v_{it} + \sum_{i=1}^{n}\sum_{t=1}^{T} \|e_1 S(t/T) V\|^2 + \sum_{i=1}^{n}\sum_{t=1}^{T} \|g_i(t/T) - \overline{g}(t/T)\|^2 \\
&\quad + \sum_{i=1}^{n}\sum_{t=1}^{T} \|\overline{g}(t/T) - e_1 S(t/T)\mathbf{G}\|^2 + 2\sum_{i=1}^{n}\sum_{t=1}^{T} v_{it}' e_1 S(t/T) V \\
&\quad + 2\sum_{i=1}^{n}\sum_{t=1}^{T} v_{it}' \left(g_i(t/T) - \overline{g}(t/T)\right) + 2\sum_{i=1}^{n}\sum_{t=1}^{T} v_{it}' \left(\overline{g}(t/T) - e_1 S(t/T)\mathbf{G}\right) \\
&\quad + 2\sum_{i=1}^{n}\sum_{t=1}^{T} (e_1 S(t/T) V)' \left(\overline{g}(t/T) - e_1 S(t/T)\mathbf{G}\right) \\
&\quad + 2\sum_{i=1}^{n}\sum_{t=1}^{T} (e_1 S(t/T) V)' \left(g_i(t/T) - \overline{g}(t/T)\right) \\
&\quad + 2\sum_{i=1}^{n}\sum_{t=1}^{T} \left(g_i(t/T) - \overline{g}(t/T)\right)' \left(\overline{g}(t/T) - e_1 S(t/T)\mathbf{G}\right) \equiv \sum_{r=1}^{10} \Pi_{nT,r}, \text{ say.}
\end{aligned}
$$

It is easy to show that: $\Pi_{nT,1} = O_P(nT)$ by the Markov inequality, $\Pi_{nT,2} = O_P(nT \log(nT))/(nTh) = o_P(nT)$, $\Pi_{nT,3} = O(nT)$ by the property of Riemann integral, $\Pi_{nT,4} = O(nTh^{2p+2}) = o(nT)$ by the Taylor expansion. For the remaining terms, it is clear that $\Pi_{nT,r} = 0$ for $r = 9, 10$, and we can show that $\sum_{r=6}^{8} \Pi_{nT,r} = O_P(nT)$ by the Cauchy–Schwarz inequality. □