Discussion Paper No. 1037

A ROBUST APPROACH TO HETEROSKEDASTICITY, ERROR SERIAL CORRELATION AND SLOPE HETEROGENEITY FOR LARGE LINEAR PANEL DATA MODELS WITH INTERACTIVE EFFECTS

Kazuhiko Hayakawa
Shuichi Nagata
Takashi Yamagata

July 2018

The Institute of Social and Economic Research
Osaka University
6-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan

A robust approach to heteroskedasticity, error serial correlation and slope heterogeneity for large linear panel data models with interactive effects*

Kazuhiko Hayakawa†
Hiroshima University

Shuichi Nagata‡
Kwansei Gakuin University

Takashi Yamagata§
University of York and Osaka University

16 July 2018

Abstract

In this paper, we propose a robust approach against heteroskedasticity, error serial correlation and slope heterogeneity for large linear panel data models. First, we establish the asymptotic validity of the Wald test based on the widely used panel heteroskedasticity and autocorrelation consistent (HAC) variance estimator of the pooled estimator under random coefficient models. Then, we show that a similar result holds with the proposed bias-corrected principal component-based estimators for models with unobserved interactive effects. Our new theoretical result justifies the use of the same slope estimator and the variance estimator, both for slope homogeneous and heterogeneous models. This robust approach can significantly reduce the model selection uncertainty for applied researchers. In addition, we propose a novel test for the correlation and dependence of the random coefficient with covariates. The test is of great importance, since the widely used estimators and/or its variance estimators can become inconsistent when the variation of coefficients depends on covariates, in general. The finite sample evidence supports the usefulness and reliability of our approach.

Key Words: panel data; slope heterogeneity; interactive effects; test for correlated random coefficients

JEL Classification: C12, C13, C23.

---

*The authors are grateful to Milda Norkutė for helpful comments.

†Department of Economics, Hiroshima University, 1-2-1 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739-8525 Japan; Tel: +81-82-424-7189; E-mail address: kazuhaya@hiroshima-u.ac.jp. This author gratefully acknowledges the financial support by JSPS KAKENHI JP17K03660.

‡School of Business Administration, Kwansei Gakuin University, 1-155 Ichiban-cho, Uegahara, Nishinomiya, Hyogo 662-8501, Japan; Tel: +81 798-54-6205; E-mail address: snagata@kwansei.ac.jp.

§Corresponding author. Department of Economics and Related Studies, University of York, York YO10 5DD, UK and ISER, Osaka University, 6-1 Mihogaoka, Ibaraki, Osaka 567-0047 Japan; Tel: +44-1904-323788; E-mail address: takashi.yamagata@york.ac.uk. This author gratefully acknowledges the financial support by JSPS KAKENHI JP18K01545 and JP15H05728.

1 Introduction

The recently increasing availability of panel data sets in which both cross-section dimension N and times series dimension T are large has produced opportunities to develop statistical methods to exploit richer information, while presenting associated technical challenges. In particular, controlling cross-sectional dependence, heterogeneity in parameters and distributions, and serial dependence has been a main focus of the literature.

The cerebrated fixed effects model permits intercept to be cross-sectionally heterogeneous whilst slope coefficients are constant across cross-section units and time. Hansen (2007) has shown that, under mild conditions, the heteroskedasticity and autocorrelation consistent (HAC) variance estimator of Arellano (1987), which is originally proposed for a short panel fixed effects estimator, will be asymptotically valid for large panels. Greenaway-McGrevy et al. (2012) propose to use the HAC estimator for the pooled principal component based (PC) estimator for the model with unobserved interactive effects.

The random-coefficient model, in which the slope coefficients are allowed to vary with the cross-sectional units, has attracted great attention in recent years.[1] It can control differences in behaviour across cross-section units which are not captured by the control variables. For such models, the estimate of interest is often the population average of slope coefficients. Interestingly, if the cross-sectional variation of slopes in the random coefficient model is independent of covariates, the fixed effects estimator is consistent to the population average of slope coefficients. A non-parametric variance-covariance estimator for such pooled estimators has been implicitly proposed in Pesaran (2006), in which the population variation of slopes is replaced by its sample counterpart – the variation of the estimates of cross-section specific slopes. The evidence has shown that the variance estimator behaves very well in finite samples.

There are some issues about this variance estimator for our robust approach. First, for the choice between the HAC and this variance estimator, the practitioner would like to know if there is slope heterogeneity or not. Second, the computation of the variance estimator requires a calculation of the individual slope estimates, and this can be costly when N and T are very large. Third, some estimation methods, such as Bai's (2009) estimator, do not permit slope heterogeneity models, and making use of statistics involving individual slope estimates might not be asymptotically justified.

In this paper, we propose a robust approach against heteroskedasticy, error serial correlation and slope heterogeneity for large linear panel data models. First, we establish the asymptotic validity of the Wald test based on the panel HAC variance estimator for the pooled estimator under random coefficient models. Then, we show that a similar result holds with the bias-corrected PC estimators for models with interactive effects, which extend the results in Westerlund and Urbain (2015) and Reese and Westerlund (2018). Our new theoretical result justifies the use of the same slope estimator and the variance estimator, both for slope homogeneous and heterogeneous models. This robust approach is expected to substantially reduce the model selection uncertainty for applied researchers.

Another main contribution of this paper is a novel test for the correlation and dependency of the random coefficient on covariates. We extend the test proposed by

[1] See Hsiao and Pesaran (2008) for an excellent survey of random coefficient panel data models.

Wooledridge (2010) by robustifying against (uncorrelated) random coefficients, proposing a Lagrange Multiplier test along with a Wald test, and developing them for the models with unobservable interactive effects. The test is of great importance, since the widely used estimators and/or its variance estimators can become inconsistent when the variation of coefficients is correlated or dependent with covariates, in general.

We have examined the finite sample performance of the estimators, tests of linear restrictions, and the LM tests for correlated random coefficients. The evidence illustrates the usefulness of our approach. In particular, for the estimation of the models with unobserved interactive effects, the size of the proposed robust Wald test using the bias-corrected PC estimators and Bai's (2009) estimator is very close to the nominal level, under both slope homogeneity and slope heterogeneity, while maintaining satisfactory power. Also, the LM tests for correlated random coefficients have correct size under both slope homogeneity and slope heterogeneity due to pure random coefficients, while exhibiting high power when the random coefficients depend on covariates.

The paper is organised as follows. The robust Wald test is proposed for standard linear panel data models in Section 2, then for the models with unobserved interactive effects in Section 3. A test for correlation of slopes with covariates is proposed in Section 4. The finite sample performance of the proposed bias-corrected estimator, the associated Wald test and the correlation test is investigated using the Monte Carlo method in Section 5. Section 6 contains some concluding remarks. Proofs of the main results are contained in Appendix, and the proofs of associated Lemmas and full experimental results are found in Online Appendices.

Notations: $\|\mathbf{A}\| = [tr(\mathbf{A}'\mathbf{A})]^{1/2}$, $\mu_{\max}(\mathbf{A})$ ($\mu_{\min}(\mathbf{A})$) is the maximum (minimum) eigenvalue of square matrix $\mathbf{A}$, "$\rightarrow_p$" denotes convergence in probability, "$\rightarrow_d$" denotes convergence in distribution, Δ is an upper bound which is a finite positive constant, $\Delta_{\min}$ is an lower bound which is a finite positive constant strictly above zero, N denotes the number of cross-section units and T denotes the number of time-series observations of panel data, $\delta_{NT} = \min\left\{\sqrt{N}, \sqrt{T}\right\}$, $(N,T) \rightarrow \infty$ denotes N and T go to infinity jointly, $\mathbf{M_A} = \mathbf{I} - \mathbf{A}(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'$, where $\mathbf{A}$ has full column rank.

2 Benchmark Panel Data Model

Consider a panel data model with cross-sectionally heterogeneous slopes:

$$y_{it}^* = \mathbf{x}_{it}^{*\prime}\boldsymbol{\beta}_i + \mathbf{f}_t^{0\prime}\boldsymbol{\lambda}_i^0 + \varepsilon_{it}, \tag{1}$$

$i = 1, 2, ..., N$, $t = 1, 2, ..., T$, $\mathbf{x}_{it}^*$ is a $k \times 1$ vector of observed covariates, $\mathbf{f}_t^0$ is a $r \times 1$ vector of time-variant but cross-sectionally invariant regressor, $\boldsymbol{\lambda}_i^0$ is a $r \times 1$ factor loading vector, which is time-invariant but cross-sectionally variant, and ε_{it} is disturbances. The $k \times 1$ slope coefficients are generated as

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{\eta}_i, \tag{2}$$

where $\boldsymbol{\eta}_i$ is independently distributed random vector across i, with $E(\boldsymbol{\eta}_i) = \mathbf{0}$. When $\boldsymbol{\eta}_i = \mathbf{0}$ for all i, it reduces to the homogeneous slope model. Throughout the paper, our interest is in the estimation and testing of the linear restrictions of $\boldsymbol{\beta}$. Now stuck the T equations of (1) to form

$$\mathbf{y}_i^* = \mathbf{X}_i^*\boldsymbol{\beta}_i + \mathbf{F}^0\boldsymbol{\lambda}_i^0 + \boldsymbol{\varepsilon}_i, \tag{3}$$

where $\mathbf{y}_i^* = (y_{i1}^*, y_{i2}^*, ..., y_{iT}^*)'$, $\mathbf{X}_i^* = (\mathbf{x}_{i1}^*, \mathbf{x}_{i2}^*, ..., \mathbf{x}_{iT}^*)'$, $\mathbf{F}^0 = (\mathbf{f}_1^0, \mathbf{f}_2^0, ..., \mathbf{f}_T^0)'$, and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iT})'$.

In this section, we assume $\mathbf{F}^0$ is observed.[2] Consider the projection matrix $\mathbf{M}_{\mathbf{F}^0} = \mathbf{I} - \mathbf{F}^0\left(\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}$. By Frisch–Waugh–Lovell theorem, denoting $\mathbf{y}_i = \mathbf{M}_{\mathbf{F}^0}\mathbf{y}_i^*$ and $\mathbf{X}_i = \mathbf{M}_{\mathbf{F}^0}\mathbf{X}_i^*$, the model of interest can be equivalently written as[3]

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \boldsymbol{\epsilon}_{f\varepsilon,i},\ \boldsymbol{\epsilon}_{f\varepsilon,i} = \mathbf{F}^0\boldsymbol{\lambda}_i^0 + \boldsymbol{\varepsilon}_i. \tag{4}$$

Remark 1 *For the discussion below, we restrict $\mathbf{f}_t^0$ to be time varying and $\boldsymbol{\lambda}_i^0$ to be cross-sectionally varying, without loss of generality. When $\mathbf{f}_t^0 = \mathbf{f}^0$ or $\boldsymbol{\lambda}_i^0 = \boldsymbol{\lambda}^0$, interactive effects will reduce to $\alpha_i = \mathbf{f}^{0\prime}\boldsymbol{\lambda}_i^0$ and $\delta_t = \mathbf{f}_t^{0\prime}\boldsymbol{\lambda}^0$, which are additive individual effects and time effects, respectively. For notational simplicity, we do not include these effects on top of interactive effects, but all the discussion below will hold by replacing $\{y_{it}^*, \mathbf{x}_{it}^{*\prime}\}$ with transformed variables $\{\ddot{y}_{it}^*, \ddot{\mathbf{x}}_{it}^{*\prime}\}$, where $\ddot{y}_{it}^* = (y_{it}^* - \bar{y}_i^* - \bar{y}_t^* + \bar{y}^*)$ and $\ddot{\mathbf{x}}_{it}^* = (\mathbf{x}_{it}^* - \bar{\mathbf{x}}_i^* - \bar{\mathbf{x}}_t^* + \bar{\mathbf{x}}^*)$ with $\bar{y}_i^* = T^{-1}\sum_{t=1}^T y_{it}^*$, $\bar{y}_t^* = N^{-1}\sum_{i=1}^N y_{it}^*$, $\bar{y}^* = N^{-1}\sum_{i=1}^N \bar{y}_i^*$, and $\bar{\mathbf{x}}_i^*$, $\bar{\mathbf{x}}_t^*$ and $\bar{\mathbf{x}}^*$ are defined analogously.*

We can rewrite the equation (4) as[4]

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{u}_i,\ \mathbf{u}_i = \mathbf{X}_i\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_{f\varepsilon,i}. \tag{5}$$

The pooled estimator of $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}} = \left(\sum_{i=1}^N \mathbf{X}_i'\mathbf{X}_i\right)^{-1}\sum_{i=1}^N \mathbf{X}_i'\mathbf{y}_i. \tag{6}$$

To analyse the asymptotic properties of $\hat{\boldsymbol{\beta}}$, we extend the assumptions in Hansen (2007) to accommodate random coefficient models as follows:

Assumption A1: $\{\mathbf{x}_{it}', \varepsilon_{it}\}$ is independent across $i = 1, 2, ..., N$ for all t, a strong mixing sequence in t with α of size $-3s/(s-4)$ for $s > 4$, with $E|\varepsilon_{it}|^{4+4\delta} \le \Delta < \infty$, $E|x_{ith}|^{8+8\delta} \le \Delta < \infty$ for all $i, t, h = 1, 2, ..., k$ and $E(\boldsymbol{\varepsilon}_i|\mathbf{X}_i) = \mathbf{0}$; $\|\boldsymbol{\beta}\| \le \Delta$; $\{\boldsymbol{\eta}_i\}$ is independent across $i = 1, 2, ..., N$ and of $\{\boldsymbol{\varepsilon}_i\}$ for all i, $E|\eta_{ih}|^{4+4\delta} \le \Delta < \infty$ and $E(\boldsymbol{\eta}_i|\mathbf{X}_i) = \mathbf{0}$.

Assumption A2: (Identification): $\mathbf{A}_{iT} = T^{-1}E(\mathbf{X}_i'\mathbf{X}_i)$ is uniformly positive definite and $\mathbf{A} = \lim_{N,T\to\infty}\mathbf{A}_{NT}$, with $\mathbf{A}_{NT} = N^{-1}\sum_{i=1}^N \mathbf{A}_{iT}$, is fixed and positive definite.

Assumption A3: (Variance Matrix 1): $\mathbf{B}_{iT} = T^{-1}E(\mathbf{X}_i'\boldsymbol{\Sigma}_{\varepsilon\varepsilon i}\mathbf{X}_i)$ and $\boldsymbol{\Sigma}_{\varepsilon\varepsilon i} = E(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'|\mathbf{X}_i)$ are uniformly positive definite and $\mathbf{B} = \lim_{N,T\to\infty}\mathbf{B}_{NT}$, with $\mathbf{B}_{NT} = N^{-1}\sum_{i=1}^N \mathbf{B}_{iT}$, is fixed and positive definite.

[2] In the next section, we consider the case in which $\mathbf{F}^0$ is not observable.

[3] Of course, as $\mathbf{X}_i'\mathbf{F}^0\boldsymbol{\lambda}_i^0 = \mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\mathbf{F}^0\boldsymbol{\lambda}_i^0 = \mathbf{0}$, we could replace $\boldsymbol{\epsilon}_{f\varepsilon,i}$ in (4) by $\boldsymbol{\varepsilon}_i$. We prefer $\boldsymbol{\epsilon}_{f\varepsilon,i}$ here as it will ease the discussions in the next section.

[4] Clearly, in this section, $\boldsymbol{\epsilon}_{f\varepsilon,i}$ in (5) can be replaced by $\boldsymbol{\varepsilon}_i$.

Assumption A4: (Variance Matrix 2): $\mathbf{C}_{iT} = T^{-2}E\left(\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\Omega}_{\eta\eta,i}\mathbf{X}_i'\mathbf{X}_i\right)$ and $\boldsymbol{\Omega}_{\eta\eta,i} = E\left(\boldsymbol{\eta}_i\boldsymbol{\eta}_i'|\mathbf{X}_i\right)$ are uniformly positive definite and $\mathbf{C} = \lim_{N,T\to\infty}\mathbf{C}_{NT}$, with $\mathbf{C}_{NT} = N^{-1}\sum_{i=1}^{N}\mathbf{C}_{iT}$, is fixed and positive definite.

Assumption A1 allows serial dependence in $\{\mathbf{x}_{it}', \varepsilon_{it}\}$ but assumes independence across i. The random coefficient is independent across i. Both the idiosyncratic errors and random coefficient are assumed to be uncorrelated with $\mathbf{x}_{it}$. Assumption A2 is a fairly standard identification condition. Assumption A3 allows conditional heteroskedasticity across i and t. Assumption A4 permits a conditionally heteroskedastic random coefficient process.

For later use, let us define the sample counterpart of $\mathbf{A}_{NT}$ and $\mathbf{A}_{iT}$ defined in Assumption A2:

$$\bar{\mathbf{A}}_{NT} = N^{-1}\sum_{i=1}^{N}\bar{\mathbf{A}}_{iT},\ \bar{\mathbf{A}}_{iT} = T^{-1}\mathbf{X}_i'\mathbf{X}_i. \tag{7}$$

Substituting (4) into (6) gives

$$\begin{aligned}\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} &= \left(\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\right)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{u}_i \qquad (8)\\ &= \bar{\mathbf{A}}_{NT}^{-1}\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\boldsymbol{\varepsilon}_i + \frac{1}{N}\sum_{i=1}^{N}\bar{\mathbf{A}}_{iT}\boldsymbol{\eta}_i\right).\end{aligned}$$

Let us consider the asymptotic properties of the first term of the second equality in (8). We state the following theorem, which is proven by Hansen (2007):

Theorem 1 *Consider model (5). Under Assumptions A1-A3, as* $(N,T)\to\infty$,

$$\bar{\mathbf{A}}_{NT}^{-1}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i'\boldsymbol{\varepsilon}_i \to_d N\left(\mathbf{0}, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\right) \tag{9}$$

where $\bar{\mathbf{A}}_{NT}$, $\mathbf{A}$, *and* $\mathbf{B}$ *are defined in (7), Assumptions A2 and A3, respectively.*

This is a very useful result, since, in the absence of slope heterogeneity $\boldsymbol{\eta}_i$, even when the dimension of $\boldsymbol{\Sigma}_{\varepsilon\varepsilon i} = E\left(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'|\mathbf{X}_i\right)$ is unbounded as $T\to\infty$ (but $\mu_{\max}\left(\boldsymbol{\Sigma}_{\varepsilon\varepsilon i}\right) \le \Delta$ with serially correlated errors), the theorem tells us that the use of the celebrated heteroskedasticity and autocorrelation consistent (HAC) variance estimator of Arellano (1987) for short panel models will be asymptotically justified for large panels.

The next theorem states the asymptotic properties of the first term of the second equality in (8).

Theorem 2 *Consider model (5). Under Assumptions A1, A2 and A4, as* $(N,T)\to\infty$,

$$\bar{\mathbf{A}}_{NT}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\bar{\mathbf{A}}_{iT}\boldsymbol{\eta}_i \to_d N\left(\mathbf{0}, \mathbf{A}^{-1}\mathbf{C}\mathbf{A}^{-1}\right) \tag{10}$$

where $\bar{\mathbf{A}}_{NT}$ *and* $\bar{\mathbf{A}}_{iT}$ *are defined in (7),* $\mathbf{A}$ *and* $\mathbf{C}$ *are defined in Assumptions A2 and A4, respectively.*

As discussed in Pesaran (2006) and Reese and Westerlund (2018), the pooled estimator $\hat{\boldsymbol{\beta}}$ is consistent to the centred value $\boldsymbol{\beta}$ under the random coefficient assumption, and the variation of $\hat{\boldsymbol{\beta}}$ due to the dispersion of slope coefficients dominates the variation due to the linear function of idiosyncratic errors. The following corollary of these two theorems clarify this point:

Corollary 1 *Consider model (5). Under Assumptions A1-A4, as* $(N, T) \to \infty$,

$$\sqrt{N}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \to_d N\left(\mathbf{0}, \mathbf{A}^{-1}\mathbf{C}\mathbf{A}^{-1}\right) \tag{11}$$

whilst under slope homogeneity, $\boldsymbol{\eta}_i = \mathbf{0}$ *for all* i,

$$\sqrt{NT}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \to_d N\left(\mathbf{0}, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}\right), \tag{12}$$

where $\hat{\boldsymbol{\beta}}$ *is defined by (6),* $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$ *are defined in Assumptions A2, A3 and A4, respectively.*

In view of this, Pesaran (2006) proposes to estimate the variance of $\hat{\boldsymbol{\beta}}$ under random coefficient assumption by

$$\tilde{\mathbf{V}}_{NT}\left(\hat{\boldsymbol{\beta}}\right) = N^{-1}\bar{\mathbf{A}}_{NT}^{-1}\ddot{\mathbf{C}}_{NT}\bar{\mathbf{A}}_{NT}^{-1}, \tag{13}$$

where

$$\ddot{\mathbf{C}}_{NT} = N^{-1}\sum_{i=1}^{N}\bar{\mathbf{A}}_{iT}\left(\hat{\boldsymbol{\beta}}_i - \overline{\hat{\boldsymbol{\beta}}}\right)\left(\hat{\boldsymbol{\beta}}_i - \overline{\hat{\boldsymbol{\beta}}}\right)'\bar{\mathbf{A}}_{iT}, \tag{14}$$

$\hat{\boldsymbol{\beta}}_i = \left(\mathbf{X}_i'\mathbf{X}_i\right)^{-1}\mathbf{X}_i'\mathbf{y}_i$ and $\overline{\hat{\boldsymbol{\beta}}} = N^{-1}\sum_{i=1}^{N}\hat{\boldsymbol{\beta}}_i$. The idea is to approximate the unobserved slope heterogeneity $\boldsymbol{\eta}_i$ by its sample counterparts, $\hat{\boldsymbol{\beta}}_i - \overline{\hat{\boldsymbol{\beta}}}$. The empirical evidence has proven that this estimator works well in finite samples.[5] There are some issues with this variance estimator for our robust approach. First, because it is different from the HAC variance estimator assuming slope homogeneity, at the choice the practitioner would like to know if there is slope heterogeneity or not. Second, the computation of the variance estimator requires a calculation of the individual slope estimates, and this can be costly when N and T are large. Third, some estimation methods, such as Bai's (2009) estimator, do not permit slope heterogeneity models and computation of statistics involving individual slope estimates might not be justified. In practice we do not necessarily have a priori information on whether slopes are homogeneous or heterogeneous, which may make the choice of the variance estimator subject to uncertainty.[6]

We propose a simple robust approach against such a choice. Based on the above discussion, under slope heterogeneity we have

$$\begin{aligned}\frac{1}{NT^2}\sum_{i=1}^{N}E\left(\mathbf{X}_i'\mathbf{u}_i\mathbf{u}_i'\mathbf{X}_i\right) &= \frac{1}{NT^2}\sum_{i=1}^{N}E\left(\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\Omega}_{\eta\eta,i}\mathbf{X}_i'\mathbf{X}_i\right) + \frac{1}{NT^2}\sum_{i=1}^{N}E\left(\mathbf{X}_i'\boldsymbol{\Sigma}_{\varepsilon\varepsilon i}\mathbf{X}_i\right) \\ &= \frac{1}{NT^2}\sum_{i=1}^{N}E\left(\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\Omega}_{\eta\eta,i}\mathbf{X}_i'\mathbf{X}_i\right) + O\left(T^{-1}\right).\end{aligned} \tag{15}$$

[5] See experimental results in Pesaran (2006), for example.

[6] Pesaran and Yamagata (2008) and Su and Chen (2013), for example, propose slope homogeneity tests, which can guide such a choice.

This suggests a new alternative estimator of $\mathbf{C}$:

$$\hat{\mathbf{C}}_{NT} = N^{-1}\sum_{i=1}^{N}\hat{\mathbf{C}}_{iT},\ \hat{\mathbf{C}}_{iT} = \frac{\mathbf{X}_i'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\mathbf{X}_i}{T^2}, \tag{16}$$

where $\hat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}$.

Under homogeneous slopes ($\boldsymbol{\eta}_i = \mathbf{0}$ for all i), $\frac{1}{NT}\sum_{i=1}^{N} E\left(\mathbf{X}_i'\mathbf{u}_i\mathbf{u}_i'\mathbf{X}_i\right) = \frac{1}{NT}\sum_{i=1}^{N} E\left(\mathbf{X}_i'\boldsymbol{\Sigma}_{\varepsilon\varepsilon i}\mathbf{X}_i\right)$ as $\mathbf{u}_i = \boldsymbol{\varepsilon}_i$, hence, following Hansen (2007), we propose the following estimator of **B:**

$$\hat{\mathbf{B}}_{NT} = N^{-1}\sum_{i=1}^{N}\hat{\mathbf{B}}_{iT},\ \hat{\mathbf{B}}_{iT} = \frac{\mathbf{X}_i'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\mathbf{X}_i}{T}. \tag{17}$$

We summarise the asymptotic properties of the estimators $\hat{\mathbf{C}}_{NT}$ and $\hat{\mathbf{B}}_{NT}$ in the following proposition:[7]

Proposition 1 *Consider the model (4) and the pooled estimator* $\hat{\boldsymbol{\beta}}$*, which is defined by (6). Under Assumptions A1-A4, under slope heterogeneity* $\hat{\mathbf{C}}_{NT}\rightarrow_p\mathbf{C}$*, whilst under slope homogeneity* ($\boldsymbol{\eta}_i = \mathbf{0}$ *for all* i) $\hat{\mathbf{B}}_{NT}\rightarrow_p\mathbf{B}$*, as* $(N,T)\rightarrow\infty$*, where* $\hat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}$*,* $\hat{\mathbf{C}}_{NT}$ *and* $\hat{\mathbf{B}}_{NT}$ *are defined by (16) and (17), and* $\mathbf{C}$ *and* $\mathbf{B}$ *are defined in Assumptions A3 and A4.*

This proposition implies that the use of a widely employed HAC variance estimator for short panel data models,

$$\hat{\mathbf{V}}_{NT}\left(\hat{\boldsymbol{\beta}}\right) = \left(\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\right)^{-1}\left[\sum_{i=1}^{N}\mathbf{X}_i'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\mathbf{X}_i\right]\left(\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\right)^{-1}, \tag{18}$$

is asymptotically justified for large panel data models under both slope homogeneity and slope heterogeneity.

When there is strong evidence that coefficients are heterogeneous, an alternative pooled estimator, such as a mean group estimator, may be preferred. In this paper we are more in line with the robust approach, which is widely employed in the literature - avoiding uncertainty in specifying and estimating 'nuisance' parameters for potential efficiency gain. As will be discussed in the next section, this approach turns out to be useful for some popular estimation methods, in particular, estimation of linear panel data models with unobserved interactive effects.

We close this section by presenting a result for the Wald test based on the proposed robust variance estimator of $\hat{\boldsymbol{\beta}}$.

Theorem 3 *Consider testing* q *linearly independent restrictions of* $\boldsymbol{\beta}$*,* $H_0 : \mathcal{R}\boldsymbol{\beta} = \mathbf{r}$ *against* $H_1 : \mathcal{R}\boldsymbol{\beta} \neq \mathbf{r}$*, where* $\mathcal{R}$ *is a* $q\times k$ *fixed matrix of full row rank. Consider the model (4) and the Wald test statistic*

$$W_{NT} = \left(\mathcal{R}\hat{\boldsymbol{\beta}} - \mathbf{r}\right)'\left\{\mathcal{R}\left[\hat{\mathbf{V}}_{NT}\left(\hat{\boldsymbol{\beta}}\right)\right]\mathcal{R}'\right\}^{-1}\left(\mathcal{R}\hat{\boldsymbol{\beta}} - \mathbf{r}\right), \tag{19}$$

where $\hat{\boldsymbol{\beta}}$ *and* $\hat{\mathbf{V}}_{NT}\left(\hat{\boldsymbol{\beta}}\right)$ *are defined by (6) and (18), respectively. Suppose that Assumptions A1-A4 hold. Then, under the* H_0*, for both heterogenous slopes and homogeneous slopes* ($\boldsymbol{\eta}_i = \mathbf{0}$ *for all* i)*,* $W_{NT}\rightarrow_d\chi_q^2$*, as* $(N,T)\rightarrow\infty$*.*

[7] The proof of the consistency of $\hat{\mathbf{B}}_{NT}$ is given by Hansen (2007).

3 Models with Unobserved Interactive Effects

When $\mathbf{F}^0$ is unobserved, it should be replaced with a suitable estimator, and in this case a further careful analysis is required. In particular, using estimated variables will result in some asymptotic biases in the pooled estimator, as discussed in Pesaran (2006), Bai (2009) and Westerlund and Urbain (2015), among others. Here we follow the discussion in Westerlund and Urbain (2015) and Reese and Westerlund (2018). Our theoretical contributions to this strand of literature are: (i) establishing the consistency of a bias-corrected estimator both under homogeneous and heterogeneous slopes;[8] (ii) showing the limit distribution of the Wald test statistic based on the HAC variance estimator both under homogeneous and heterogeneous slopes,[9] and; (iii) proposing a new test for correlation and dependence of the random coefficients with the regressors (in the next section).

In this section we assume that $\mathbf{X}_i^*$ has a linear factor structure,[10]

$$\mathbf{X}_i^* = \mathbf{F}^0\mathbf{\Gamma}_i^{0\prime} + \mathbf{V}_i. \tag{20}$$

By combining (4) and (20), we have a system of $m = k+1$ equations:

$$\mathbf{Z}_i^* = \mathbf{F}^0\mathbf{G}_i^0 + \mathbf{E}_i \tag{21}$$

where $\mathbf{Z}_i^* = (\mathbf{y}_i^*, \mathbf{X}_i^*)$, $\mathbf{G}_i^0 = \left(\mathbf{\Gamma}_i^{0\prime}\boldsymbol{\beta}_i + \boldsymbol{\lambda}_i^0, \mathbf{\Gamma}_i^{0\prime}\right)$, $\mathbf{E}_i = (\mathbf{V}_i\boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \mathbf{V}_i)$. For later usage, define

$$\mathbf{\Upsilon}_N^0 = N^{-1}\sum_{i=1}^N \mathbf{G}_i^0\mathbf{G}_i^{0\prime}. \tag{22}$$

In line with Bai (2009) and Norkute et al (2018), we replace Assumptions A1-A4 with the followings:

Assumption B1 (idiosyncratic error in y): (i) ε_{it} is independently distributed across i; (ii) $E(\varepsilon_{it}) = 0$ and $E|\varepsilon_{it}|^{8+\delta} \leq \Delta < \infty$; (iii) $T^{-1}\sum_{s=1}^T\sum_{t=1}^T E|\varepsilon_{is}\varepsilon_{it}|^{1+\delta} \leq \Delta < \infty$; (iv) $E\left|N^{-1/2}\sum_{i=1}^N[\varepsilon_{is}\varepsilon_{it} - E(\varepsilon_{is}\varepsilon_{it})]\right|^4 \leq \Delta < \infty$ for every t and s; (v) $N^{-1}T^{-2}\sum_{i=1}^N\sum_{t=1}^T\sum_{s=1}^T\sum_{r=1}^T\sum_{w=1}^T|\text{cov}(\varepsilon_{is}\varepsilon_{it}, \varepsilon_{ir}\varepsilon_{iw})| \leq \Delta < \infty$; (vi) $\Sigma_{\varepsilon\varepsilon,i} = E(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i')$ is positive definite and its largest eigenvalue is bounded, uniformly every i and T.

Assumption B2 (idiosyncratic error in x): (i) $v_{\ell it}$ is independently distributed across i and group-wise independent from ε_{it}; (ii) $E(v_{\ell it}) = 0$ and $E|v_{\ell it}|^{8+\delta} \leq \Delta < \infty$; (iii) $T^{-1}\sum_{s=1}^T\sum_{t=1}^T E|v_{\ell is}v_{\ell it}|^{1+\delta} \leq \Delta < \infty$; (iv) $E\left|N^{-1/2}\sum_{i=1}^N[v_{\ell is}v_{\ell it} - E(v_{\ell is}v_{\ell it})]\right|^4 \leq$

[8] Westerlund and Urbain (2015; supplement) only prove the consistency of the bias-corrected estimator under the slope homogeneity. Pesaran (2006) and Reese and Westerlund (2015) provide a proof of consistency of the non-bias-corrected pooled estimator.

[9] Theorem 2 in Reese and Westerlund (2015) shows the limit distribution of the pooled estimator allowing weak factors, but does not discuss estimation of asymptotic variance and associated Wald test, nor bias-correction.

[10] Bai (2009) does not impose such a structure and this generality introduces two extra bias terms. In the Monte Carlo section, we apply our approach to Bai's non-linear estimator.

$\Delta < \infty$ for every ℓ, t and s; (v) $N^{-1}T^{-2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{w=1}^{T}\left|\text{cov}\left(v_{\ell is}v_{\ell it}, v_{\ell ir}v_{\ell iw}\right)\right| \leq \Delta < \infty$; (vi) the largest eigenvalue of $E\left(\mathbf{v}_{\ell i}\mathbf{v}_{\ell i}'\right)$ is bounded uniformly for every ℓ, i and T.

Assumption B3 (factor components): (i) $E\left\|\mathbf{f}_t^0\right\|^4 \leq \Delta < \infty$ and $T^{-1}\sum_{t=1}^{T}\mathbf{f}_t^0\mathbf{f}_t^{0\prime} \rightarrow_p \mathbf{\Sigma}_f$ as $T \rightarrow \infty$ which is a fixed positive definite matrix, $\mathbf{f}_t^0$ is group-wise independent from $\mathbf{v}_{it}$ and ε_{it}; (ii) denoting $\mathbf{H}_i^0 = \left[\mathbf{\Gamma}_i^{0\prime}, \boldsymbol{\lambda}_i^0\right]'$, $E\left(\mathbf{H}_i^0\right) = \mathbf{0}$, $\mathbf{H}_i^0$ is independent across i, $E\left\|\mathbf{H}_i^0\right\|^4 \leq \Delta < \infty$ and $N^{-1}\sum_{i=1}^{N}\mathbf{H}_i^{0\prime}\mathbf{H}_i^0 \rightarrow_p \mathbf{\Omega}_H = \begin{bmatrix} \mathbf{\Omega}_{\mathbf{\Gamma\Gamma}}^0 & \boldsymbol{\omega}_{\mathbf{\Gamma}\boldsymbol{\lambda}}^0 \\ \boldsymbol{\omega}_{\mathbf{\Gamma}\boldsymbol{\lambda}}^{0\prime} & \omega_{\boldsymbol{\lambda\lambda}}^0 \end{bmatrix}$, which is a fixed positive definite matrix, $\mathbf{H}_i^0$ is group-wise independent from $\mathbf{v}_{it}$ and ε_{it}; (iii) $\mathbf{\Upsilon}_N^0 \rightarrow_p \mathbf{\Upsilon}^0$ as $N \rightarrow \infty$, which is a fixed positive definite matrix.

Assumption B4 (random coefficient): $\boldsymbol{\eta}_i$ is independent across i, $E\left(\boldsymbol{\eta}_i\right) = \mathbf{0}$, $E(\boldsymbol{\eta}_i\boldsymbol{\eta}_i') = \bar{\mathbf{\Omega}}_{\eta\eta,i}$ which is a fixed positive definite matrix uniformly for every i, $E\left\|\boldsymbol{\eta}_i\right\|^4 \leq \Delta < \infty$ and $\left\|\boldsymbol{\beta}\right\|^4 \leq \Delta < \infty$, and $\boldsymbol{\eta}_i$ is group-wise independent of ε_{it}, $\mathbf{v}_{it}$ and $\mathbf{H}_i^0$.

Assumption B5 (identification and Variance Matrices): $\tilde{\mathbf{A}}_{iT} = T^{-1}E\left(\mathbf{V}_i'\mathbf{V}_i\right)$, $\tilde{\mathbf{B}}_{iT} = T^{-1}E\left(\mathbf{V}_i'\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{V}_i\right)$ and $\tilde{\mathbf{C}}_{iT} = T^{-2}E\left(\mathbf{V}_i'\mathbf{V}_i\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\mathbf{V}_i'\mathbf{V}_i\right)$ are uniformly positive definite and $\tilde{\mathbf{A}} = \lim_{N,T\rightarrow\infty}N^{-1}\sum_{i=1}^{N}\tilde{\mathbf{A}}_{iT}$, $\tilde{\mathbf{B}} = \lim_{N,T\rightarrow\infty}N^{-1}\sum_{i=1}^{N}\tilde{\mathbf{B}}_{iT}$ and $\tilde{\mathbf{C}} = \lim_{N,T\rightarrow\infty}N^{-1}\sum_{i=1}^{N}\tilde{\mathbf{C}}_{iT}$ are fixed and positive definite.

Idiosyncratic errors ε_{it} and $\mathbf{v}_{it}$ are independent groups of each other, independent over i, but allowed to be serially correlated as structured by Assumptions B1 and B2. Assumption B3 implies there are r factors, and the factor loadings $\mathbf{\Gamma}_i^0$ and $\boldsymbol{\lambda}_i^0$ have mean zero without loss of generality and are allowed to be correlated with each other. Assumption B4 implies that the random coefficients can be heteroskedastic but should be independent of all other cross-section varying variables. Assumption B5 corresponds to Assumptions A2-A4 in Section 2.

For any invertible $r \times r$ matrix $\mathbf{R}$, define

$$\mathbf{F} = \mathbf{F}^0\mathbf{R},\ \mathbf{G}_i = \mathbf{R}^{-1}\mathbf{G}_i^0, \tag{23}$$

such that $T^{-1}\mathbf{F}'\mathbf{F} = \mathbf{I}_r$ and $\sum_{i=1}^{N}\mathbf{G}_i\mathbf{G}_i'$ is diagonal. Then, $\mathbf{M}_{\mathbf{F}} = \mathbf{M}_{\mathbf{F}^0}$, so that $\mathbf{M}_{\mathbf{F}}\mathbf{F}^0 = \mathbf{M}_{\mathbf{F}^0}\mathbf{F} = \mathbf{0}$. The solutions to the minimisation problem,

$$arg \min_{\mathbf{F}\subset\mathcal{F},\mathbf{G}_i\subset\mathcal{G}} \frac{1}{NT}\sum_{\ell=1}^{m}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(z_{it\ell} - \mathbf{f}_t'\mathbf{g}_{i\ell}\right)^2,$$

$$\text{subjectto } T^{-1}\mathbf{F}'\mathbf{F} = \mathbf{I}_r \text{ and } \sum_{i=1}^{N}\mathbf{G}_i\mathbf{G}_i' \text{ being diagonal}, \tag{24}$$

with $z_{it\ell}^*$ being $(t,\ell)th$ element of $\mathbf{Z}_i^*$, are given by $\hat{\mathbf{F}}$, which is $\sqrt{T}$ times the eigenvectors corresponding to the r largest eigenvalues of the $T \times T$ matrix $N^{-1}\sum_{i=1}^{N}\mathbf{Z}_i^*\mathbf{Z}_i^{*\prime}$, and $\hat{\mathbf{G}}_i = \hat{\mathbf{F}}'\mathbf{Z}_i^*/T$, thus, $\hat{\mathbf{E}}_i = \mathbf{M}_{\hat{\mathbf{F}}}\mathbf{Z}_i^*$.[11]

[11] In the standard literature, factor loadings and the idiosyncratic errors in factor models are assumed to be independent (e.g. Bai and Ng, 2002). Under slope heterogeneity, due to the presence of $\boldsymbol{\beta}_i$ in $\mathbf{g}_{i1}$ and $\mathbf{e}_{i1}$, they are uncorrelated but not independent. As is shown in Lemma B.3, this does not change the convergence rate of the factor estimators. For example, under both slope heterogeneity and homogeneity, $T^{-1/2}\left\|\hat{\mathbf{F}} - \mathbf{F}^0\mathbf{R}\right\| = O_p\left(\delta_{NT}^{-1}\right)$.

Defining the transformed variables

$$\hat{\mathbf{y}}_i = \mathbf{M}_{\hat{\mathbf{F}}}\mathbf{y}_i^*,\ \hat{\mathbf{X}}_i = \mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*, \tag{25}$$

the pooled estimator is obtained as

$$\hat{\boldsymbol{\beta}}_{PC} = \left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{y}}_i. \tag{26}$$

Noting that $\hat{\mathbf{X}}_i'\hat{\mathbf{y}}_i = \hat{\mathbf{X}}_i'\mathbf{y}_i^*$ and $\mathbf{u}_i = \mathbf{X}_i^*\boldsymbol{\eta}_i + \epsilon_{f\varepsilon,i}$, we have

$$\begin{aligned}\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta} &= \left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\sum_{i=1}^{N}\hat{\mathbf{X}}_i\mathbf{u}_i \\ &= \left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\left[\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*\boldsymbol{\eta}_i + \sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i}\right].\end{aligned} \tag{27}$$

As discussed in Bai (2009) and Westerlund and Urbain (2015), there will be asymptotic bias under homogeneous slopes when $N/T \to c \in (0, \Delta]$. Greenaway-McGrevy et al (2012) consider the same model with serially correlated errors, but do not derive asymptotic bias. We extend the results of Westerlund and Urbain, proposing a bias-correction which is asymptotically justified with both homogeneous slopes and heterogeneous slopes.

Proposition 2 *Consider model (5). Under Assumptions B1-B5, as* $(N,T) \to \infty$ *and* $N/T \to c \in (0, \Delta]$,

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\boldsymbol{\varepsilon}_i + \sqrt{\frac{T}{N}}\boldsymbol{\xi}_{NT} + o_p(1) \tag{28}$$

where

$$\begin{aligned}\boldsymbol{\xi}_{NT} = &-\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{1i}^0\sigma_{\varepsilon i}^2 + \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{G}_j^0\bar{\boldsymbol{\Omega}}_{EE,j}\boldsymbol{G}_j^{0\prime}\right)\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i^0 \\ &-\frac{1}{N}\sum_{i=1}^{N}\bar{\boldsymbol{\Omega}}_{VE,i}\boldsymbol{G}_i^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i^0,\end{aligned} \tag{29}$$

$\boldsymbol{g}_{1i}^0$ *is the first column vector of* $\boldsymbol{G}_i^0$,

$$\sigma_{\varepsilon i}^2 = E\left(T^{-1}\boldsymbol{\varepsilon}_i'\boldsymbol{\varepsilon}_i\right),\ \bar{\boldsymbol{\Omega}}_{EE,i} = E\left(T^{-1}\mathbf{E}_i'\mathbf{E}_i\right),\ \bar{\boldsymbol{\Omega}}_{VE,i} = E\left(T^{-1}\mathbf{V}_i'\mathbf{E}_i\right). \tag{30}$$

Observe that, under slope heterogeneity, $\boldsymbol{G}_i^0$ and $\mathbf{E}_i$ are functions of $\boldsymbol{\beta}_i$. This is the reason why the expression of the bias is different from that in Westerlund and Urbain (2015).[12] Based on Proposition 2, we propose to use the following bias-corrected estimator:

$$\tilde{\boldsymbol{\beta}}_{PC} = \hat{\boldsymbol{\beta}}_{PC} - \frac{1}{N}\hat{\mathbf{c}}_{NT}, \tag{31}$$

[12]It is easily seen that, under slope homogeneity, the bias-term $\boldsymbol{\xi}_{NT}$ and that of Westerlund and Urbain (2015, Theorem 1) are asymptotically the same.

where

$$\hat{\mathbf{c}}_{NT} = \left(\frac{1}{NT} \sum_{i=1}^{N} \hat{\mathbf{X}}_i' \hat{\mathbf{X}}_i \right)^{-1} \hat{\boldsymbol{\xi}}_{NT}, \tag{32}$$

$$\hat{\boldsymbol{\xi}}_{NT} = -\frac{1}{N} \sum_{i=1}^{N} \hat{\boldsymbol{\Gamma}}_i \hat{\boldsymbol{\Upsilon}}_N^{-1} \hat{\boldsymbol{g}}_{1i} \hat{\sigma}_{ui}^{\dagger 2} + \frac{1}{N} \sum_{i=1}^{N} \hat{\boldsymbol{\Gamma}}_i \hat{\boldsymbol{\Upsilon}}_N^{-1} \left(\frac{1}{N} \sum_{j=1}^{N} \hat{\boldsymbol{G}}_j \hat{\boldsymbol{\Omega}}_{EE,j} \hat{\boldsymbol{G}}_j' \right) \hat{\boldsymbol{\Upsilon}}_N^{-1} \hat{\boldsymbol{\lambda}}_i^{\dagger} \tag{33}$$

$$-\frac{1}{N} \sum_{i=1}^{N} \hat{\boldsymbol{\Omega}}_{VE,i} \hat{\boldsymbol{G}}_i' \hat{\boldsymbol{\Upsilon}}_N^{-1} \hat{\boldsymbol{\lambda}}_i^{\dagger}$$

with

$$\hat{\boldsymbol{\Gamma}}_i' = \frac{\hat{\mathbf{F}}' \mathbf{X}_i^*}{T}, \ \hat{\boldsymbol{\Upsilon}}_N = N^{-1} \sum_{i=1}^{N} \hat{\boldsymbol{G}}_i \hat{\boldsymbol{G}}_i', \ \hat{\boldsymbol{G}}_i = \frac{\hat{\mathbf{F}}' \mathbf{Z}_i^*}{T}, \ \hat{\boldsymbol{g}}_{1i} = \frac{\hat{\mathbf{F}}' \mathbf{y}_i^*}{T} \tag{34}$$

$$\hat{\sigma}_{ui}^{\dagger 2} = \frac{\hat{\mathbf{u}}_{PC,i}' \mathbf{M}_{\hat{\mathbf{F}}} \hat{\mathbf{u}}_{PC,i}}{T}, \ \hat{\boldsymbol{\Omega}}_{EE,i} = \frac{\hat{\mathbf{E}}_i' \hat{\mathbf{E}}_i}{T}, \ \hat{\mathbf{E}}_i = \mathbf{M}_{\hat{\mathbf{F}}} \mathbf{Z}_i^* \tag{35}$$

$$\hat{\boldsymbol{\lambda}}_i^{\dagger} = \frac{\hat{\mathbf{F}}' \hat{\mathbf{u}}_{PC,i}}{T}, \ \hat{\boldsymbol{\Omega}}_{VE,i} = \frac{\hat{\mathbf{V}}_i' \hat{\mathbf{E}}_i}{T}, \ \hat{\mathbf{V}}_i = \mathbf{M}_{\hat{\mathbf{F}}} \mathbf{X}_i^*, \tag{36}$$

$\hat{\mathbf{u}}_{PC,i} = \mathbf{y}_i^* - \mathbf{X}_i^* \hat{\boldsymbol{\beta}}_{PC}$. This estimator is different from those proposed by Westerlund and Urbain (2015), to allow slope heterogeneity. Note that the probability limit of estimators with superscript "†", $\hat{\sigma}_{ui}^{\dagger 2}$ and $\hat{\boldsymbol{\lambda}}_i^{\dagger}$, will be different for the slope heterogeneous case, because they are functions of $\mathbf{u}_i = \mathbf{X}_i^{*\prime} \boldsymbol{\eta}_i + \boldsymbol{\epsilon}_{f\varepsilon,i}$. The following proposition shows that $\hat{\mathbf{c}}_{NT}$ is consistent to the bias given by (29) under slope homogeneity, and the limit of $\hat{\mathbf{c}}_{NT}$ remains bounded under slope heterogeneity.

Proposition 3 *Under Assumptions B1-B5, as $(N, T) \to \infty$ and $N/T \to c \in (0, \Delta]$,*

$$\hat{\mathbf{c}}_{NT} - \mathbf{c} \to_p \mathbf{0} \tag{37}$$

where $\mathbf{c} = \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{V}_i' \mathbf{V}_i \right)^{-1} \boldsymbol{\xi}_{NT}^{\dagger}$,

$$\begin{aligned} \boldsymbol{\xi}_{NT}^{\dagger} &= -\frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\Gamma}_i^0 \left(\boldsymbol{\Upsilon}_N^0 \right)^{-1} \boldsymbol{g}_{1i}^0 \sigma_{ui}^2 \\ &+ \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\Gamma}_i^0 \left(\boldsymbol{\Upsilon}_N^0 \right)^{-1} \left(\frac{1}{N} \sum_{j=1}^{N} \boldsymbol{G}_j^0 \bar{\boldsymbol{\Omega}}_{EE,i} \boldsymbol{G}_j^{0\prime} \right) \left(\boldsymbol{\Upsilon}_N^0 \right)^{-1} \left(\boldsymbol{\lambda}_i^0 + \Gamma_i^0 \boldsymbol{\eta}_i \right) \\ &- \frac{1}{N} \sum_{i=1}^{N} \bar{\boldsymbol{\Omega}}_{VE,i} \boldsymbol{G}_i^{0\prime} \left(\boldsymbol{\Upsilon}_N^0 \right)^{-1} \left(\boldsymbol{\lambda}_i^0 + \Gamma_i^0 \boldsymbol{\eta}_i \right) \end{aligned} \tag{38}$$

with

$$\sigma_{ui}^2 = \sigma_{\varepsilon i}^2 + tr \left(\tilde{\mathbf{A}}_{iT} \bar{\boldsymbol{\Omega}}_{\eta\eta,i} \right). \tag{39}$$

Remark 4 *Under slope homogeneity, $\boldsymbol{\eta}_i = \mathbf{0}$ for all i, it is easily seen that $\hat{\boldsymbol{\xi}}_{NT} - \boldsymbol{\xi}_{NT} = o_p(1)$, as $\hat{\boldsymbol{\xi}}_{NT} - \boldsymbol{\xi}_{NT}^{\dagger} = o_p(1)$ and $\boldsymbol{\xi}_{NT}^{\dagger} - \boldsymbol{\xi}_{NT} = o_p(1)$, thus, the limiting distribution of $\sqrt{NT}\left(\tilde{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)$ is centred at zero. Under slope heterogeneity, $\hat{\boldsymbol{\xi}}_{NT} - \boldsymbol{\xi}_{NT}^{\dagger} = o_p(1)$, where $E\left\|\boldsymbol{\xi}_{NT}^{\dagger}\right\|^2$ is bounded, hence, the limiting distribution of $\sqrt{N}\left(\tilde{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)$ is centred at zero.*

Now we are ready to state the asymptotic normality result of the bias corrected estimator under slope homogeneity:

Theorem 4 *Under Assumptions B1-B5, under homogeneous slopes ($\boldsymbol{\eta}_i = \mathbf{0}$ for all i), as $(N,T) \to \infty$ and $N/T \to c \in (0, \Delta]$,*

$$\sqrt{NT}\left(\tilde{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right) \to_d N\left(\mathbf{0}, \tilde{\mathbf{A}}^{-1}\tilde{\mathbf{B}}\tilde{\mathbf{A}}^{-1}\right). \tag{40}$$

where $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ are defined in Assumption B5.

Next, consider the case of slope heterogeneity. Noting that $\mathbf{u}_i = \mathbf{X}_i^{*\prime}\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_{f\varepsilon,i}$, by Proposition 2 and Lemma B.11, we have

$$\begin{aligned} \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{u}_i &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\mathbf{X}_i^{*}}{T}\boldsymbol{\eta}_i + \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} + o_p(1) \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{V}_i}{T}\boldsymbol{\eta}_i + o_p(1). \end{aligned} \tag{41}$$

Together with Proposition 3, the asymptotic normality of the bias-corrected estimator under slope heterogeneity is established in the following theorem:

Theorem 5 *Under Assumptions B1-B5, as $(N,T) \to \infty$ and $N/T \to c \in (0, \Delta]$,*

$$\sqrt{N}\left(\tilde{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right) \to_d N\left(\mathbf{0}, \tilde{\mathbf{A}}^{-1}\tilde{\mathbf{C}}\tilde{\mathbf{A}}^{-1}\right) \tag{42}$$

where $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{C}}$ are defined in Assumption B5.

We propose the heteroskedasticity, autocorrelation and slope heterogeneity robust variance estimator for the model with unobserved interactive effects, which is given by

$$\hat{\mathbf{V}}_{PC} = \left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{u}}_{PC,i}\hat{\mathbf{u}}_{PC,i}'\hat{\mathbf{X}}_i\right)\left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}, \tag{43}$$

where $\hat{\mathbf{u}}_{PC,i} = \mathbf{y}_i^* - \mathbf{X}_i^*\tilde{\boldsymbol{\beta}}_{PC}$. The asymptotic justification of the use of this variance estimator is established in the following theorem.[13]

[13] Replacing $\hat{\mathbf{u}}_i$ in (43) with $\tilde{\mathbf{u}}_i = \mathbf{y}_i^* - \mathbf{X}_i^*\tilde{\boldsymbol{\beta}}_{PC}$ will not alter the results.

Theorem 6 *Consider testing* q *linearly independent restrictions of* $\boldsymbol{\beta}$, $H_0 : \mathcal{R}\boldsymbol{\beta} = \mathbf{r}$ *against* $H_1 : \mathcal{R}\boldsymbol{\beta} \neq \mathbf{r}$, *where* $\mathcal{R}$ *is a* $q \times k$ *fixed matrix of full row rank. Consider the model (4) and the Wald statistic*

$$W_{PC} = \left(\mathcal{R}\tilde{\boldsymbol{\beta}}_{PC} - \mathbf{r}\right)' \left(\mathcal{R}\hat{\mathbf{V}}_{PC}\mathcal{R}'\right)^{-1} \left(\mathcal{R}\tilde{\boldsymbol{\beta}}_{PC} - \mathbf{r}\right), \tag{44}$$

where $\tilde{\boldsymbol{\beta}}_{PC}$ *and* $\hat{\mathbf{V}}_{PC}$ *are defined by (31) and (43), respectively. Suppose that Assumptions B1-B5 hold. Then, under the* H_0, *for both heterogenous slopes and homogeneous slopes* ($\boldsymbol{\eta}_i = \mathbf{0}$ *for all* i), $W_{PC} \rightarrow_d \chi^2_q$, *as* $(N,T) \rightarrow \infty$, *as* $T/N \rightarrow c \in (0, \Delta]$.

Remark 7 *Our approach is also robust against mixtures of homogeneous and heterogeneous slopes.*[14] *To see this, consider the case in which the* k *slopes are partitioned in such a way that* $k = k_1 + k_2$, *without loss of generality, where* $\boldsymbol{\beta}_i = (\boldsymbol{\beta}'_{1i}, \boldsymbol{\beta}'_2)'$, $\boldsymbol{\beta}_{1i} = \boldsymbol{\beta}_1 + \boldsymbol{\eta}_{1i}$, $E(\boldsymbol{\eta}_{1i}) = \mathbf{0}$ *and* $Var(\boldsymbol{\eta}_{1i}) = \boldsymbol{\Omega}_{1i}$, *with* $\boldsymbol{\beta} = (\boldsymbol{\beta}'_1, \boldsymbol{\beta}'_2)'$. *Define a scaling diagonal matrix of order* k *as*

$$\mathbf{D} = \begin{bmatrix} \sqrt{N}\mathbf{I}_{k_1} & \mathbf{0} \\ \mathbf{0} & \sqrt{NT}\mathbf{I}_{k_2} \end{bmatrix}, \tag{45}$$

so that

$$\mathbf{D}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)'\mathbf{D} = \left(\mathbf{D}\bar{\mathbf{A}}^{-1}_{NT}\mathbf{D}^{-1}\right)\left(\mathbf{D}\sum_{i=1}^{N}\frac{\mathbf{X}'_i\mathbf{u}_i\mathbf{u}'_i\mathbf{X}_i}{T^2N^2}\mathbf{D}\right)\left(\mathbf{D}^{-1}\bar{\mathbf{A}}^{-1}_{NT}\mathbf{D}\right). \tag{46}$$

It is easily seen that $\mathbf{D}\bar{\mathbf{A}}^{-1}_{NT}\mathbf{D}^{-1} = \bar{\mathbf{A}}^{-1}_{NT}$. *Recalling that* $\mathbf{u}_i = \mathbf{X}_i\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i$, $\boldsymbol{\eta}_i = (\boldsymbol{\eta}'_{1i}, \mathbf{0}')'$ *and* $E(\boldsymbol{\eta}_i\boldsymbol{\varepsilon}'_i) = \mathbf{0}$, *the probability limit of the middle term is*

$$p\lim_{N,T\rightarrow\infty}\sum_{i=1}^{N}\mathbf{D}\frac{\mathbf{X}'_i\mathbf{u}_i\mathbf{u}'_i\mathbf{X}_i}{T^2N^2}\mathbf{D} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{22} \end{pmatrix}, \tag{47}$$

where

$$\mathbf{C}_{11} = p\lim_{N,T\rightarrow\infty} N^{-1}\sum_{i=1}^{N}\left(\frac{\mathbf{X}'_{1i}\mathbf{X}_{1i}}{T}\boldsymbol{\eta}_{1i}\boldsymbol{\eta}'_{1i}\frac{\mathbf{X}'_{1i}\mathbf{X}_{1i}}{T}\right), \tag{48}$$

$$\mathbf{C}_{22} = p\lim_{N,T\rightarrow\infty} N^{-1}\sum_{i=1}^{N}\frac{\mathbf{X}'_{2i}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}'_i\mathbf{X}_{2i}}{T}. \tag{49}$$

Therefore, the asymptotic normality of $\mathbf{D}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)$, *the consistency of the HAC estimator and the asymptotic validity of Wald test hold with mixtures of homogeneous and heterogeneous slopes.*

[14] We do not consider cross-sectional and/or time-series structural breaks in $\boldsymbol{\beta}_i$ which is beyond the scope of this paper.

4 Wald and LM tests for Correlation of Random Coefficients with Covariates

As discussed earlier, the proposed robust approach works for random coefficients. If it is fixed cross-sectionlly varying coefficients or correlated random coefficients with $\mathbf{X}_i^*$, the approach may not work. To see this, consider the model (5) but without factor components. We have $\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = \left(\sum_{i=1}^N \mathbf{X}_i^{*\prime}\mathbf{X}_i^*\right)^{-1} \sum_{i=1}^N \mathbf{X}_i^{*\prime}\left[\mathbf{X}_i^*\boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i\right]$. If $E\left(\boldsymbol{\eta}_i|\mathbf{X}_i^*\right) \neq \mathbf{0}$, $E\left[\mathbf{X}_i^{*\prime}\mathbf{X}_i^* E\left(\boldsymbol{\eta}_i|\mathbf{X}_i^*\right)\right]$ is not necessarily zero, and in general it renders $\hat{\boldsymbol{\beta}}$ biased.

In view of this, we propose novel tests for correlation or dependence of random coefficients with covariates, substantially extending the test proposed by Wooledridge (2010; Ch11.7.4). The main distinctions of our tests from Wooledridge's are: (i) our tests are robust against (uncorrelated) random coefficients;[15] (ii) we propose a Lagrange Multiplier test along with a Wald test; (iii) ours permit $E\left(\boldsymbol{\eta}_i|\mathbf{X}_i^*\right)$ to be a non-linear function of $\mathbf{X}_i^*$.

More generally, suppose that the random part of the coefficients is modeled as

$$\boldsymbol{\eta}_i = \mathbf{h}\left(\mathbf{X}_i^*\right) - \boldsymbol{\mu}_{\mathbf{h}} + \boldsymbol{\zeta}_i \tag{50}$$

with $E\left[\mathbf{h}\left(\mathbf{X}_i^*\right)\right] = \boldsymbol{\mu}_{\mathbf{h}}$ and $E\left(\boldsymbol{\zeta}_i|\mathbf{X}_i^*\right) = \mathbf{0}$, where various forms of function of $\mathbf{X}_i^*$ can be entertained. For the testing purpose, we consider $\mathbf{h}\left(\mathbf{X}_i^*\right) = \boldsymbol{\Xi}_i\boldsymbol{\delta}$ with

$$\boldsymbol{\Xi}_i = \left(\bar{\mathbf{x}}_i^{(1)}; \bar{\mathbf{x}}_i^{(2)}; ..., \bar{\mathbf{x}}_i^{(g)}\right), \tag{51}$$

$\bar{\mathbf{x}}_i^{(g)} = \left(\bar{x}_{i1}^{(g)}, \bar{x}_{i2}^{(g)}, ..., \bar{x}_{ik}^{(g)}\right)'$, $\bar{x}_{ih}^{(g)} = T^{-1}\sum_{t=1}^T x_{ith}^g$.[16] Note that x_{ith} is the (t,h) element of the defactored regressor, $\mathbf{X}_i = \mathbf{M_F}\mathbf{X}_i^*$.[17] Initially assuming that $\mathbf{F}^0$ is observable, consider an augmented regression

$$\mathbf{y}_i = \mathbf{W}_i\boldsymbol{\theta} + \boldsymbol{\epsilon}_i, \tag{52}$$

where $\mathbf{W}_i = \left[\mathbf{X}_i, \mathbf{L}_i\right]$ with

$$\mathbf{L}_i = \mathbf{X}_i\left(\boldsymbol{\Xi}_i - \bar{\boldsymbol{\Xi}}\right), \tag{53}$$

$\bar{\boldsymbol{\Xi}} = N^{-1}\sum_{i=1}^N \boldsymbol{\Xi}_i$, $\boldsymbol{\theta} = \left(\boldsymbol{\beta}', \boldsymbol{\delta}'\right)'$, and the associated unrestricted estimator $\hat{\boldsymbol{\theta}} = \left(\widehat{\hat{\boldsymbol{\beta}}}', \hat{\boldsymbol{\delta}}\right)' = \left(\mathbf{W}_i'\mathbf{W}_i\right)^{-1}\mathbf{W}_i'\mathbf{y}_i$. Under the null hypothesis of $H_0 : \boldsymbol{\delta} = \mathbf{0}$ and Assumptions A1-A4, for homogeneous or heterogeneous slopes, Theorem 3 establishes that

$$W_{CRC}^{(g)} = \hat{\boldsymbol{\delta}}'\hat{\mathbf{V}}_{\delta\delta}^{-1}\hat{\boldsymbol{\delta}} \to_d \chi_g^2 \tag{54}$$

as $(N,T) \to \infty$, where $\hat{\mathbf{V}}_{\delta\delta}$ is defined as the bottom right partition of $\hat{\mathbf{V}}_{NT}\left(\hat{\boldsymbol{\theta}}\right) = \begin{pmatrix} \hat{\mathbf{V}}_{\beta\beta} & \hat{\mathbf{V}}_{\beta\delta} \\ \hat{\mathbf{V}}_{\delta\beta} & \hat{\mathbf{V}}_{\delta\delta} \end{pmatrix} = \left(\sum_{i=1}^N \mathbf{W}_i'\mathbf{W}_i\right)^{-1}\left(\sum_{i=1}^N \mathbf{W}_i'\hat{\boldsymbol{\epsilon}}_i\hat{\boldsymbol{\epsilon}}_i'\mathbf{W}_i\right)\left(\sum_{i=1}^N \mathbf{W}_i'\mathbf{W}_i\right)^{-1}$, $\hat{\boldsymbol{\epsilon}}_i = \mathbf{y}_i - \mathbf{W}_i\hat{\boldsymbol{\theta}}$.

[15] Wooldridge (2010;p.386) points out that the drawback of his test is that it cannot detect heterogeneity in $\boldsymbol{\beta}_i$ that is uncorrelated with $\bar{\mathbf{x}}_i$. In our robustified test, this becomes the desirable property.

[16] Cross product terms, such as $T^{-1}\sum_{t=1}^T x_{ith}^{(g)}x_{itj}^{(f)}$ for $h \neq j$, could be included in $\boldsymbol{\Xi}_i$.

[17] For the model with fixed effects, the test variable $\boldsymbol{\Xi}_i$ should not be based on within-transformed $\mathbf{X}_i^*$, otherwise $\bar{\mathbf{x}}_i^{(1)} = \mathbf{0}$ for all i.

For the estimated factor case, the test statistic is computed based on $\left(\hat{\mathbf{y}}_i, \hat{\mathbf{X}}_i\right)$, $\hat{\mathbf{W}}_i = [\hat{\mathbf{X}}_i, \hat{\mathbf{L}}_i]$ with $\hat{\mathbf{L}}_i = \hat{\mathbf{X}}_i\left(\hat{\boldsymbol{\Xi}}_i - \overline{\hat{\boldsymbol{\Xi}}}\right)$, $\hat{\boldsymbol{\Xi}}_i$ based on $\hat{\boldsymbol{X}}_i$, and $\tilde{\boldsymbol{\theta}}_{PC} = \left(\tilde{\hat{\boldsymbol{\beta}}}'_{PC}, \tilde{\delta}_{PC}\right)' = \left(\hat{\mathbf{W}}_i'\hat{\mathbf{W}}_i\right)^{-1}\hat{\mathbf{W}}_i'\hat{\mathbf{y}}_i$, which is the bias-corrected PC estimator discussed in Section 3.

We also consider the Lagrange Multiplier (LM) or Score test of the correlated random coefficient. One of the advantages of employing the LM test is that, unlike the Wald test, computation of the LM test only requires the estimation results of the null model. The LM test statistic with observable factors is defined as

$$LM_{CRC}^{(g)} = \left(\sum_{i=1}^{N}\mathbf{L}_i'\hat{\mathbf{u}}_i\right)'\left(\sum_{i=1}^{N}\mathbf{K}_i'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\mathbf{K}_i\right)^{-1}\left(\sum_{i=1}^{N}\mathbf{L}_i'\hat{\mathbf{u}}_i\right) \tag{55}$$

where $\hat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}$ with $\hat{\boldsymbol{\beta}} = \left(\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\right)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{y}_i$ and

$$\mathbf{K}_i' = \mathbf{L}_i' - \left(\sum_{i=1}^{N}\mathbf{L}_i'\mathbf{X}_i\right)\left(\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\right)^{-1}\mathbf{X}_i'. \tag{56}$$

For the PC estimator, the LM test statistic is given by

$$LM_{CRC_{PC}}^{(g)} = \left(\sum_{i=1}^{N}\hat{\mathbf{L}}_i'\tilde{\mathbf{u}}_{PC,i}\right)'\left(\sum_{i=1}^{N}\hat{\mathbf{K}}_i'\hat{\mathbf{u}}_{PC,i}\hat{\mathbf{u}}_{PC,i}'\hat{\mathbf{K}}_i\right)^{-1}\left(\sum_{i=1}^{N}\hat{\mathbf{L}}_i'\tilde{\mathbf{u}}_{PC,i}\right) \tag{57}$$

where $\hat{\mathbf{u}}_{PC,i} = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}_{PC}$ with $\hat{\boldsymbol{\beta}}_{PC} = \left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{y}}_i$, $\tilde{\mathbf{u}}_{PC,i} = \mathbf{y}_i - \mathbf{X}_i\tilde{\boldsymbol{\beta}}_{PC}$ with $\tilde{\boldsymbol{\beta}}_{PC}$ being the bias corrected estimator, and

$$\hat{\mathbf{K}}_i' = \hat{\mathbf{L}}_i' - \left(\sum_{i=1}^{N}\hat{\mathbf{L}}_i'\hat{\mathbf{X}}_i\right)\left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\hat{\mathbf{X}}_i'. \tag{58}$$

By the standard discussion of asymptotic equivalence of the LM and Wald tests, it is readily established that under the null hypothesis $LM_{CRC} \to_d \chi_g^2$ as $(N,T) \to \infty$, and $LM_{CRC_{PC}} \to_d \chi_g^2$ as $(N,T) \to \infty$ such that $N/T \to c \in (0, \Delta]$. It may be sufficient to consider $g = 2$ to approximate the function $\mathbf{g}\left(\mathbf{X}_i^*\right)$ for our testing purpose.

When the test is rejected in favour of alternatives, it is preferable to employ estimators which are consistent when variation of $\boldsymbol{\beta}_i$ is dependent on covariates. For the estimation of the models with observed factors, the mean group estimator proposed by Chamberlain (1982) and Pesaran and Smith (1995) would be possible choices.

5 Monte Carlo Experiments

In this section we investigate the finite sample performance of our robust approach against slope heterogeneity, error serial correlation and heteroskedasticity. We consider the performance of the following estimators: (two-way) fixed effects estimator $\hat{\boldsymbol{\beta}}_{FE}$, which is the pooled ordinary least square (OLS) estimator of within-transformed and cross-sectionally demeaned variables; the bias-non-corrected PC estimator $\hat{\boldsymbol{\beta}}_{PC}$ (defined by (26)) and its bias-corrected version $\tilde{\boldsymbol{\beta}}_{PC}$ (defined by (31)); Bai's (2009) iterative PC estimator, both bias-non-corrected $\hat{\boldsymbol{\beta}}_{Bai}$ and the bias-corrected estimator $\tilde{\boldsymbol{\beta}}_{Bai}$.[18] Bai's estimator does not require the linear factor structure in $\mathbf{X}_i^*$ unlike the PC estimator, and its algorithm estimates $\boldsymbol{\beta}$ and $\mathbf{F}$ iteratively from the residual $\mathbf{u}_i$, given the initial value of $\boldsymbol{\beta}$. This generality results in additional bias terms. In all the experiments, we assume that the number of factors r is known.[19]

In particular, we examine bias and root mean square errors (RMSE) of the estimators, and empirical size and power of the (Wald) test for linear restrictions of slope coefficients, as well as the performance of the LM test for correlation and dependence of slope coefficients with covariates.[20]

5.1 Design

Consider the following data generating process:

$$y_{it}^* = \sum_{h=1}^{k} x_{ith}^* \beta_{ih} + \sum_{\ell=1}^{r} f_{t\ell}\lambda_{i\ell} + \sigma_{\varepsilon,it}\varepsilon_{it}, \; i = 1, 2, ..., N; t = 1, 2, ..., T \tag{59}$$

where $\lambda_{i\ell} \sim iidN(0,1)$, $f_{t\ell} = \rho_f f_{t-1,\ell} + \sqrt{1-\rho_f^2}\nu_{t\ell}$, $\nu_{t\ell} \sim iidN(0,1)$ with $f_{0,\ell} \sim iidN(0,1)$ for $\ell = 1,..,r$, $\varepsilon_{it} = \rho_\varepsilon \varepsilon_{it-1} + \sqrt{1-\rho_\varepsilon^2}\xi_{it}$, $\xi_{it} \sim iidN(0,1)$ with $\varepsilon_{i0} \sim iidN(0,1)$, and

$$\sigma_{\varepsilon,it} = (\kappa_{\varepsilon,i}\kappa_{\varepsilon,t})^{1/2}, \; \kappa_{\varepsilon,i} \sim iidU(0.5, 1.5) \text{ and } \kappa_{\varepsilon,t} = 0.5 + t/T. \tag{60}$$

The regressors x_{ith}, $h = 1, 2, .., k$, are generated as

$$x_{ith}^* = \sum_{\ell=1}^{r} f_{t\ell}\gamma_{ih\ell} + \phi\sigma_{\mathrm{v},it}\mathrm{v}_{ith}, \tag{61}$$

where $\mathrm{v}_{ith} = \rho_{\mathrm{v}}\mathrm{v}_{it-1,h} + \sqrt{1-\rho_{\mathrm{v}}^2}\varpi_{it,h}$. We consider two types of distribution for $\varpi_{it,h}$: (i) $\varpi_{it,h} = \left(\varpi_{it,h}^* - c\right)/\sqrt{2c}$, $\varpi_{it,h}^* \sim iid\chi_c^2$ and $\mathrm{v}_{i0,h} = \left(\mathrm{v}_{i0,h}^* - c\right)/\sqrt{2c}$, $\mathrm{v}_{i0,h}^* \sim iid\chi_c^2$ with $c = 6$, and (ii) $\varpi_{it,h} \sim iidN(0,1)$ with $\mathrm{v}_{i0,h} \sim iidN(0,1)$. The factor loadings in x_{ith}^* are generated as

$$\gamma_{ih\ell} = 0.7\lambda_{i\ell} + \left(1 - 0.7^2\right)^{1/2}\varphi_{ih\ell}, \tag{62}$$

$\varphi_{ih\ell} \sim iidN(0,1)$ for $h = 1,..,k$ and $\ell = 1,..,r$, so that they are correlated with factor loadings in y_{it}^*.

$$\sigma_{\mathrm{v},it} = (\kappa_{\mathrm{v},i}\kappa_{\mathrm{v},t})^{1/2}, \; \kappa_{\mathrm{v},i} \sim iidU(0.5, 1.5) \text{ and } \kappa_{\mathrm{v},t} = 0.5 + t/T, \tag{63}$$

[18]See Appendix D for the definition of $\hat{\boldsymbol{\beta}}_{Bai}$ and $\tilde{\boldsymbol{\beta}}_{Bai}$. We take the error-serial correlation into our consideration for the bias correction.

[19]The Pesaran's (2006) CCE estimator is not considered in our experiments, since, to our knowledge, feasible analytical bias correction for the pooled estimator under slope homogeneity is not available.

[20]The finite sample performance of the Wald version of the correlated random effects test is much worse than the LM test version. Therefore, its summary results are not reported.

and $\phi^2 = \{2, 3\}$. Finally we have

$$\beta_{ih} = \beta_h + \sigma_\eta \left(\sqrt{1 - \rho_{x\eta}^2} \eta_{ih} + \rho_{x\eta} w_{ih} \right), \tag{64}$$

$\eta_{ih} \sim iidN(0, 1)$ for $h = 1, .., k$, and

$$w_{ih} = \frac{1}{\sqrt{q}} \sum_{p=1}^{q} \frac{z_{ih,p} - \bar{z}_{h,p}}{s_{zh,p}}, \tag{65}$$

where $\bar{z}_{h,p} = N^{-1} \sum_{i=1}^{N} z_{ih,p}$, $s_{zh,p}^2 = (N-1)^{-1} \sum_{i=1}^{N} (z_{ih,p} - \bar{z}_{h,p})^2$. We consider $z_{ih,p} = T^{-1} \sum_{t=1}^{T} (x_{ith}^*)^p$.

We set $k = 2$ (two regressors) for all the experiments. We consider two sets of design: the model without factors ($r = 0$) to examine the fixed effects estimator, and the model with two factors ($r = 2$) to examine the PC and Bai's estimators. As recommended in Remark 1, before the estimation the data is all within transformed and cross-sectionally demeaned, to make the results invariant to the inclusion of (additive) individual effects and time effects.

In view of the sensitiveness of the finite sample behaviour of the PC estimator to the parameter values (β_1, β_2),[21] we consider three combinations of (β_1, β_2): $(1, 3)$, $(0, 0)$ and $(-1, -3)$.[22]

To look into the bias and RMSE of the estimators, and the size and power of the test of linear restrictions for the estimators, we consider the following sets of designs:

(A) homogeneous slopes ($\sigma_\eta = 0$ in (64));

(B) heterogeneous slopes ($\sigma_\eta = 0.2$ in (64)).

In order to see the effects of dependence of β_i with the regressors upon the bias of the estimators and the associated tests, we set $\rho_{x\eta} = 0.5$ in (64). To investigate the effects of the symmetry of the distribution upon the performance of the estimators and the tests, we consider two types of distribution of disturbances in x_{ith} :

(C) $\left(\varpi_{it,h}^* - 6\right) / \sqrt{12}$, $\varpi_{it,h}^* \sim iid\chi_6^2$, with $\rho_{x\eta} = 0.5$

(D) $\varpi_{it,h} \sim iidN(0, 1)$, with $\rho_{x\eta} = 0.5$.

For designs (C) and (D), we consider two types of dependence of β_{ih} upon regressors: β_{ih} is a linear function of the following cross-sectionally standardised values: (i) $T^{-1} \sum_{t=1}^{T} (x_{ith}^*)$ (i.e., $q = 1$ and $p = 1$ in (65)) and (ii) $T^{-1} \sum_{t=1}^{T} (x_{ith}^*)^2$ (i.e., $q = 1$ and $p = 2$ in (65)).

Finally, the size and the power of the LM tests with degrees $g = 1, 2$, are examined as the set **(E)**. The empirical size is obtained using designs (A) and (B), and the empirical power is computed by designs (C) and (D).

We consider all the combinations of $N = 50, 100, 200$ and $T = 25, 50, 100, 200$. Throughout the experiments, we set $\rho_f = 0.5$, $\rho_\varepsilon = 0.5$ and $\rho_v = 0.5$. To save space, we report the results with $\phi^2 = 2$ only.[23] All the tests are conducted at the five per cent significance level. All the experimental results are based on 2,000 replications.

[21] See discussions in Westerlund and Urbain (2015) for more details.

[22] As the FE estimator is much less sensitive to the change of (β_1, β_2), the results for $(1, 3)$ are only examined and reported.

[23] The results with $\phi^2 = 3$ are qualitatively very similar to those with $\phi^2 = 2$, which are available upon request from the authors.

5.2 Results

Table 1 summarises the performance of the Fixed Effect estimator for the model of $(\beta_1, \beta_2) = (1, 3)$, with time-series and cross-section heteroskedastic, serially correlated errors in the absence of interactive effects. Panels A reports the bias, the root mean square error (RMSE) of estimates of β_1, and the size of the Wald test for $H_0 : \beta_1 = 1$ and the power for $H_0 : \beta_1 = 0.95$, under homogeneous slopes, and Panel B under heterogeneous random slopes. The results for β_2 are qualitatively similar and not reported. As predicted by the theory, the Wald test based on the HAC variance estimator has correct size both under slope homogeneity and heterogeneity. Panels C&D report the bias of the estimates and the size of the Wald test for $H_0 : \beta_1 = 1$, to see the effects of dependence between random coefficients and regressors. In Panel C the regressors are generated by asymmetric disturbances and in Panel D, they are drawn from symmetric distribution. In Panel C, when $\boldsymbol{\eta}_i$ depends on $\sum_{t=1}^{T} x_{ith}^*$, the fixed effects estimator exhibits systematic bias, but in Panel D, it does not. This is because when the third moment of x_{ith}^* is zero, by construction $E\left[\mathbf{X}_i^{*\prime}\mathbf{X}_i^*\boldsymbol{\eta}_i\right] = \mathbf{0}$ which makes the estimator unbiased. However, as can be seen in Panel D, the size of the test declines systematically as sample size rises, which suggests that the HAC variance estimates will not be consistent. When $\boldsymbol{\eta}_i$ is a linear function of $\sum_{t=1}^{T}(x_{ith}^*)^2$, regardless of the shape of the distribution of regressors, it exhibits serious bias in estimates (see Panels C&D). Therefore, it is of great importance to statistically check the evidence of dependence of $\boldsymbol{\beta}_i$ with regressors. The performance of the proposed LM test for correlation and dependence of random coefficients with regressors is summarised in Panel E. As can be seen, it has correct size with slope homogeneity and random coefficients, and the LM test with $g = 2$ has high power against both types of dependence of $\boldsymbol{\beta}_i$, $\sum_{t=1}^{T} x_{ith}^*$ and $\sum_{t=1}^{T}\left(x_{ith}^*\right)^2$, whilst the LM test with $g = 1$ lacks power when $\boldsymbol{\beta}_i$ depends on $\sum_{t=1}^{T}\left(x_{ith}^*\right)^2$ only. Therefore, it is recommended to employ $g = 2$ in practice.

Let us turn our attention to the estimation of the models with unobservable interactive effects. The relevant results are reported in Tables 2-4. Each table contains Panels A-E, which correspond to the panels in Table 1. Tables 2-4 employ different parameter values of (β_1, β_2). Table 2 summarises the results for $(\beta_1, \beta_2) = (1, 3)$, Table 3 for $(-1, -3)$ and Table 4 for $(0, 0)$. To illustrate the effectiveness of the bias-correction, we report the results both for bias-non-corrected and bias-corrected estimators.

Consider Panel A of Table 1, which deals with the slope homogeneous case. First look at the bias of the estimators. Non-bias-corrected estimator $(\hat{\beta}_{Bai})$ has very little bias and the magnitude of correction is very small. As reported in Bai (2009), the bias-corrected estimator $(\tilde{\beta}_{Bai})$ has very small bias and it becomes smaller as N and/or T rise. On the other hand, the bias-non-corrected PC estimator $(\hat{\beta}_{PC})$ has a larger magnitude of bias, in line with the results reported in Appendix A of Westerlund and Urbain (2015). Nonetheless, the bias-corrected estimator $(\tilde{\beta}_{PC})$ successfully reduces the bias. In terms of RMSE, $\tilde{\beta}_{Bai}$ and $\tilde{\beta}_{PC}$ are very similar for all the combinations of (N, T). The size of the Wald test based on $\hat{\beta}_{Bai}$ and $\tilde{\beta}_{Bai}$ is close to nominal level. Due to the bias, the size of the test based on $\hat{\beta}_{PC}$ has moderate size distortion, which is successfully corrected by the bias-correction - the size of the test based on $\tilde{\beta}_{PC}$ is much closer to the nominal level.

Now let us turn our attention to the random coefficient model, the results of which are summarised in Panel B, Table 2. The magnitude of the bias of the estimators under slope heterogeneity is larger than under slope homogeneity, especially with small N and T, but it gets smaller as N and T increase. As in the homogeneous slope case, the bias of both $\hat{\beta}_{Bai}$ and $\tilde{\beta}_{Bai}$ is relatively small, whilst the bias of $\hat{\beta}_{PC}$ is much larger than that of $\hat{\beta}_{Bai}$. Interestingly, the bias-corrected PC estimator successfully reduces the bias of $\hat{\beta}_{PC}$ for heterogeneous slope models as well (see Panel B of Tables 2-4). In general, the reduction of the bias increases the variation of the estimator, and when the bias in $\hat{\beta}_{PC}$ is relatively small, in terms of RMSE, the performances of $\hat{\beta}_{PC}$ and $\tilde{\beta}_{PC}$ are very similar. This slight loss of efficiency is revealed in the power comparison of the Wald test. When both Wald tests based on $\hat{\beta}_{PC}$ ($\hat{\beta}_{Bai}$) and $\tilde{\beta}_{PC}$ ($\tilde{\beta}_{Bai}$) have correct size, the power of the test based on $\hat{\beta}_{PC}$ ($\hat{\beta}_{Bai}$) is marginally higher than that based on $\tilde{\beta}_{PC}$ ($\tilde{\beta}_{Bai}$). However, when the bias in $\hat{\beta}_{PC}$ is relatively large (see, for example, Panel B of Table 3), as the bias-correction is very effective, the RMSE of $\tilde{\beta}_{PC}$ becomes much smaller than that of $\hat{\beta}_{PC}$. Furthermore, due to the finite sample bias of $\hat{\beta}_{PC}$, there can be severe size distortion in Wald tests (see, for example, Panel B of Table 3). In addition, the power based on the PC estimator tends to be higher than that of Bai's estimator. This is likely due to the fact that the PC estimator uses information by exploiting the factor structure in regressors, whilst Bai's estimator does not. The properties of the results reported in Panels C, D and E are very similar to those commented earlier on the corresponding panels in Table 1.

Finally, we comment on the sensitivity of the estimators to the variation of the parameter values.[24] It is revealed by our simulation that Bai's estimates are invariant to the changes of the parameter values of (β_1, β_2). Namely, the results related to Bai's estimators in Tables 2-4 are numerically identical. On the other hand, comparing the results in Panels A and B of Tables 2-4, it can be seen that the bias of $\hat{\beta}_{PC}$ is sensitive to the values of (β_1, β_2). The bias of $\hat{\beta}_{PC}$ is positive in Table 2 with $(\beta_1, \beta_2) = (1, 3)$ and in Table 3 with $(-1, -3)$, while the bias in Table 4 $(0, 0)$ is negative. Note that in the case of $(\beta_1, \beta_2) = (-1, -3)$ the bias of $\hat{\boldsymbol{\beta}}_{PC}$ is relatively large in magnitude and it requires a larger sample size for the bias-corrected estimator to satisfactorily reduce it. To sum up, the proposed bias-correction of the PC estimator is quite effective for both slope homogeneous and heterogeneous cases, and the efficacy of $\tilde{\beta}_{PC}$ is mostly comparable to that of $\tilde{\beta}_{Bai}$. However, in view of the sensitivity of the finite sample performance of the PC estimates to the (centred) value of slope coefficients, the proposed robust approach based on Bai's (2009) estimator might be preferred in practice.

6 Concluding Remarks

In this paper, we have proposed a robust approach against heteroskedasticity, error serial correlation and slope heterogeneity for large linear panel data models. First, we have established the asymptotic validity of the Wald test based on the panel HAC variance estimator of the pooled estimator under random coefficient models. Then, we have shown that a similar result holds with the proposed bias-corrected principal component-based pooled estimators for models with unobserved interactive effects. Our new theoretical

[24] Such sensitivities are reported in Westerlund and Urbain (2015) for the PC estimator. We note two points. First, for all the design employed in Westerlund and Urbain, the mean of the factor loadings is well away from zero, which could exaggerate the magnitude of the bias of the estimator. For the PC approach, we recommend to within-transform and cross-sectionally demean the data before the estimation, which would lessen such sensitivities.

result has justified the use of the same slope estimator and the variance estimator, both for slope homogeneous and heterogeneous models. This robust approach can significantly reduce the model selection uncertainty for applied researchers.

In addition, we have proposed a novel test for correlation and dependence of the random coefficient with covariates. The test is of great importance, since the widely used estimators and/or its variance estimators can become inconsistent when the variation of coefficients depends on covariates, in general.

We have examined the finite sample performance of the estimators, tests of linear restrictions, and the LM tests for correlated random coefficients. The evidence illustrates the usefulness of our approach. In particular, for the estimation of the models with unobserved interactive effects, the size of the proposed robust Wald test using the bias-corrected PC estimators and Bai's (2009) estimator is very close to the nominal level, under both slope homogeneity and slope heterogeneity, while maintaining satisfactory power. Also, the LM tests for correlated random coefficients have correct size under both slope homogeneity and slope heterogeneity due to pure random coefficients, while exhibiting high power when the random coefficients depend on covariates. In view of the sensitivity of the finite sample performance of the PC estimates to the (centred) value of slope coefficients, the proposed robust approach based on Bai's (2009) estimator might be preferred in practice.

As emphasised in the paper, when the test of correlated random coefficient rejects the null in favour of alternatives, it is preferable to employ estimators which are consistent when variation of slopes is dependent on covariates. For the estimation of the models with observed factors, the mean group estimator proposed by Chamberlain (1982) and Pesaran and Smith (1995) would be possible choices. For the estimation of the models with unobserved factors, to our knowledge, no satisfactory alternative estimators have been proposed in the literature. Thus, developing such an estimator will be an important future research theme.

Table 1: Summary results of Fixed Effects estimator for the model with $\{\beta_1, \beta_2\} = \{1, 3\}$, heteroskedastic and serially correlated errors

Panel A: Homogeneous Slopes, $\beta_{ih} = \beta_h$ for all i, $h = 1, 2$												
for β_1	Bias (x100)			RMSE (x100)			Size			Power		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{FE}$												
25	-0.148	-0.115	-0.005	2.572	1.848	1.276	5.5	5.9	4.9	47.4	76.1	96.9
50	-0.077	-0.040	0.008	1.853	1.305	0.920	6.0	5.4	5.2	76.0	96.6	100.0
100	-0.061	-0.015	-0.001	1.372	0.954	0.674	5.9	5.7	5.4	95.5	100.0	100.0
200	-0.017	0.004	0.007	0.955	0.677	0.479	5.4	6.1	5.3	100.0	100.0	100.0
Panel B: Heterogeneous Slopes, $\beta_{ih} = \beta_h + \eta_{ih}$ with $\eta_{ih} \sim iidN\,(0, 0.04)$ for all i, $h = 1, 2$												
for β_1	Bias (x100)			RMSE (x100)			Size			Power		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{FE}$												
25	-0.038	-0.064	0.045	4.122	2.966	2.146	5.6	5.5	5.4	23.0	39.3	66.6
50	-0.015	-0.010	0.057	3.679	2.592	1.869	6.6	5.4	5.1	30.2	50.5	78.6
100	-0.031	0.009	0.039	3.327	2.328	1.661	5.7	4.6	4.8	33.0	57.3	85.2
200	0.025	0.037	0.050	3.129	2.194	1.562	6.1	5.5	4.9	36.5	63.2	90.0

Notes for Panels A and B: Data is generated as $y^*_{it} = x^*_{it,1}\beta_{i1} + x^*_{it,2}\beta_{i2} + \sigma_{\varepsilon,it}\varepsilon_{it}$, $i = 1, ..., N$, $t = 1, ..., T$, $\varepsilon_{it} = \rho_\varepsilon \varepsilon_{it-1} + \sqrt{1 - \rho_\varepsilon^2}\xi_{it}$, $\xi_{it} \sim iidN(0, 1)$ with $\varepsilon_{i0} \sim iidN(0, 1)$, $\sigma_{\varepsilon,it} = (\kappa_{\varepsilon,i}\kappa_{\varepsilon,t})^{1/2}$, $\kappa_{\varepsilon,i} \sim iidU\,(0.5, 1.5)$ and $\kappa_{\varepsilon,t} = 0.5 + t/T$; $x^*_{it,h} = \phi\sigma_{\text{v},it}\text{v}_{it,h}$, where $\text{v}_{it,h} = \rho_\text{v}\text{v}_{it-1,h} + \sqrt{1 - \rho_\text{v}^2}\varpi_{it,h}$, $\varpi_{it,h} \sim iid\left(\chi^2_6 - 6\right)/\sqrt{12}$ with $\text{v}_{i0,h} \sim iid\left(\chi^2_6 - 6\right)/\sqrt{12}$, $\sigma_{\text{v},it} = (\kappa_{\text{v},i}\kappa_{\text{v},t})^{1/2}$, $\kappa_{\text{v},i} \sim iidU\,(0.5, 1.5)$ and $\kappa_{\text{v},t} = 0.5 + t/T$. We set $\rho_\varepsilon = \rho_\text{v} = 0.5$ and $\phi^2 = 2$. $\hat{\beta}_{FE}$ is the pooled regression of within-transformed and cross-sectionally demeaned variables. The size is rejection frequency of the proposed Wald test (defined by (19)) for $H_0 : \beta_1 = 1$ and the power for $H_0 : \beta_1 = 0.95$, based on the 5% level test. All results are based on 2000 replications.

Table 1 continued

Panel C: Correlated Heterogeneous Slopes, $\rho_{x\eta}=0.5$, x^*_{ith} generated using χ^2_6												
	(i) β_{ih} is function of $\sum_t x^*_{ith}$						(ii) β_{ih} is function of $\sum_t (x^*_{ith})^2$					
for β_1	Bias (x100)			Size			Bias (x100)			Size		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{FE}$												
25	0.068	0.019	0.147	5.2	5.0	4.4	1.145	1.158	1.273	6.0	6.9	8.5
50	0.107	0.088	0.164	5.7	4.6	3.3	1.208	1.247	1.315	7.8	6.8	9.4
100	0.075	0.099	0.153	5.5	3.6	3.8	1.236	1.317	1.350	7.7	7.6	9.9
200	0.151	0.145	0.188	5.2	4.3	3.2	1.398	1.450	1.473	7.4	8.9	12.7

Panel D: Correlated Heterogeneous Slopes, $\rho_{x\eta}=0.5$, x^*_{ith} generated using $N(0,1)$												
	(i) β_{ih} is function of $\sum_t x^*_{ith}$						(ii) β_{ih} is function of $\sum_t (x^*_{ith})^2$					
for β_1	Bias (x100)			Size			Bias (x100)			Size		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{FE}$												
25	0.046	0.022	0.031	6.6	5.2	4.7	1.213	1.185	1.205	7.6	7.4	7.6
50	-0.045	-0.008	-0.016	6.4	4.6	3.5	1.134	1.168	1.183	6.8	6.2	8.5
100	-0.033	-0.021	-0.015	5.7	4.7	2.9	1.192	1.194	1.222	6.1	7.6	8.7
200	-0.062	-0.049	-0.043	6.0	4.6	3.3	1.235	1.236	1.266	7.1	7.4	9.8

Notes for Panels C and D: The data generating process (DGP) is the same as that for Panel B, except $\beta_{ih}=\beta_h+\sigma_\eta\left(\sqrt{1-\rho^2_{x\eta}}\eta_{ih}+\rho_{x\eta}w_{ih}\right)$, $\eta_{ih}\sim iidN(0,1)$ for $h=1,2$, $w_{ih}=\frac{z_{ih,p}-\bar{z}_{h,p}}{s_{zh,p}}$, where $\bar{z}_{h,p}=N^{-1}\sum_{i=1}^N z_{ih,p}$, $s^2_{zh,p}=(N-1)^{-1}\sum_{i=1}^N(z_{ih,p}-\bar{z}_{h,p})^2$, $z_{ih,p}=T^{-1}\sum_{t=1}^T\left(x^*_{it,h}\right)^p$, $p=1,2$. The DGP for Panel D is identical to of Panel C, except that $\varpi_{it,h}\sim iidN(0,1)$.

Table 1 continued

Panel E: Size and Power of the LM test of Correlated Random Coefficient																								
	(A) Size			(B) Size			(C) Power, x_{ith} generated using χ_6^2						(D) Power, x_{ith} generated using $N(0,1)$					
	Slope Homo			Slope Hetero			$\sum_t x_{ith}^*$			$\sum_t (x_{ith}^*)^2$			$\sum_t x_{ith}^*$			$\sum_t (x_{ith}^*)^2$		
T,N	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
$LM_{CRC}^{(1)}$																		
25	5.1	5.4	4.9	4.9	5.3	5.1	87.1	99.7	100.0	5.1	4.5	4.9	89.2	99.5	100.0	4.8	5.4	5.3
50	5.5	5.6	5.2	4.4	5.2	5.7	96.7	100.0	100.0	5.2	4.4	5.4	96.8	100.0	100.0	4.5	4.4	5.9
100	5.7	4.9	4.9	4.6	4.9	5.7	99.0	100.0	100.0	4.8	4.7	5.8	99.3	100.0	100.0	4.9	4.6	5.8
200	4.2	4.4	5.4	4.5	5.1	4.9	99.7	100.0	100.0	5.0	4.9	5.6	99.8	100.0	100.0	5.7	4.8	5.8
$LM_{CRC}^{(2)}$																		
25	4.1	3.9	4.6	3.7	4.2	4.6	75.6	98.6	100.0	37.4	76.2	98.5	76.9	99.0	100.0	43.3	83.4	98.9
50	3.6	5.5	4.8	3.6	4.4	4.1	90.0	99.8	100.0	57.2	92.6	99.9	91.3	99.9	100.0	61.6	94.5	99.8
100	4.0	4.1	4.3	3.3	4.1	4.9	95.9	100.0	100.0	72.3	96.7	99.9	97.4	100.0	100.0	73.2	97.2	100.0
200	3.4	3.8	4.5	4.1	4.1	5.0	98.4	100.0	100.0	82.8	98.6	100.0	98.9	100.0	100.0	82.6	98.9	100.0

Notes for Panel E: The results of column blocks (A), (B), (C) and (D) are obtained using the same DGPs as Panels A, B, C and D, respectively. $LM_{CRC}^{(g)}$ is the proposed LM tests of correlated random effects defined by (55). The test statistics are referred to the 95% quantile of χ_g^2 distribution, $g = 1, 2$. All results are based on 2000 replications.

Table 2: Summary results of Bai and PC estimators for the model with $\{\beta_1, \beta_2\} = \{1, 3\}$, interactive effects, heteroskedastic and serially correlated errors

Panel A: Homogeneous Slopes, $\beta_{ih} = \beta_h$ for all i, $h = 1, 2$												
for β_1	Bias (x100)			RMSE (x100)			Size			Power		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{Bai}$												
25	0.048	0.015	0.091	2.717	1.932	1.336	6.9	6.5	5.5	50.5	76.2	97.1
50	0.013	0.018	0.048	1.919	1.341	0.942	6.9	6.4	5.6	76.5	96.6	100.0
100	-0.021	0.018	0.021	1.413	0.972	0.681	7.4	6.4	5.6	95.1	99.9	100.0
200	0.007	0.024	0.021	0.989	0.688	0.485	7.3	6.2	5.5	100.0	100.0	100.0
$\tilde{\beta}_{Bai}$												
25	0.038	0.006	0.082	2.712	1.931	1.334	6.9	6.4	5.5	50.2	76.1	97.1
50	0.003	0.009	0.040	1.918	1.341	0.942	6.8	6.2	5.6	76.3	96.7	100.0
100	-0.031	0.010	0.015	1.415	0.973	0.682	7.3	6.6	5.5	95.1	99.9	100.0
200	-0.008	0.014	0.014	0.996	0.689	0.485	7.1	6.1	5.4	99.9	100.0	100.0
$\hat{\beta}_{PC}$												
25	0.318	0.156	0.158	2.703	1.931	1.337	6.3	6.0	5.3	52.5	78.2	97.3
50	0.274	0.147	0.110	1.932	1.339	0.945	6.4	5.7	5.1	78.4	97.2	100.0
100	0.253	0.151	0.088	1.425	0.976	0.683	7.2	6.1	5.9	96.7	99.9	100.0
200	0.290	0.164	0.087	1.021	0.703	0.490	7.1	6.7	5.7	99.9	100.0	100.0
$\tilde{\beta}_{PC}$												
25	0.122	0.035	0.092	2.731	1.941	1.337	6.3	6.1	5.6	49.5	76.3	97.1
50	0.071	0.023	0.042	1.942	1.340	0.943	6.3	6.0	4.9	75.6	96.7	100.0
100	0.047	0.024	0.019	1.424	0.974	0.681	6.9	6.0	5.8	95.8	99.9	100.0
200	0.086	0.038	0.018	0.996	0.691	0.485	5.7	6.1	5.7	99.9	100.0	100.0

Notes for Panel A: Data is generated as $y^*_{it} = \sum_{h=1}^{2} x^*_{ith}\beta_{ih} + \sum_{\ell=1}^{2} f_{t\ell}\lambda_{i\ell} + \sigma_{\varepsilon,it}\varepsilon_{it}$, $i = 1, 2, ..., N; t = 1, 2, ..., T$, where $\lambda_{i\ell} \sim iidN(0,1)$, $f_{t\ell} = \rho_f f_{t-1,\ell} + \sqrt{1-\rho_f^2}\nu_{t\ell}$, $\nu_{t\ell} \sim iidN(0,1)$ with $f_{0,\ell} \sim iidN(0,1)$ for $\ell = 1, .., r$, $\varepsilon_{it} = \rho_\varepsilon \varepsilon_{it-1} + \sqrt{1-\rho_\varepsilon^2}\xi_{it}$, $\xi_{it} \sim iidN(0,1)$ with $\varepsilon_{i0} \sim iidN(0,1)$, and $\sigma_{\varepsilon,it} = (\kappa_{\varepsilon,i}\kappa_{\varepsilon,t})^{1/2}$, $\kappa_{\varepsilon,i} \sim iidU(0.5, 1.5)$ and $\kappa_{\varepsilon,t} = 0.5 + t/T$; $x^*_{ith} = \sum_{\ell=1}^{r} f_{t\ell}\gamma_{ih\ell} + \phi\sigma_{v,it}v_{ith}$,where $v_{ith} = \rho_v v_{it-1,h} + \sqrt{1-\rho_v^2}\varpi_{it,h}$, $\varpi_{it,h} \sim iid\left(\chi_6^2 - 6\right)/\sqrt{12}$ with $v_{i0,h} \sim iid\left(\chi_6^2 - 6\right)/\sqrt{12}$, $\gamma_{ih\ell} = 0.7\lambda_{i\ell} + \left(1 - 0.7^2\right)^{1/2}\varphi_{ih\ell}$, $\varphi_{ih\ell} \sim iidN(0,1)$, $\sigma_{v,it} = (\kappa_{v,i}\kappa_{v,t})^{1/2}$, $\kappa_{v,i} \sim iidU(0.5, 1.5)$ and $\kappa_{v,t} = 0.5 + t/T$, $\phi^2 = 2$. $\hat{\beta}_{Bai}$ is non-bias-corrected and $\tilde{\beta}_{Bai}$ is bias-corrected estimator proposed by Bai (2009) and $\hat{\beta}_{PC}$ is the PC estimator defined by (C.4) and $\tilde{\beta}_{PC}$ is the proposed bias corrected estimator defined by (31). The size is rejection frequency of the proposed Wald test (defined by (19)) for $H_0 : \beta_1 = 1$ and the power for $H_0 : \beta_1 = 0.95$, based on the 5% level test. All results are based on 2000 replications.

Table 2 continued

Panel B: Heterogeneous Slopes, $\beta_{ih} = \beta_h + \eta_{ih}$ with $\eta_{ih} \sim iidN(0, 0.04)$ for all i, $h = 1, 2$												
for β_1	Bias (x100)			RMSE (x100)			Size			Power		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{Bai}$												
25	0.002	-0.009	0.108	4.228	3.049	2.190	7.8	6.3	6.2	26.2	41.8	67.3
50	-0.095	-0.038	0.052	3.741	2.627	1.882	8.5	6.7	6.1	32.9	52.8	79.0
100	-0.170	-0.050	0.013	3.364	2.344	1.668	7.7	5.8	5.2	36.5	58.8	85.7
200	-0.134	-0.037	0.017	3.152	2.198	1.564	8.1	6.5	5.3	39.8	64.4	90.3
$\tilde{\beta}_{Bai}$												
25	-0.016	-0.024	0.096	4.227	3.049	2.189	7.7	6.3	6.2	25.9	41.6	66.9
50	-0.116	-0.056	0.038	3.742	2.629	1.882	8.4	6.9	5.9	32.8	52.6	78.8
100	-0.196	-0.070	-0.001	3.365	2.345	1.668	7.6	5.7	5.3	36.2	58.6	85.4
200	-0.181	-0.071	-0.004	3.153	2.200	1.565	7.7	6.4	5.3	39.4	63.3	90.0
$\hat{\beta}_{PC}$												
25	0.335	0.161	0.188	4.160	3.038	2.194	6.3	6.0	6.1	25.2	43.1	68.0
50	0.251	0.138	0.134	3.700	2.609	1.882	7.2	5.9	5.7	33.4	53.8	79.1
100	0.201	0.134	0.105	3.325	2.332	1.663	6.5	5.2	5.1	38.2	60.5	86.6
200	0.247	0.155	0.108	3.120	2.194	1.562	6.1	6.0	5.2	41.8	66.0	91.1
$\tilde{\beta}_{PC}$												
25	0.200	0.071	0.138	4.208	3.055	2.199	6.6	6.2	6.3	24.8	41.9	67.1
50	0.107	0.044	0.082	3.736	2.620	1.885	7.2	5.7	5.8	32.0	52.6	78.2
100	0.053	0.038	0.052	3.355	2.340	1.665	6.4	5.2	5.0	36.7	58.9	86.0
200	0.100	0.060	0.055	3.142	2.199	1.563	6.5	6.0	5.1	39.7	64.5	90.5

Notes for Panel B: See notes to Panel A.

Table 2 continued

Panel C: Correlated Heterogeneous Slopes, $\rho_{x\eta} = 0.5$, x^*_{ith} generated using χ^2_6												
	(i) β_{ih} is function of $\sum_t x^*_{ith}$						(ii) β_{ih} is function of $\sum_t (x^*_{ith})^2$					
for β_1	Bias (x100)			Size, $H_0 : \beta_1 = 1$			Bias (x100)			Size, $H_0 : \beta_1 = 1$		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{Bai}$												
25	0.439	0.456	0.576	7.5	6.1	5.4	0.591	0.724	0.918	6.7	6.1	6.9
50	0.364	0.456	0.546	8.6	5.9	5.1	0.556	0.754	0.911	8.4	6.5	7.1
100	0.352	0.497	0.571	7.4	5.4	5.4	0.526	0.789	0.918	7.8	6.7	7.0
200	0.476	0.594	0.660	7.3	6.2	5.2	0.650	0.881	1.008	7.4	7.6	8.4
$\tilde{\beta}_{Bai}$												
25	0.424	0.442	0.563	7.5	6.2	5.5	0.580	0.712	0.907	6.7	6.2	6.8
50	0.347	0.439	0.533	8.6	5.9	5.1	0.540	0.739	0.898	8.4	6.4	6.9
100	0.332	0.479	0.558	7.4	5.5	5.4	0.508	0.767	0.904	7.9	6.7	7.0
200	0.444	0.566	0.641	7.1	5.9	5.0	0.619	0.843	0.978	7.7	7.5	8.0
$\hat{\beta}_{PC}$												
25	0.795	0.634	0.657	6.2	5.9	5.7	1.117	1.002	1.059	6.3	6.6	7.2
50	0.732	0.640	0.631	6.9	5.3	4.9	1.059	1.010	1.036	7.5	6.5	7.5
100	0.739	0.688	0.665	6.8	5.0	5.4	1.043	1.046	1.047	7.3	6.5	7.6
200	0.869	0.790	0.753	6.9	5.9	5.3	1.172	1.146	1.136	7.0	7.8	9.0
$\tilde{\beta}_{PC}$												
25	0.671	0.551	0.611	6.5	5.8	5.6	0.918	0.880	0.992	6.6	6.4	7.0
50	0.600	0.553	0.583	7.0	5.2	4.9	0.850	0.882	0.966	7.3	6.2	7.1
100	0.605	0.600	0.616	6.6	4.9	5.2	0.830	0.916	0.976	7.2	6.3	6.9
200	0.736	0.703	0.704	6.7	5.9	5.2	0.959	1.015	1.065	6.7	7.5	8.4

Notes for Panel C: The data generating process (DGP) is the same as Panel B, except $\beta_{ih} = \beta_h + \sigma_\eta \left(\sqrt{1-\rho^2_{x\eta}} \eta_{ih} + \rho_{x\eta} w_{ih} \right)$, $\eta_{ih} \sim iid \left(\chi^2_6 - 6\right)/\sqrt{12}$ for $h = 1, 2$, $w_{ih} = \frac{z_{ih,p} - \bar{z}_{h,p}}{s_{zh,p}}$, where $\bar{z}_{h,p} = N^{-1} \sum_{i=1}^N z_{ih,p}$, $s^2_{zh,p} = (N-1)^{-1} \sum_{i=1}^N \left(z_{ih,p} - \bar{z}_{h,p}\right)^2$, $z_{ih,p} = T^{-1} \sum_{t=1}^T \left(x^*_{it,h}\right)^p$, $p = 1, 2$.

Table 2 continued

Panel D: Correlated Heterogeneous Slopes, $\rho_{x\eta} = 0.5$, x^*_{ith} generated using $N(0,1)$												
	(i) β_{ih} is function of $\sum_t x^*_{ith}$						(ii) β_{ih} is function of $\sum_t (x^*_{ith})^2$					
for β_1	Bias (x100)			Size, $H_0 : \beta_1 = 1$			Bias (x100)			Size, $H_0 : \beta_1 = 1$		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{Bai}$												
25	-0.006	-0.016	0.083	8.6	5.9	4.3	0.605	0.736	0.873	9.2	6.5	6.5
50	-0.174	-0.090	-0.014	7.7	6.2	4.3	0.460	0.678	0.804	8.2	6.8	6.2
100	-0.177	-0.114	-0.033	7.5	5.1	3.6	0.497	0.698	0.818	8.2	6.5	6.4
200	-0.227	-0.159	-0.079	7.4	5.5	4.0	0.525	0.709	0.831	8.3	6.4	6.6
$\tilde{\beta}_{Bai}$												
25	-0.022	-0.030	0.071	8.5	6.0	4.2	0.592	0.723	0.862	9.2	6.4	6.5
50	-0.191	-0.104	-0.026	7.9	6.2	4.3	0.443	0.665	0.793	8.1	6.8	6.2
100	-0.200	-0.131	-0.046	7.4	5.3	3.5	0.474	0.675	0.804	8.1	6.3	6.3
200	-0.266	-0.186	-0.097	7.3	5.5	4.1	0.486	0.669	0.800	8.4	6.1	6.6
$\hat{\beta}_{PC}$												
25	0.350	0.164	0.168	7.2	5.4	4.3	1.137	1.014	1.017	8.2	6.9	6.4
50	0.213	0.110	0.077	6.2	5.5	4.2	0.987	0.951	0.937	8.0	6.7	6.6
100	0.218	0.086	0.061	6.3	5.0	3.4	1.028	0.965	0.948	6.8	6.3	6.8
200	0.176	0.044	0.020	5.7	4.7	3.4	1.057	0.979	0.964	7.2	6.3	6.9
$\tilde{\beta}_{PC}$												
25	0.224	0.080	0.122	7.5	5.5	4.3	0.944	0.895	0.952	8.8	6.7	6.2
50	0.070	0.020	0.028	6.2	5.6	4.1	0.774	0.824	0.868	7.5	6.4	6.2
100	0.073	-0.007	0.010	6.5	4.9	3.4	0.812	0.835	0.878	6.6	6.3	6.3
200	0.029	-0.050	-0.031	6.0	5.0	3.6	0.839	0.847	0.893	7.3	5.6	6.6

See notes to Panel C. The DGP for Panel D is identical to of Panel C, except that $\varpi_{it,h} \sim iidN(0,1)$.

Table 2 continued

Panel E: Size and Power of the LM test of Correlated Random Coefficient																								
	(A) Size			(B) Size			(C) Power, x_{ith} generated using χ^2_6						(D) Power, x_{ith} generated using $N(0,1)$					
	Slope Homo			Slope Hetero			$\sum_t x^*_{ith}$			$\sum_t (x^*_{ith})^2$			$\sum_t x^*_{ith}$			$\sum_t (x^*_{ith})^2$		
T,N	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
$LM^{(1)}_{CRC_{Bai}}$																		
25	5.0	5.6	5.8	5.0	4.3	4.3	22.7	37.7	54.7	6.0	5.4	5.9	22.8	40.3	56.0	5.1	5.7	4.6
50	5.8	5.2	5.8	5.6	5.1	5.0	40.4	59.6	75.9	5.9	5.2	6.1	40.0	61.5	76.7	5.2	5.2	5.3
100	5.8	4.8	5.1	4.8	4.4	4.8	71.7	89.5	96.1	5.1	5.4	6.1	71.6	89.7	96.2	5.2	5.7	5.3
200	3.7	4.5	4.7	5.3	5.3	5.3	99.6	100.0	100.0	5.8	7.1	10.2	99.8	100.0	100.0	4.5	5.0	4.8
$LM^{(2)}_{CRC_{Bai}}$																		
25	3.8	4.7	5.3	4.0	4.2	5.1	17.0	32.7	49.5	10.7	20.6	38.5	16.2	32.6	51.0	12.3	24.6	47.2
50	4.0	5.0	5.8	4.8	4.8	4.3	31.7	55.0	72.2	13.6	25.7	48.7	30.3	53.3	72.9	15.6	30.6	55.3
100	4.5	4.8	4.6	4.5	4.6	5.4	59.7	85.2	95.1	14.8	32.2	58.9	60.6	86.3	95.1	17.3	35.4	63.6
200	3.9	5.0	4.7	6.2	4.6	4.6	98.3	100.0	100.0	17.5	40.3	74.5	98.6	100.0	100.0	17.2	38.5	70.4
$LM^{(1)}_{CRC_{PC}}$																		
25	5.1	6.0	6.1	5.1	4.7	4.9	22.7	37.4	53.9	5.7	5.2	5.5	23.2	39.7	56.5	4.8	5.9	4.5
50	5.7	5.2	6.1	5.4	4.8	5.4	39.1	59.9	75.7	5.3	5.5	6.1	40.4	61.3	76.7	5.3	5.2	5.3
100	5.2	5.1	4.8	4.7	4.5	4.8	70.1	88.9	95.8	4.8	4.8	5.8	72.6	90.2	96.2	5.2	5.5	5.2
200	4.0	4.7	4.4	5.2	5.1	5.4	99.7	100.0	100.0	4.9	6.8	9.9	99.7	100.0	100.0	4.6	5.0	4.9
$LM^{(2)}_{CRC_{PC}}$																		
25	4.4	4.6	5.6	4.4	4.2	5.0	17.4	32.2	49.2	9.4	18.9	37.4	16.7	32.4	50.7	11.2	23.1	46.6
50	5.1	5.4	6.0	5.2	5.0	4.4	31.2	54.1	72.4	10.7	23.3	46.4	31.5	53.5	72.9	12.4	28.0	53.0
100	5.5	6.3	4.9	5.0	4.5	5.3	58.3	84.4	95.0	11.2	27.3	54.6	59.8	86.1	95.2	12.4	29.7	59.5
200	6.0	6.1	5.4	6.1	4.4	4.6	98.2	100.0	100.0	13.1	33.5	69.8	98.5	100.0	100.0	10.9	31.8	65.8

Notes for Panel E: The results of column blocks (A), (B), (C) and (D) are obtained using the same DGPs as Panels A, B, C and D, respectively. The $LM^{(g)}_{CRC_{Bai}}$ and $LM^{(g)}_{CRC_{PC}}$ are the proposed LM tests of correlated random effects defined by (57), based on bias-corrected Bai (2009) estimator (defined by (C.4)) and the proposed bias-corrected PC estimator (defined by (31)), respectively. The test statistics are referred to the 95% quantile of χ^2_g distribution, $g=1,2$. All results are based on 2000 replications.

Table 3: Summary results of PC estimators for the model with $\{\beta_1, \beta_2\} = \{-1, -3\}$, interactive effects, heteroskedastic and serially correlated errors

Panel A: Homogeneous Slopes, $\beta_{ih} = \beta_h$ for all i, $h = 1, 2$												
for β_1	Bias (x100)			RMSE (x100)			Size			Power		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{PC}$												
25	2.907	1.573	1.014	4.151	2.559	1.707	20.6	15.3	12.8	79.4	92.3	99.4
50	2.489	1.279	0.715	3.230	1.876	1.194	26.0	17.2	12.1	96.6	99.6	100.0
100	2.356	1.199	0.625	2.800	1.559	0.930	38.9	26.0	17.0	99.9	100.0	100.0
200	2.338	1.182	0.599	2.577	1.379	0.773	58.6	39.5	24.8	100.0	100.0	100.0
$\tilde{\beta}_{PC}$												
25	1.067	0.595	0.509	3.214	2.131	1.475	9.0	7.6	8.2	59.1	80.9	98.0
50	0.577	0.265	0.192	2.155	1.408	0.981	7.2	6.8	5.8	78.8	97.0	100.0
100	0.421	0.168	0.093	1.559	1.016	0.699	7.8	6.4	6.0	96.5	99.9	100.0
200	0.416	0.152	0.065	1.131	0.729	0.496	7.1	6.2	5.6	99.9	100.0	100.0
Panel B: Heterogeneous Slopes, $\beta_{ih} = \beta_h + \eta_{ih}$ with $\eta_{ih} \sim iidN(0, 0.04)$ for all i, $h = 1, 2$												
for β_1	Bias (x100)			RMSE (x100)			Size			Power		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{PC}$												
25	3.191	1.720	1.112	5.413	3.566	2.489	13.4	10.5	9.5	50.1	61.0	80.9
50	2.728	1.406	0.806	4.672	2.986	2.058	14.0	9.4	8.1	56.8	70.0	87.4
100	2.562	1.315	0.710	4.248	2.695	1.813	14.2	9.5	7.3	62.6	77.1	93.3
200	2.562	1.307	0.688	4.079	2.571	1.710	14.7	9.3	7.2	68.2	83.0	95.5
$\tilde{\beta}_{PC}$												
25	1.255	0.684	0.576	4.609	3.217	2.309	8.0	7.5	7.2	33.4	49.3	73.4
50	0.715	0.333	0.250	3.872	2.658	1.914	7.8	6.0	6.0	37.2	56.0	80.4
100	0.523	0.224	0.145	3.427	2.359	1.675	6.4	5.2	5.5	39.8	61.7	87.2
200	0.537	0.216	0.121	3.207	2.218	1.570	6.5	6.0	5.2	43.3	66.2	90.8

Notes to Table 3: The DGPs for Panels A-E are identical to those in Table 2, except that $\{\beta_1, \beta_2\} = \{-1, -3\}$. The performance of Bai's estimators is not reported, since the results are identical to those in Table 2. See notes to Panels A and B in Table 2. The size is rejection frequency of the proposed Wald test (defined by (19)) for $H_0 : \beta_1 = -1$ and the power for $H_0 : \beta_1 = -1.05$, based on the 5% level test. All results are based on 2000 replications.

Table 3 continued

Panel C: Correlated Heterogeneous Slopes, $\rho_{x\eta}=0.5$, x^*_{ith} generated using χ^2_6												
	(i) β_{ih} is function of $\sum_t x^*_{ith}$						(ii) β_{ih} is function of $\sum_t (x^*_{ith})^2$					
for β_1	Bias (x100)			Size, $H_0:\beta_1=-1$			Bias (x100)			Size, $H_0:\beta_1=-1$		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{PC}$												
25	3.638	2.183	1.576	15.0	12.0	10.9	4.432	2.862	2.196	17.2	15.7	17.0
50	3.208	1.906	1.301	16.0	11.2	9.2	3.868	2.460	1.814	16.9	14.5	14.4
100	3.094	1.870	1.270	16.6	12.7	9.5	3.685	2.374	1.730	18.7	15.5	15.3
200	3.173	1.937	1.332	17.9	12.8	10.8	3.743	2.430	1.783	19.9	17.0	17.1
$\tilde{\beta}_{PC}$												
25	1.721	1.158	1.044	9.0	7.6	7.6	2.172	1.656	1.571	8.8	9.0	11.1
50	1.215	0.843	0.750	8.0	5.8	5.9	1.514	1.206	1.165	7.6	6.9	7.8
100	1.079	0.792	0.711	7.1	5.7	5.4	1.297	1.099	1.071	6.6	6.1	7.0
200	1.176	0.862	0.772	7.7	6.2	5.4	1.370	1.156	1.122	6.1	6.9	8.2
Panel D: Correlated Heterogeneous Slopes, $\rho_{x\eta}=0.5$, x^*_{ith} generated using $N(0,1)$												
	β_{ih} is a function of $\sum_t x^*_{ith}$						β_{ih} is a function of $\sum_t x^{*2}_{ith}$					
for β_1	Bias (x100)			Size, $H_0:\beta_1=-1$			Bias (x100)			Size, $H_0:\beta_1=-1$		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{PC}$												
25	3.237	1.735	1.086	15.7	10.0	7.3	4.480	2.888	2.141	19.5	16.3	15.7
50	2.754	1.410	0.757	13.4	8.7	6.0	3.831	2.426	1.716	17.7	14.4	14.3
100	2.607	1.291	0.672	13.2	8.1	4.8	3.682	2.305	1.629	17.5	14.9	13.7
200	2.508	1.216	0.609	13.0	7.4	4.8	3.631	2.271	1.613	18.9	16.6	14.5
$\tilde{\beta}_{PC}$												
25	1.307	0.699	0.550	9.2	6.8	5.5	2.242	1.690	1.522	10.3	9.3	9.8
50	0.733	0.329	0.199	7.0	5.9	4.3	1.473	1.170	1.068	7.4	6.7	7.4
100	0.555	0.192	0.104	6.6	5.1	3.3	1.285	1.027	0.970	6.4	6.1	6.5
200	0.462	0.114	0.038	5.9	4.7	3.5	1.240	0.990	0.951	6.0	5.6	6.6

See notes to Panel C and Panel D in Table 2..

Table 3 continued

Panel E: Size and Power of the LM test of Correlated Random Coefficient																								
	(A) Size			(B) Size			(C) Power, x_{ith} generated using χ^2_6						(D) Power, x_{ith} generated using $N(0,1)$					
	Slope Homo			Slope Hetero			$\sum_t x^*_{ith}$			$\sum_t (x^*_{ith})^2$			$\sum_t x^*_{ith}$			$\sum_t (x^*_{ith})^2$		
T,N	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
$LM^{(1)}_{CRC_{PC}}$																		
25	4.9	5.7	5.8	5.0	4.6	4.9	22.1	36.5	53.0	4.6	4.9	5.1	23.0	39.0	57.0	4.4	5.2	4.5
50	6.0	5.1	6.2	5.2	4.8	5.3	39.0	59.2	75.4	4.5	5.0	6.0	39.6	60.9	76.5	4.9	4.7	4.8
100	4.8	5.3	4.9	4.5	4.7	5.1	69.8	88.8	95.6	4.6	4.5	5.6	71.2	90.0	96.1	4.2	5.1	5.1
200	3.3	4.6	4.3	5.1	4.9	5.4	99.6	100.0	100.0	4.9	6.0	9.9	99.7	100.0	100.0	3.8	4.8	4.9
$LM^{(2)}_{CRC_{PC}}$																		
25	4.7	5.0	6.0	4.7	5.0	4.9	17.0	31.9	48.9	15.1	25.6	43.8	17.2	32.1	50.3	19.4	33.0	53.8
50	4.5	5.5	6.5	4.9	4.8	4.6	30.4	53.2	72.3	18.4	30.1	52.8	30.2	53.4	72.4	21.5	36.9	59.5
100	4.6	5.5	4.6	4.3	4.5	5.2	57.8	84.3	95.1	20.4	36.3	62.7	60.0	85.8	95.3	21.6	39.7	67.8
200	3.5	4.9	4.8	4.7	4.3	4.1	98.2	100.0	100.0	22.0	44.0	76.0	98.0	100.0	100.0	21.5	41.5	72.6

Notes for Panel E: See notes to Panel E in Table 2.

Table 4: Summary results of PC estimators for the model with $\{\beta_1, \beta_2\} = \{0, 0\}$, interactive effects, heteroskedastic and serially correlated errors

Panel A: Homogeneous Slopes, $\beta_{ih} = \beta_h$ for all i, $h = 1, 2$												
for β_1	Bias (x100)			RMSE (x100)			Size			Power		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{PC}$												
25	-0.506	-0.205	0.026	2.707	1.929	1.326	5.4	6.3	4.8	41.7	72.6	97.0
50	-0.622	-0.285	-0.094	1.997	1.358	0.941	7.5	6.1	5.3	65.2	95.0	100.0
100	-0.671	-0.301	-0.136	1.547	1.010	0.691	10.1	7.7	6.0	89.1	99.8	100.0
200	-0.638	-0.296	-0.141	1.167	0.743	0.502	11.7	8.0	6.4	99.6	100.0	100.0
$\tilde{\beta}_{PC}$												
25	0.045	0.074	0.168	2.673	1.924	1.339	6.2	6.0	5.4	49.3	76.8	97.5
50	-0.046	0.007	0.054	1.906	1.331	0.939	6.1	5.7	5.2	74.5	96.7	100.0
100	-0.083	-0.003	0.015	1.404	0.967	0.679	6.5	6.0	5.6	94.6	99.9	100.0
200	-0.045	0.004	0.011	0.981	0.683	0.482	5.6	5.9	5.7	99.9	100.0	100.0
Panel B: Heterogeneous Slopes, $\beta_{ih} = \beta_h + \eta_{ih}$ with $\eta_{ih} \sim iidN(0, 0.04)$ for all i, $h = 1, 2$												
for β_1	Bias (x100)			RMSE (x100)			Size			Power		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{PC}$												
25	-0.449	-0.177	0.066	4.143	3.020	2.178	6.0	5.9	5.8	20.7	38.3	66.0
50	-0.614	-0.281	-0.062	3.715	2.609	1.872	7.1	6.2	5.5	25.7	47.4	76.6
100	-0.697	-0.306	-0.112	3.371	2.338	1.661	6.6	5.6	5.2	28.1	53.9	83.6
200	-0.655	-0.292	-0.114	3.159	2.197	1.559	7.4	6.0	5.1	31.3	58.8	88.5
$\tilde{\beta}_{PC}$												
25	0.147	0.126	0.221	4.173	3.039	2.195	6.8	6.2	6.1	25.6	42.7	68.8
50	0.007	0.037	0.099	3.711	2.614	1.881	7.2	6.0	5.9	31.6	53.0	79.1
100	-0.063	0.018	0.053	3.344	2.337	1.665	6.8	5.2	5.2	35.5	58.7	85.9
200	-0.016	0.035	0.052	3.135	2.196	1.562	7.3	6.1	5.2	39.0	64.7	90.6

Notes to Table 4: The DGPs for Panels A-E are identical to those in Table 2, except that $\{\beta_1, \beta_2\} = \{0, 0\}$. The performance of Bai's estimators is not reported, since the results are identical to those in Table 2. See notes to Panels A and B in Table 2. The size is rejection frequency of the proposed Wald test (defined by (19)) for $H_0 : \beta_1 = 0$ and the power for $H_0 : \beta_1 = -0.05$, based on the 5% level test. All results are based on 2000 replications.

Table 4 continued

Panel C: Correlated Heterogeneous Slopes, $\rho_{x\eta} = 0.5$, x_{ith}^* generated using χ_6^2												
	(i) β_{ih} is function of $\sum_t x_{ith}^*$						(ii) β_{ih} is function of $\sum_t (x_{ith}^*)^2$					
for β_1	Bias (x100)			Size, $H_0 : \beta_1 = 0$			Bias (x100)			Size, $H_0 : \beta_1 = 0$		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{PC}$												
25	0.004	0.293	0.534	6.0	5.6	5.0	0.120	0.549	0.875	5.4	5.6	6.4
50	-0.139	0.220	0.434	6.6	5.0	4.4	-0.004	0.488	0.786	7.4	5.8	5.9
100	-0.162	0.247	0.447	5.9	4.4	4.6	-0.052	0.506	0.779	6.3	5.1	5.7
200	-0.036	0.342	0.531	5.7	4.5	4.2	0.070	0.597	0.862	6.1	6.2	6.5
$\tilde{\beta}_{PC}$												
25	0.614	0.604	0.693	6.7	5.9	5.8	0.813	0.912	1.064	6.6	6.7	7.7
50	0.497	0.545	0.600	7.2	5.2	5.1	0.712	0.864	0.981	8.1	6.6	7.4
100	0.487	0.579	0.616	6.6	5.1	5.4	0.677	0.889	0.976	7.5	6.7	7.3
200	0.620	0.677	0.702	6.6	6.0	5.2	0.806	0.983	1.062	7.6	7.9	8.6
Panel D: Correlated Heterogeneous Slopes, $\rho_{x\eta} = 0.5$, x_{ith}^* generated using $N(0,1)$												
	β_{ih} is a function of $\sum_t x_{ith}^*$						β_{ih} is a function of $\sum_t (x_{ith}^*)^2$					
for β_1	Bias (x100)			Size, $H_0 : \beta_1 = 0$			Bias (x100)			Size, $H_0 : \beta_1 = 0$		
T,N	50	100	200	50	100	200	50	100	200	50	100	200
$\hat{\beta}_{PC}$												
25	-0.446	-0.182	0.040	7.2	5.3	4.0	0.129	0.556	0.828	7.4	5.7	5.6
50	-0.662	-0.315	-0.122	6.2	5.8	4.0	-0.092	0.420	0.682	7.1	5.9	5.4
100	-0.675	-0.359	-0.158	6.7	4.8	3.3	-0.071	0.415	0.676	6.8	5.2	5.2
200	-0.726	-0.406	-0.204	6.6	5.0	3.7	-0.052	0.424	0.687	7.2	5.1	5.4
$\tilde{\beta}_{PC}$												
25	0.163	0.129	0.199	7.2	5.5	4.4	0.833	0.925	1.018	8.5	7.2	6.9
50	-0.034	0.007	0.041	6.2	5.8	4.1	0.628	0.799	0.877	7.9	6.7	6.5
100	-0.035	-0.031	0.008	6.3	4.8	3.3	0.661	0.800	0.874	7.4	6.5	6.6
200	-0.082	-0.076	-0.036	6.4	4.9	3.9	0.685	0.811	0.886	8.3	6.1	6.9

See notes to Panel C and Panel D in Table 2.

Table 4 continued

Panel E: Size and Power of the LM test of Correlated Random Coefficient																								
	Size			Size			Power, x_{ith} generated using χ_6^2						Power, x_{ith} generated using $N(0,1)$					
	Slope Homo			Slope Hetero			$\sum_t x_{ith}^*$			$\sum_t (x_{ith}^*)^2$			$\sum_t x_{ith}^*$			$\sum_t (x_{ith}^*)^2$		
T,N	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
$LM_{CRC_{PC}}^{(1)}$																		
25	5.2	5.6	6.3	5.4	4.6	4.7	22.8	37.3	54.1	6.1	5.5	5.4	23.7	39.6	56.4	4.9	6.0	4.4
50	5.8	5.4	5.9	5.9	4.9	5.2	39.7	60.0	75.7	5.6	5.3	6.2	40.7	61.6	76.7	5.3	5.1	5.3
100	5.4	5.0	5.0	4.8	4.7	4.8	71.2	89.2	95.8	4.7	5.3	6.0	73.0	89.9	96.2	5.0	5.4	5.3
200	3.4	4.4	4.6	5.5	5.3	5.3	99.7	100.0	100.0	5.5	7.2	10.1	99.7	100.0	100.0	4.8	4.8	5.0
$LM_{CRC_{PC}}^{(2)}$																		
25	5.0	4.8	5.6	4.2	4.4	4.9	17.2	32.9	49.0	9.4	19.0	37.3	17.5	32.4	50.9	10.6	22.8	46.2
50	4.5	5.4	5.8	4.9	5.1	4.2	31.7	54.2	72.4	10.1	23.0	46.4	31.1	53.8	72.9	12.6	27.4	53.0
100	5.3	5.9	4.6	4.8	4.6	5.4	59.3	84.7	95.0	11.3	27.6	54.9	60.8	86.3	95.1	12.8	29.7	59.2
200	5.4	5.5	5.2	6.0	4.5	4.7	98.4	100.0	100.0	13.5	34.2	70.7	98.8	100.0	100.0	12.0	32.2	66.6

See notes to Panel E in Table 2.

Appendix A: Lemmas and Proofs of Main Results for Section 2

We rely on the law of large numbers and central limit theorem results, which are stated in Lemmas A.1 and A.2, which are given and proved in Hansen (2007). The results which are stated as Lemmas A.3-A.6 are discussed and proven in Hansen (2007), but replicated here for convenience. The proof of main results, which are readily proven based on the lemmas, are given in A.2. We provide proofs of LemmaA.8 in A.3.[25]

A.1: Lemmas for Section 2

Lemma A.1 *Suppose* $\{W_{i,T}\}$ *are independent across* $i = 1, 2, ..., N$ *for all* T *with* $E(W_{i,T}) = \mu_{i,T}$ *and* $E|W_{i,T}|^{1+\delta} < \Delta < \infty$ *for some* $\delta > 0$ *and all* i, T. *Then* $N^{-1}\sum_{i=1}^{N}(W_{i,T} - \mu_{i,T}) \xrightarrow{p} 0$ *as* $(N,T) \xrightarrow{j} \infty$.

Lemma A.2 *Suppose* $\{\mathbf{w}_{i,T}\}$, $h \times 1$ *random vectors, are independent across* $i = 1, 2, ..., N$ *for all* T *with* $E(\mathbf{w}_{i,T}) = \mathbf{0}$, $E(\mathbf{w}_{i,T}\mathbf{w}_{i,T}') = \Sigma_{i,T}$ *and* $E\|\mathbf{w}_{i,T}\|^{2+\delta} < \Delta < \infty$ *for some* $\delta > 0$ *and all* i, T. *Assume* $\Sigma = \lim_{N,T\to\infty} N^{-1}\sum_{i=1}^{N}\Sigma_{i,T}$ *is positive definite and the smallest eigenvalue of* Σ *is strictly positive. Then,* $N^{-1/2}\sum_{i=1}^{N}\mathbf{w}_{i,T} \xrightarrow{d} N(\mathbf{0}, \mathbf{\Sigma})$ *as* $(N,T) \xrightarrow{j} \infty$.

Lemma A.3 *Let* $\{w_t\}$ *be a strong mixing sequence with* $E(w_t) = 0$, $E|w_t|^{s+\delta} < \Delta \leq \infty$ *and mixing coefficient* $\alpha(m)$ *of size* $(1-c)r/(r-c)$ *where* $c \in 2\mathbb{N}$, $s \leq c < r$. *Then, there is a constant* C *depending only on* s *and* $\alpha(m)$ *such that* $E\left|\sum_{t=1}^{T} w_t\right|^s \leq C\ D(s,\delta,T)$, *where* $D(s,\delta,T)$ *is as defined in Doukhan (1994) and satisfying* $D(s,\delta,T) = O(T)$ *for* $s \leq 2$ *and* $D(s,\delta,T) = O(T^{s/2})$ *for* $s > 2$.

Lemma A.4 *Under Assumptions A1 and A2,* $\bar{\mathbf{A}}_{NT} - \mathbf{A} \to_p \mathbf{0}$ *and* $\bar{\mathbf{A}}_{NT}^{-1} - \mathbf{A}^{-1} \to_p \mathbf{0}$ *as* $(N,T) \to \infty$, *where* $\bar{\mathbf{A}}_{NT}$ *and* $\mathbf{A}$ *are defined by (7) and in Assumption A2, respectively.*

Lemma A.5 *Under Assumptions A1-A3,* $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i'\boldsymbol{\varepsilon}_i \to_d N(\mathbf{0}, \mathbf{B})$, *where* $\mathbf{B}$ *is defined in Assumption A3.*

Lemma A.6 *Under Assumptions A1-A3,* $N^{-1}\sum_{i=1}^{N}\hat{\mathbf{B}}_{i,T} - \mathbf{B} \to_p \mathbf{0}$ *as* $(N,T) \to \infty$, *where* $\hat{\mathbf{B}}_{i,T} = T^{-1}\mathbf{X}_i'\hat{\varepsilon}_i\hat{\varepsilon}_i'\mathbf{X}_i$ *with* $\hat{\boldsymbol{\varepsilon}}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}$ *with* $\boldsymbol{\eta}_i = \mathbf{0}$ *for all* i, *and* $\mathbf{B}$ *is defined in Assumption A3.*

Lemma A.7 *Under Assumptions A1-A4,* $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i \to_d N(\mathbf{0}, \mathbf{C})$, *where* $\mathbf{C}$ *is defined in Assumption A4.*

Lemma A.8 *Under Assumptions A1-A4,* $N^{-1}\sum_{i=1}^{N}\hat{\mathbf{C}}_{i,T} - \mathbf{C} \to_p \mathbf{0}$ *as* $(N,T) \to \infty$, *where* $\hat{\mathbf{C}}_{i,T} = T^{-2}\mathbf{X}_i'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\mathbf{X}_i$ *with* $\hat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}$ *and* $\mathbf{C}$ *is defined in Assumption A4.*

A.2: Proofs of Main Results in Section 2

Proof of Theorem 1. Applying Lemmas A.4 and A.5, the result immediately follows. ■
Proof of Theorem 2. Applying Lemmas A.4 and A.7, the result immediately follows. ■
Proof of Proposition 1. Applying Lemmas A.8 and A.6, the result immediately follows. ■
Proof of Theorem 3. In the case of both slope homogeneity and slope heterogeneity, using Theorems 1&2, together with Proposition 1, $\left[\hat{\mathbf{V}}_{NT}\left(\hat{\boldsymbol{\beta}}\right)\right]^{-1/2}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \to_d N(\mathbf{0}, \mathbf{I}_k)$ as $(N,T) \to \infty$. It is straightforward to impose the linear restriction $H_0 : \mathcal{R}\boldsymbol{\beta} = \mathbf{r}$ and show that under the null, $\left[\mathcal{R}\hat{\mathbf{V}}_{NT}\left(\hat{\boldsymbol{\beta}}\right)\mathcal{R}'\right]^{-1/2}\left(\mathcal{R}\hat{\boldsymbol{\beta}} - \mathbf{r}\right) \to_d N(\mathbf{0}, \mathbf{I}_q)$ which implies that $\left(\mathcal{R}\hat{\boldsymbol{\beta}} - \mathbf{r}\right)'\left[\mathcal{R}\hat{\mathbf{V}}_{NT}\left(\hat{\boldsymbol{\beta}}\right)\mathcal{R}'\right]^{-1}\left(\mathcal{R}\hat{\boldsymbol{\beta}} - \mathbf{r}\right) \to_d \chi_q^2$ as $(N,T) \to \infty$. This completes the proof. ■

[25] Proof of other Lemmas in this subsection is provided in Appendix C.1 for convenience.

A.3: Proofs of Lemmas in Section A.1

Proof of Lemma A.8. We write

$$\begin{aligned} N^{-1}\sum_{i=1}^{N}\hat{\mathbf{C}}_{i,T} &= N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{u}_i\mathbf{u}_i'\mathbf{X}_i - N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\mathbf{u}_i'\mathbf{X}_i \\ &\quad -N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{u}_i\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbf{X}_i'\mathbf{X}_i + N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbf{X}_i \\ &= \mathbf{E}_1-\mathbf{E}_2-\mathbf{E}_3+\mathbf{E}_4. \end{aligned}$$

Recall $\mathbf{u}_i=\mathbf{X}_i\boldsymbol{\eta}_i+\boldsymbol{\varepsilon}_i$. First

$$\begin{aligned} \mathbf{E}_3 &= N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbf{X}_i'\mathbf{X}_i + N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbf{X}_i'\mathbf{X}_i \qquad (A.1)\\ &= \mathbf{E}_{31}+\mathbf{E}_{32}, \text{ say.} \qquad (A.2) \end{aligned}$$

$$vec\left(\mathbf{E}_{31}\right)=N^{-1}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\otimes\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\boldsymbol{\eta}_i\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)$$

but $E\left\|\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\otimes\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\boldsymbol{\eta}_i\right\|^{1+\delta}\leq\left(E\left\|\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\right\|^{2+2\delta}E\left\|\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\boldsymbol{\eta}_i\right\|^{2+2\delta}\right)^{1/2}\leq\Delta$ by (C.1) and (C.6). As $\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)=O_p\left(N^{-1/2}\right)$, $\mathbf{E}_{31}=O_p\left(N^{-1/2}\right)$. Similarly $vec\left(\mathbf{E}_{32}\right)=N^{-1}T^{-1/2}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\otimes\frac{\mathbf{X}_i'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)=O_p\left(N^{-1/2}T^{-1/2}\right)$, thus, $\mathbf{E}_3=O_p\left(N^{-1/2}\right)+O_p\left(N^{-1/2}T^{-1/2}\right)$. It is easily seen that $\mathbf{E}_2=O_p\left(N^{-1/2}\right)+O_p\left(N^{-1/2}T^{-1/2}\right)$. $\left\|\mathbf{E}_4\right\|\leq N^{-1}T^{-1}\sum_{i=1}^{N}\left\|T^{-1/2}\mathbf{X}_i'\right\|^2\left\|\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right\|^2=O_p\left(N^{-1}T^{-1}\right)$. Finally,

$$\begin{aligned} \mathbf{E}_1 &= N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\left(\mathbf{X}_i\boldsymbol{\eta}_i+\boldsymbol{\varepsilon}_i\right)\left(\mathbf{X}_i\boldsymbol{\eta}_i+\boldsymbol{\varepsilon}_i\right)'\mathbf{X}_i \\ &= N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\mathbf{X}_i'\mathbf{X}_i + N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{X}_i \\ &\quad +N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i\boldsymbol{\varepsilon}_i'\mathbf{X}_i + N^{-1}T^{-2}\sum_{i=1}^{N}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\boldsymbol{\eta}_i'\mathbf{X}_i'\mathbf{X}_i \\ &= \mathbf{G}_1+\mathbf{G}_2+\mathbf{G}_3+\mathbf{G}_4, \text{ say.} \end{aligned}$$

Since, $E\left\|T^{-3/2}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i\boldsymbol{\varepsilon}_i'\mathbf{X}_i\right\|^{1+\delta}\leq E\left(\left\|T^{-1}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i\right\|\left\|T^{-1/2}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\right\|\right)^{1+\delta}\leq E\left\|T^{-1}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i\right\|^{2+2\delta}E\left\|T^{-1/2}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\right\|^{2+2\delta}\leq\Delta$ by by (C.1) and (C.6), $\mathbf{G}_3=O_p\left(T^{-1/2}\right)$. By a similar derivation, it is easily seen that $\mathbf{G}_4=O_p\left(T^{-1/2}\right)$. By (C.1), $\mathbf{G}_2=O_p\left(T^{-1}\right)$. Finally by (C.6), $\mathbf{G}_1-\mathbf{C}\to_p\mathbf{0}$, and the required result follows. ■

Appendix B: Lemmas and Proofs of Main Results in Section 3

B.1: Lemmas

Proof of the consistency of factor estimators and other related results are in line with the discussion in Bai (2009).

Lemma B.1 *(i)* $T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}\left[\sigma_{\ell N}\left(s,t\right)\right]^2\leq\Delta$*, where* $\sigma_{\ell N}\left(s,t\right)=N^{-1}\sum_{i=1}^{N}E\left(e_{it\ell}e_{is\ell}\right)$ *for all* $\ell=1,...,m$*;*

(ii) $E\left(T^{-1}\sum_{t=1}^{T}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}e_{it\ell}\mathbf{g}_{i\ell}^{0}\right\|^{2}\right)\leq\Delta$ *for all* $\ell=1,...,m$;

(iii) $E\left(\left\|\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\sum_{i=1}^{N}e_{it\ell}\mathbf{g}_{i\ell}^{0}\right\|^{2}\right)\leq\Delta$ *for all* $\ell=1,...,m$;

Lemma B.2

$$\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^{0}=\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\mathbf{F}^{0}\boldsymbol{g}_{\ell i}^{0}\mathbf{e}_{\ell i}'\hat{\mathbf{F}}\hat{\mathbf{Q}}+\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\mathbf{e}_{\ell i}\boldsymbol{g}_{\ell i}^{0\prime}\mathbf{F}^{0\prime}\hat{\mathbf{F}}\hat{\mathbf{Q}}+\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\mathbf{e}_{\ell i}\mathbf{e}_{\ell i}'\hat{\mathbf{F}}\hat{\mathbf{Q}},$$

where

$$\hat{\mathbf{Q}}=\left(\boldsymbol{\Upsilon}_{N}^{0}\boldsymbol{\Lambda}_{0\hat{F}}\right)^{-1},\ \boldsymbol{\Lambda}_{0\hat{F}}=T^{-1}\mathbf{F}^{0\prime}\hat{\mathbf{F}},\ \text{and } \mathbf{R}=\left(\boldsymbol{\Xi}_{NT}\hat{\mathbf{Q}}\right)^{-1}, \tag{B.1}$$

$\boldsymbol{\Xi}_{NT}$ *is* r *eigenvectors of* $N^{-1}\sum_{i=1}^{N}\mathbf{Z}_{i}^{*}\mathbf{Z}_{i}^{*\prime}$ *corresponding to the first* r *largest eigenvalues, which is invertible.*

Lemma B.3 $T^{-s/2}\left\|\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right\|^{s}=O_{p}\left(\delta_{NT}^{-s}\right)$, $s=1,2$.

Lemma B.4 $N^{-1/2}\left\|\sum_{i=1}^{N}\hat{\mathbf{G}}_{i}-\mathbf{R}^{-1}\mathbf{G}_{i}^{0}\right\|^{s}=O_{p}\left(\delta_{NT}^{-s}\right)$, $s=1,2$.

Lemma B.5 $\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\hat{\mathbf{F}}\right\|=O_{p}\left(\delta_{NT}^{-2}\right)$, $\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\mathbf{F}^{0}\right\|=O_{p}\left(\delta_{NT}^{-2}\right)$, $\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\boldsymbol{\varepsilon}_{i}\right\|=O_{p}\left(\delta_{NT}^{-2}\right)$, $\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\mathbf{v}_{\ell i}\right\|=O_{p}\left(\delta_{NT}^{-2}\right)$, $\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\mathbf{e}_{\ell i}\right\|=O_{p}\left(\delta_{NT}^{-2}\right)$, $\left\|\mathbf{R}\mathbf{R}'-\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\right\|=O_{p}\left(\delta_{NT}^{-2}\right)$.

Lemma B.6 $\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\mathbf{v}_{\ell i}\boldsymbol{\gamma}_{\ell i}^{0\prime}\right\|=O_{p}\left(N^{-1/2}\right)+O_{p}\left(\delta_{NT}^{-2}\right)$, $\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\mathbf{e}_{\ell i}\mathbf{g}_{\ell i}^{0\prime}\right\|=O_{p}\left(N^{-1/2}\right)+O_{p}\left(\delta_{NT}^{-2}\right)$ *for all* ℓ.

Lemma B.7 $\left\|\frac{\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\mathbf{F}^{0}}{\sqrt{T}}\right\|^{2}\leq\Delta$, $\left\|\frac{\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\mathbf{v}_{\ell i}}{\sqrt{T}}\right\|^{2}\leq\Delta$, $\left\|\frac{\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\boldsymbol{\varepsilon}_{i}}{\sqrt{T}}\right\|^{2}\leq\Delta$, $\left\|\frac{\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\mathbf{e}_{\ell i}}{\sqrt{T}}\right\|^{2}\leq\Delta$ *for all* ℓ *where* $\Sigma_{EE}=N^{-1}\sum_{i=1}^{N}\mathbf{E}_{i}\mathbf{E}_{i}'$ *and* $\bar{\Sigma}_{EE}=N^{-1}\sum_{i=1}^{N}E\left(\mathbf{E}_{i}\mathbf{E}_{i}'\right)$.

Lemma B.8 $\frac{\mathbf{F}^{0\prime}\hat{\mathbf{F}}}{T}\rightarrow_{p}\boldsymbol{\Lambda}$ *as* $(N,T)\rightarrow\infty$, $\boldsymbol{\Lambda}$ *is fixed and positive definite.*

Lemma B.9 $\left\|\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^{0}}\right\|=O_{p}\left(\delta_{NT}^{-1}\right)$.

Lemma B.10 $\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_{i}^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^{0}}\right)\mathbf{X}_{i}^{*}=O_{p}\left(\delta_{NT}^{-1}\right)$. *In the following lemmas, we consider the properly scaled limiting properties of these three terms.*

Lemma B.11 *Under Assumptions B1-B5,*

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_{i}^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_{i}^{*}\boldsymbol{\eta}_{i}=\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_{i}^{*\prime}\mathbf{M}_{\mathbf{F}^{0}}\mathbf{X}_{i}^{*}\boldsymbol{\eta}_{i}+O_{p}\left(\delta_{NT}^{-1}\right). \tag{B.2}$$

Lemma B.12 *Under Assumptions B1-B5,*

$$\begin{aligned}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{i}^{0}\mathbf{F}^{0\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} &= -\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{i}^{0}\left(\boldsymbol{\Upsilon}_{N}^{0}\right)^{-1}\boldsymbol{G}_{j}^{0}\frac{\mathbf{E}_{j}'\mathbf{M}_{\mathbf{F}^{0}}\boldsymbol{\varepsilon}_{i}}{T} \\ &\quad -\sqrt{\frac{T}{N}}\frac{1}{N}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_{i}^{0}\left(\boldsymbol{\Upsilon}_{N}^{0}\right)^{-1}\boldsymbol{G}_{n}^{0}\frac{\mathbf{E}_{n}'\mathbf{E}_{j}}{T}\boldsymbol{G}_{j}^{0\prime}\left(\boldsymbol{\Upsilon}_{N}^{0}\right)^{-1}\boldsymbol{\lambda}_{i}^{0}+o_{p}(1).\end{aligned} \tag{B.3}$$

Lemma B.13 *Under Assumptions B1-B5,*

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i}=\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_i+\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{E}_j}{T}\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i^0+o_p(1). \tag{B.4}$$

Lemma B.14 *Under Assumptions B1-B5,*

$$\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_i}{T}=\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{1i}^0\sigma_{\varepsilon i}^2+o_p(1). \tag{B.5}$$

where $\boldsymbol{g}_{1i}^0$ *is the first column of* $\boldsymbol{G}_i^0$ *and* $\sigma_{\varepsilon i}^2=T^{-1}\sum_{t=1}^{T}E\left(\varepsilon_{it}^2\right)$*. In a similar manner,*

$$\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i\boldsymbol{\eta}_i}{T}=\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{1i}^0\psi_{v\eta i}+o_p(1) \tag{B.6}$$

$$\psi_{v\eta i}=tr\left(\boldsymbol{\Sigma}_{vvi}\boldsymbol{\Omega}_{\eta\eta,i}\right).$$

Lemma B.15 *Under Assumptions B1-B5,*

$$\begin{aligned}&\sqrt{\frac{T}{N}}\frac{1}{N}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_n^0\frac{\mathbf{E}_n'\mathbf{E}_j}{T}\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i\\ =\;&\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{G}_j^0\bar{\boldsymbol{\Omega}}_{EE,j}\boldsymbol{G}_j^{0\prime}\right)\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i+o_p(1)\end{aligned} \tag{B.7}$$

where $\bar{\boldsymbol{\Omega}}_{EE,i}=E\left(\mathbf{E}_i'\mathbf{E}_i/T\right)$*.*

Lemma B.16 *Under Assumptions B1-B5,*

$$\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{E}_j}{T}\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i=\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\bar{\boldsymbol{\Omega}}_{VE,i}\boldsymbol{G}_i^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i+o_p(1) \tag{B.8}$$

where $\bar{\boldsymbol{\Omega}}_{VE,i}=E\left(\mathbf{V}_i'\mathbf{E}_i/T\right)$*.*

Lemma B.17 *Under Assumptions B1-B5,* $\hat{\boldsymbol{\xi}}_{NT}-\boldsymbol{\xi}_{NT}^{\dagger}\to_p\mathbf{0}$ *as* $(N,T)\to\infty$ *such that* $T/N\to c\in(0,\Delta]$*.*

B.2: Proof of Main Results in Section 3

Proof of Proposition 2. First, using $\mathbf{X}_i^*=\mathbf{F}^0\boldsymbol{\Gamma}_i^{0\prime}+\mathbf{V}_i$, we can write

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon i}=\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\mathbf{F}^{0\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon i}+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon i}. \tag{B.9}$$

Substituting (B.3) and (B.4) in Lemmas B.12&B.13 into (B.9) gives

$$\begin{aligned}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon i}\;=\;&-\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_i}{T}\\&-\sqrt{\frac{T}{N}}\frac{1}{N}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_n^0\frac{\mathbf{E}_n'\mathbf{E}_j}{T}\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i^0\\&+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_i+\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{E}_j}{T}\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i^0+o_p(1).\end{aligned} \tag{B.10}$$

Further substituting (B.5), (B.7) and (B.8) into (B.10) gives the required result.

Proof of Proposition 3. Recall that $\hat{\mathbf{c}}_{NT}=\left(\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\hat{\boldsymbol{\xi}}_{NT}$. By Lemma B.10, $\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i-\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i=o_p(1)$ and, by continuous mapping theorem, $\left(\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}-\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i\right)^{-1}=o_p(1)$. By Lemma B.17, $\hat{\boldsymbol{\xi}}_{NT}-\boldsymbol{\xi}_{NT}^{\dagger}=o_p(1)$, therefore, $\hat{\mathbf{c}}_{NT}-\mathbf{c}=\left(\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\hat{\boldsymbol{\xi}}_{NT}-\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\right)^{-1}\boldsymbol{\xi}_{NT}^{\dagger}=o_p(1)$, as required. ■

Proof of Theorem 4. Under slope homogeneity, recalling $\tilde{\boldsymbol{\beta}}_{PC}=\hat{\boldsymbol{\beta}}_{PC}-\frac{1}{N}\hat{\mathbf{c}}_{NT}$ with $\hat{\mathbf{c}}_{NT}=\left(\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\hat{\boldsymbol{\xi}}_{NT}$, we have

$$\sqrt{NT}\left(\tilde{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)=\left(\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i}-\sqrt{\frac{T}{N}}\hat{\boldsymbol{\xi}}_{NT}\right),$$

but by Lemma B.10 and continuous mapping theorem, $\left(\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}-\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i\right)^{-1}=o_p(1)$, and using Proposition 2, we can write

$$\sqrt{NT}\left(\tilde{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)=\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i\right)^{-1}\left[\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\boldsymbol{\varepsilon}_i-\sqrt{\frac{T}{N}}\left(\hat{\boldsymbol{\xi}}_{NT}-\boldsymbol{\xi}_{NT}\right)\right]+o_p(1),$$

where $\boldsymbol{\xi}_{NT}$ is defined by (29). By Lemma B.17,$\hat{\boldsymbol{\xi}}_{NT}-\boldsymbol{\xi}_{NT}=o_p(1)$ with slope homogeneity, we have

$$\sqrt{NT}\left(\tilde{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)=\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i\right)^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\boldsymbol{\varepsilon}_i\right)+o_p(1)$$

as $(N,T)\to\infty$ such that $T/N\to c\in(0,\Delta]$. By Assumption B5, $\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i\to_p\mathbf{A}$, and $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\boldsymbol{\varepsilon}_i\to_d N\left(\mathbf{0},\tilde{\mathbf{B}}\right)$ by Lemma A.2 and Assumption B5, as $(N,T)\to\infty$ such that $T/N\to c\in(0,\Delta]$, the required result follows. ■

Proof of Theorem 5. Under heterogeneous slopes, recalling $\tilde{\boldsymbol{\beta}}_{PC}=\hat{\boldsymbol{\beta}}_{PC}-\frac{1}{N}\hat{\mathbf{c}}_{NT}$ with $\hat{\mathbf{c}}_{NT}=\left(\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\hat{\boldsymbol{\xi}}_{NT}$, we have

$$\begin{aligned}\sqrt{N}\left(\tilde{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\ &=\ \left(\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\left[\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^{*}}{T}\boldsymbol{\eta}_i\right.\\ &\left.\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i}-\frac{1}{\sqrt{N}}\hat{\boldsymbol{\xi}}_{NT}\right].\end{aligned}\tag{B.11}$$

By Lemmas B.10 & B.11 and Proposition 2 we have

$$\begin{aligned}\sqrt{N}\left(\tilde{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\ &=\ \left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i\right)^{-1}\left[\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{V}_i}{T}\boldsymbol{\eta}_i\right.\\ &\left.\quad+\left\{\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\boldsymbol{\varepsilon}_i+\frac{1}{\sqrt{N}}\left(\hat{\boldsymbol{\xi}}_{NT}-\boldsymbol{\xi}_{NT}^{\dagger}\right)\right\}\right]+o_p(1).\end{aligned}\tag{B.12}$$

Because $\hat{\boldsymbol{\xi}}_{NT}-\boldsymbol{\xi}_{NT}^{\dagger}=o_p(1)$ with $\boldsymbol{\xi}_{NT}^{\dagger}=O_p(1)$ by Proposition 3 and $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\boldsymbol{\varepsilon}_i=O_p(1)$ by Assumption B5, inside of the curly brackets is $O_p\left(T^{-1/2}\right)+o_p\left(N^{-1/2}\right)$. Therefore,

$$\sqrt{N}\left(\tilde{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)=\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{V}_i}{T}\boldsymbol{\eta}_i+o_p(1)$$

as $(N,T) \to \infty$ such that $T/N \to c \in (0,\Delta]$. By Assumption B5, $\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i \to_p \mathbf{A}$ and $\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i\right)^{-1} \to_p \mathbf{A}^{-1}$, and $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{V}_i}{T}\boldsymbol{\eta}_i \to_d N\left(\mathbf{0},\tilde{\mathbf{C}}\right)$ by Lemma A.2 and Assumption B5, as $(N,T)\to\infty$ such that $T/N \to c \in (0,\Delta]$, the required result follows. ■

Proof of Theorem 6. Consider

$$\hat{\mathbf{V}}_{PC} = \left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}\left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\hat{\mathbf{X}}_i\right)\left(\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1}, \tag{B.13}$$

where $\hat{\mathbf{u}}_i = \mathbf{y}_i^* - \mathbf{X}_i^*\hat{\boldsymbol{\beta}}_{PC}$. Under homogeneous or heterogeneous slopes, by Lemma B.10 and continuous mapping theorem, $\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i - \frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{V}_i = o_p(1)$, then $\left(\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\right)^{-1} - \mathbf{A}^{-1} = o_p(1)$. First consider the slope homogeneous case. Noting $\hat{\mathbf{X}}_i'\hat{\mathbf{u}}_i = \hat{\mathbf{X}}_i'\left[\boldsymbol{\epsilon}_{\varepsilon fi} - \mathbf{X}_i^*\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)\right]$ we have

$$\begin{aligned}\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\hat{\mathbf{X}}_i &= (NT)^{-1}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\boldsymbol{\epsilon}_{\varepsilon fi}\boldsymbol{\epsilon}_{\varepsilon fi}'\hat{\mathbf{X}}_i - (NT)^{-1}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)\boldsymbol{\epsilon}_{\varepsilon fi}'\hat{\mathbf{X}}_i \qquad \text{(B.14)}\\ &\quad - (NT)^{-1}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\boldsymbol{\epsilon}_{\varepsilon fi}\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)'\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i + (NT)^{-1}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)'\hat{\mathbf{X}}_i\\ &= \mathbf{A}_1 - \mathbf{A}_2 - \mathbf{A}_3 + \mathbf{A}_4. \qquad \text{(B.15)}\end{aligned}$$

$$\begin{aligned}\mathbf{A}_3 &= TN^{-1}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{\varepsilon fi}}{T}\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)'\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\\ &= TN^{-1}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\epsilon}_{\varepsilon fi}}{T}\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)'\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\mathbf{X}_i^*}{T}\\ &\quad + TN^{-1}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{\varepsilon fi}}{T}\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)'\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\\ &\quad + TN^{-1}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\epsilon}_{\varepsilon fi}}{T}\left(\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right)'\frac{\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}^0}\right)\mathbf{X}_i^*}{T}\\ &= \mathbf{A}_{31} + \mathbf{A}_{32} + \mathbf{A}_{33}.\end{aligned}$$

but as

$$E\left\|\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\right\|^2 \le E\left\|\frac{\mathbf{X}_i^{*\prime}\mathbf{X}_i^*}{T}\right\|^2 + E\left\|\frac{\mathbf{X}_i^{*\prime}\hat{\mathbf{F}}}{T}\right\|^4 \le E\left\|\frac{\mathbf{X}_i^{*\prime}}{T}\right\|^8 + r^2 \le \Delta$$

by Assumptions B1-B5 and

$$\left\|\frac{\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{\varepsilon fi}}{T}\right\|^2 \le \left\|\left(N^{-1/2}\sum_{i=1}^{N}\frac{\boldsymbol{\epsilon}_{\varepsilon fi}'\otimes\mathbf{X}_i^{*\prime}}{T}\right)\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}^0}\right)\right\|^2 \le \left\|N^{-1/2}\sum_{i=1}^{N}\frac{\boldsymbol{\epsilon}_{\varepsilon fi}'\otimes\mathbf{X}_i^{*\prime}}{T}\right\|^2\left\|\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}^0}\right\|^2,$$

we have

$$\begin{aligned}\left\|\mathbf{A}_{32}\right\| &\le TN^{-1}\sum_{i=1}^{N}\left\|\frac{\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{\varepsilon fi}}{T}\right\|\left\|\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\right\|\left\|\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right\|\\ &\le T\left(N^{-1}\sum_{i=1}^{N}\left\|\frac{\boldsymbol{\epsilon}_{\varepsilon fi}'\otimes\mathbf{X}_i^{*\prime}}{T}\right\|^2\left\|\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}^0}\right\|^2\right)^{1/2}\left(N^{-1}\sum_{i=1}^{N}\left\|\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\right\|^2\right)^{1/2}\left\|\hat{\boldsymbol{\beta}}_{PC} - \boldsymbol{\beta}\right\|\\ &= TO_p\left(\delta_{NT}^{-1}\right)O_p\left(\frac{1}{\sqrt{NT}}\right) = O_p\left(\sqrt{\frac{T}{N}}\right)O_p\left(\delta_{NT}^{-1}\right) = O_p\left(\delta_{NT}^{-1}\right).\end{aligned}$$

In a similar manner, it is easily shown that $\|\mathbf{A}_{33}\| = O_p\left(\delta_{NT}^{-1}\right)$.

$$\begin{aligned}
\|\mathbf{A}_{31}\| &\leq \sqrt{T}N^{-1}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_{\varepsilon i}}{\sqrt{T}}\right\|\left\|\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i^*}{T}\right\|\left\|\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right\| \\
&\leq \left(N^{-1}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_{\varepsilon i}}{\sqrt{T}}\right\|\right)^{1/2}\left(N^{-1}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i^*}{T}\right\|\right)^{1/2}\left\|\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right\| \\
&= \sqrt{T}O_p\left(\frac{1}{\sqrt{NT}}\right) = O_p\left(\frac{1}{\sqrt{N}}\right)
\end{aligned}$$

hence, $\mathbf{A}_3 = o_p(1)$. In a similar manner, it is easily shown that $\mathbf{A}_2 = o_p(1)$.

$$\|\mathbf{A}_4\| \leq N^{-1}\sum_{i=1}^{N}\left\|\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\right\|\left\|\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right\|^2 = O_p\left(\frac{1}{NT}\right).$$

Next, consider $\mathbf{A}_1$.

$$\begin{aligned}
\mathbf{A}_1 &= (NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\epsilon}_{\varepsilon fi}\boldsymbol{\epsilon}_{\varepsilon fi}'\mathbf{M}_{\mathbf{F}^0}\mathbf{X}_i^* \\
&\quad + (NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{\varepsilon fi}\boldsymbol{\epsilon}_{\varepsilon fi}'\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^* \\
&\quad + (NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\epsilon}_{\varepsilon fi}\boldsymbol{\epsilon}_{\varepsilon fi}'\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\mathbf{X}_i^* \\
&= \mathbf{A}_{11}+\mathbf{A}_{12}+\mathbf{A}_{13}
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{A}_{12}\| &\leq \left\|(NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{\varepsilon fi}\boldsymbol{\epsilon}_{\varepsilon fi}'\mathbf{M}_{\mathbf{F}^0}\mathbf{X}_i^*\right\| \\
&\quad + \left\|(NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{\varepsilon fi}\boldsymbol{\epsilon}_{\varepsilon fi}'\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\mathbf{X}_i^*\right\| \\
&\leq \left(N^{-1}\sum_{i=1}^{N}\left\|\frac{\boldsymbol{\epsilon}_{\varepsilon fi}'\otimes\mathbf{X}_i^{*\prime}}{T}\right\|^2\left\|\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right\|^2\right)^{1/2}\left(\frac{1}{NT}\sum_{i=1}^{N}\left\|\boldsymbol{\varepsilon}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i\right\|^2\right)^{1/2} = O_p\left(\delta_{NT}^{-1}\right).
\end{aligned}$$

In a similar manner, it is easily shown that $\mathbf{A}_{13} = O_p\left(\delta_{NT}^{-1}\right)$. Therefore, we have

$$\begin{aligned}
\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\hat{\mathbf{X}}_i &= \frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\epsilon}_{\varepsilon fi}\boldsymbol{\epsilon}_{\varepsilon fi}'\mathbf{M}_{\mathbf{F}^0}\mathbf{X}_i^* + o_p(1) \\
&= \frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{V}_i + o_p(1).
\end{aligned}$$

Using Lemma A.1, $\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{V}_i - \tilde{\mathbf{B}} = o_p(1)$, and we conclude that, under Assumptions B1-B5,

$$(NT)^{-1}\hat{\mathbf{V}}_{PC} - \tilde{\mathbf{A}}^{-1}\tilde{\mathbf{B}}\tilde{\mathbf{A}}^{-1} \to_p \mathbf{0} \tag{B.16}$$

as $(N,T)\to\infty$ such that $T/N \to c(0,\Delta]$. ■ ■

Now consider the case with heterogeneous slopes. Noting $\hat{\mathbf{X}}_i'\hat{\mathbf{u}}_i = \hat{\mathbf{X}}_i'\left[\mathbf{X}_i^*\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_{\varepsilon fi} - \mathbf{X}_i^*\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\right]$

we have

$$\frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\hat{\mathbf{X}}_i = \frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\left(\mathbf{X}_i^*\boldsymbol{\eta}_i+\boldsymbol{\epsilon}_{\varepsilon fi}\right)\left(\mathbf{X}_i^*\boldsymbol{\eta}_i+\boldsymbol{\epsilon}_{\varepsilon fi}\right)'\hat{\mathbf{X}}_i - \frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\left(\mathbf{X}_i^*\boldsymbol{\eta}_i+\boldsymbol{\epsilon}_{\varepsilon fi}\right)'\hat{\mathbf{X}}_i \tag{B.17}$$

$$-\frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\left(\mathbf{X}_i^*\boldsymbol{\eta}_i+\boldsymbol{\epsilon}_{\varepsilon fi}\right)\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)'\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i + \frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)'\hat{\mathbf{X}}_i$$

$$= \mathbf{B}_1-\mathbf{B}_2-\mathbf{B}_3+\mathbf{B}_4. \tag{B.18}$$

Consider $\mathbf{B}_3$.

$$\begin{aligned}\mathbf{B}_3 &= \frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\mathbf{X}_i^*\boldsymbol{\eta}_i\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i+\frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\boldsymbol{\epsilon}_{\varepsilon fi}\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\hat{\mathbf{X}}_i'\hat{\mathbf{X}}_i\\ &= \mathbf{B}_{31}+\mathbf{B}_{32}.\end{aligned}$$

$\mathbf{B}_{32}$ is similar to $\mathbf{A}_3/T$, but noting $\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}=O_p\left(N^{-1/2}\right)$, we have $\mathbf{B}_{32}=O_p\left(\delta_{NT}^{-1}\right)T^{-1/2}=o_p\left(1\right)$.

$$\left\|\mathbf{B}_{31}\right\|\leq\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\eta}_i\right\|\left\|\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right\|\left\|\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\right\|^2=O\left(N^{-1/2}\right),$$

thus $\mathbf{B}_3=o_p\left(1\right)$. In a similar manner, it is easily shown that $\mathbf{B}_2=o_p\left(1\right)$ and $\mathbf{B}_4=o_p\left(1\right)$. Now consider $\mathbf{B}_1$.

$$\begin{aligned}\mathbf{B}_1 &= \frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\mathbf{X}_i^*\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\mathbf{X}_i^{*\prime}\hat{\mathbf{X}}_i+\frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\mathbf{X}_i^*\boldsymbol{\eta}_i\boldsymbol{\epsilon}_{\varepsilon fi}'\hat{\mathbf{X}}_i\\ &\quad+\frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\boldsymbol{\epsilon}_{\varepsilon fi}\boldsymbol{\eta}_i'\mathbf{X}_i^{*\prime}\hat{\mathbf{X}}_i+\frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i'\boldsymbol{\epsilon}_{\varepsilon fi}\boldsymbol{\epsilon}_{\varepsilon fi}'\hat{\mathbf{X}}_i\\ &= \mathbf{B}_{11}+\mathbf{B}_{12}+\mathbf{B}_{13}+\mathbf{B}_{14}.\end{aligned}$$

$$\begin{aligned}\mathbf{B}_{13} &= \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_i}{T}\boldsymbol{\eta}_i'\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}+\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{\varepsilon fi}}{T}\boldsymbol{\eta}_i'\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\\ &= \mathbf{C}_1+\mathbf{C}_2\end{aligned}$$

but

$$\left\|\mathbf{C}_2\right\|\leq\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\boldsymbol{\epsilon}_{\varepsilon fi}'\otimes\mathbf{X}_i^{*\prime}}{T}\right\|\left\|\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right\|\left\|\boldsymbol{\eta}_i\right\|\left\|\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\right\|=O_p\left(\delta_{NT}^{-1}\right)$$

and

$$\left\|\mathbf{C}_1\right\|\leq\frac{1}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\boldsymbol{\eta}_i'\right\|\left\|\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\right\|=O_p\left(T^{-1/2}\right),$$

thus, $\mathbf{B}_{13}=o_p\left(1\right)$. In a similar manner, it is easily shown that $\mathbf{B}_{12}=o_p\left(1\right)$. Also, $\mathbf{B}_{14}=\mathbf{A}_1/T=O_p\left(T^{-1}\right)$.

$$\begin{aligned}\mathbf{B}_{11} &= \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i}{T}\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i}{T}+\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\mathbf{X}_i^*}{T}\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*}{T}\\ &\quad+\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\mathbf{X}_i^*}{T}\boldsymbol{\eta}_i\boldsymbol{\eta}_i'\frac{\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\mathbf{X}_i^*}{T}\\ &= \mathbf{D}_1+\mathbf{D}_2+\mathbf{D}_3.\end{aligned}$$

$$
\begin{aligned}
\mathbf{D}_2 &= \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\mathbf{X}_i^{*}}{T}\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\prime}\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\mathbf{X}_i^{*}}{T} \\
&\quad +\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\mathbf{X}_i^{*}}{T}\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\prime}\frac{\mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\mathbf{X}_i^{*}}{T} \\
&= \mathbf{D}_{21}+\mathbf{D}_{22},
\end{aligned}
$$

but

$$
\left\|\mathbf{D}_{21}\right\| \leq \frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{X}_i^{*\prime}}{\sqrt{T}}\right\|^2\left\|\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right\|\left\|\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\prime}\right\|\left\|\frac{\mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}^0}\mathbf{X}_i^{*}}{T}\right\| = O_p\left(\delta_{NT}^{-1}\right)
$$

and

$$
\left\|\mathbf{D}_{22}\right\| \leq \frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{X}_i^{*\prime}}{\sqrt{T}}\right\|^4\left\|\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right\|^2\left\|\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\prime}\right\| = O_p\left(\delta_{NT}^{-2}\right),
$$

thus, $\mathbf{D}_2 = o_p(1)$. In a similar manner, it is easily shown that $\mathbf{D}_3 = O_p\left(\delta_{NT}^{-1}\right)$. To sum up

$$
\begin{aligned}
\frac{1}{NT^2}\sum_{i=1}^{N}\hat{\mathbf{X}}_i^{\prime}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i^{\prime}\hat{\mathbf{X}}_i &= \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{V}_i^{\prime}\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i}{T}\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\prime}\frac{\mathbf{V}_i^{\prime}\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i}{T}+o_p(1) \\
&= \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{V}_i^{\prime}\mathbf{V}_i}{T}\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\prime}\frac{\mathbf{V}_i^{\prime}\mathbf{V}_i}{T}+o_p(1).
\end{aligned}
$$

Using Lemma A.1, $\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{V}_i^{\prime}\mathbf{V}_i}{T}\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\prime}\frac{\mathbf{V}_i^{\prime}\mathbf{V}_i}{T}-\tilde{\mathbf{C}} = o_p(1)$, and we conclude that, under Assumptions B1-B5,

$$
N^{-1}\hat{\mathbf{V}}_{PC}-\tilde{\mathbf{A}}^{-1}\tilde{\mathbf{C}}\tilde{\mathbf{A}}^{-1} \rightarrow_p \mathbf{0} \tag{B.19}
$$

as $(N,T)\rightarrow\infty$ such that $T/N \rightarrow c(0,\Delta]$.

In the case of both slope homogeneity and slope heterogeneity, using Theorems 4 and 5, together with (B.16) and (B.19), $\hat{\mathbf{V}}_{PC}^{-1/2}\left(\tilde{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right) \rightarrow_d N\left(\mathbf{0},\mathbf{I}_k\right)$. It is straightforward to impose the linear restriction $H_0 : \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$ and show that under the null, $\left(\mathbf{R}\hat{\mathbf{V}}_{PC}\mathbf{R}\right)^{-1/2}\left(\mathbf{R}\tilde{\boldsymbol{\beta}}_{PC}-\mathbf{r}\right) \rightarrow_d N\left(\mathbf{0},\mathbf{I}_q\right)$ which implies that $\left(\mathbf{R}\tilde{\boldsymbol{\beta}}_{PC}-\mathbf{r}\right)^{\prime}\left(\mathbf{R}\hat{\mathbf{V}}_{PC}\mathbf{R}\right)^{-1}\left(\mathbf{R}\tilde{\boldsymbol{\beta}}_{PC}-\mathbf{r}\right) \rightarrow_d \chi_q^2$ as $(N,T)\rightarrow\infty$ such that $T/N \rightarrow c(0,\Delta]$. This completes the proof.

References:

Arellano, M., 1987, Computing robust standard errors for within-group estimators, Oxford Bulletin of Economics and Statistics, 49, 431-434.

Bai, J., 2009, Panel data models with interactive effects. Econometrica, 77, 1229-1279.

Bai, J. and S. Ng, 2002, Determining the number of factors in approximate factor models. Econometrica, 70, 191-221.

Chamberlain, G., 1982, Multivariate regression models for panel data. Journal of Econometrics, 18, 5-46.

Doukhan, P., 1994, Mixing: properties and examples. Springer-Verlag, New York.

Greenaway-McGrevy, R., Han, C., Sul, D., 2012, Asymptotic distribution of factor augmented estimators for panel regression. Journal of Econometrics, 169, 48–53.

Hsiao, C., Pesaran, M. H., 2008. Random coefficient models. In: Matyas, L., Sevestre, P., eds. The Econometrics of Panel Data. Vol. 46. Berlin: Springer Berlin Heidelberg.

Hansen, C.B., 2007, Asymptotic properties of a robust variance matrix estimator for panel data with T is large. Journal of Econometrics, 141, 597-620.

Kapetanios, G., Pesaran, M. H., 2005, Alternative approaches to estimation and inference in large multifactor panels: Small sample results with an application to modelling of asset returns. CESifoWorking Paper Series 1416.

Norkute, M., Sarafidis, V., Yamagata, T., 2018, Instrumental variable estimation of dynamic linear panel data models with defactored regressors and a multifactor error structure. ISER, Osaka University.

Pesaran, M. H., 2006, Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica, 74, 967–1012.

Pesaran, M. H., Smith, Ron, 1995, Estimating long-run relationships from dynamic heterogeneous panels. Journal of Econometrics, 68, 79–113.

Pesaran, M.H., Yamagata, T., 2008, Testing slope homogeneity in large panels. Journal of Econometrics, 142, 50-93.

Reese, S. and J. Westerlund, 2018, Estimation of factor-augmented panel regressions with weakly influential factors. Econometric Reviews, 37, 401-465.

Su, L., Chen, Q., 2013, Testing homogeneity in panel data models with interactive fixed effects. Econometric Theory, 29, 1079–1135.

Westerlund, J. and J. Urbain, 2015, Cross-sectional averages versus principal components. Journal of Econometrics, 185, 372-377.

Wooledridge, J., 2010, Econometric Analysis of Cross Section and Panel Data, MIT press.

Supplementary Appendix

for

"A robust approach to heteroskedasticity, error serial correlation and slope heterogeneity for large linear panel data models with interactive effects"

by K. Hayakawa, S. Nagata and T. Yamagata

Appendix C: Proof of Lemmas for Section 3

In what follows, we repeatedly use Cauchy-Schwarz inequality, triangular inequality, Minkowski in equality, Holder's inequality, and other well-established results: for conformable matrices $\mathbf{ABC}$, $vec(\mathbf{ABC}) = (\mathbf{C}' \otimes \mathbf{A})\, vec(\mathbf{B})$, $E\|\mathbf{A}\otimes\mathbf{B}\|^s \le \left(E\|\mathbf{A}\|^{2s} E\|\mathbf{B}\|^{2s}\right)^{1/2}$, for square matrices, $\|\mathbf{AB}\| \le \|\mathbf{A}\|\,\mu_{\max}\|\mathbf{B}\|$.

C1: Proof of Lemmas for Section 2

Proof of Lemma A.4. $E\left\|\bar{\mathbf{A}}_{iT}\right\|^{1+\delta} = E\left\|T^{-1}\mathbf{X}_i'\mathbf{X}_i\right\|^{1+\delta} \le E\left\|T^{-1/2}\mathbf{X}_i\right\|^{2+2\delta} \le E\left|tr\left(T^{-1}\mathbf{X}_i'\mathbf{X}_i\right)\right|^{1+\delta}$ $= T^{-(1+\delta)}\left[\left(E\left|\sum_{h=1}^{k}\sum_{t=1}^{T}x_{ith}^2\right|^{1+\delta}\right)^{\frac{1}{1+\delta}}\right]^{1+\delta} \le T^{-(1+\delta)}\left[\sum_{h=1}^{k}\sum_{t=1}^{T}\left(E\left|x_{ith}^2\right|^{1+\delta}\right)^{\frac{1}{1+\delta}}\right]^{1+\delta} \le k^{1+\delta}\Delta < \infty$ using Holder's and Minkowski's inequality and Assumption A2, then applying Lemma A.1 gives $\bar{\mathbf{A}}_{NT} - \mathbf{A} \to_p \mathbf{0}$. Applying continuous mapping theorem yields $\bar{\mathbf{A}}_{NT}^{-1} - \mathbf{A}^{-1} \to_p \mathbf{0}$. ■

Proof of Lemma A.5. We have

$$E\left\|T^{-1/2}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\right\|^{2+2\delta} \le E\left|\sum_{h=1}^{k}\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}x_{ith}\varepsilon_{it}\right|^2\right|^{1+\delta} \le \left[\sum_{h=1}^{k}\left(E\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}x_{ith}\varepsilon_{it}\right|^{2+2\delta}\right)^{\frac{1}{1+\delta}}\right]^{1+\delta} \le k^{1+\delta}\left(T^{-\frac{2+2\delta}{2}}C\ D(s,\delta,T)\right) \le \Delta < \infty, \tag{C.1}$$

where the third inequality follows, because, by Assumption A1, $E(x_{ith}\varepsilon_{it}) = 0$, $E|x_{ith}\varepsilon_{it}|^{s+\delta} \le E|x_{ith}|^{2s+2\delta} E|\varepsilon_{it}|^{2s+2\delta} \le 2\Delta^{2s+2\delta} < \infty$ for $s > 2$ and all $h = 1, ..., k$, and using Lemma A.3 $E\left|\sum_{t=1}^{T}x_{ith}\varepsilon_{it}\right|^{2+2\delta} = C\ D(s,\delta,T) = O\left(T^{\frac{2+2\delta}{2}}\right)$. Therefore $E\left\|T^{-1/2}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\right\|^{2+2\delta} \le \Delta$ and together with Assumption A3, applying Lemma A.2 the result follows. ■

Proof of Lemma A.6. We write

$$N^{-1}\sum_{i=1}^{N}\hat{\mathbf{B}}_{i,T} = (NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{X}_i - (NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{X}_i\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\boldsymbol{\varepsilon}_i'\mathbf{X}_i \tag{C.2}$$

$$- (NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbf{X}_i'\mathbf{X}_i + (NT)^{-1}\sum_{i=1}^{N}\mathbf{X}_i'\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)'\mathbf{X}_i$$

$$= \mathbf{D}_1 - \mathbf{D}_2 - \mathbf{D}_3 + \mathbf{D}_4. \tag{C.3}$$

First

$$vec(\mathbf{D}_3) = N^{-1}T^{-1/2}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\otimes\frac{\mathbf{X}_i'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right)\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right). \tag{C.4}$$

$E\left(\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\otimes\frac{\mathbf{X}_i'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right)=\mathbf{0}$ by Assumptions A1 and A2. Noting

$$\begin{aligned} E\left\|T^{-1}\mathbf{X}_i'\mathbf{X}_i\right\|^{2+2\delta} &\leq E\left\|T^{-1/2}\mathbf{X}_i\right\|^{4+4\delta}=E\left|tr\left(T^{-1}\mathbf{X}_i'\mathbf{X}_i\right)\right|^{2+2\delta}=T^{-(1+\delta)}\left[\left(E\left|\sum_{h=1}^{k}\sum_{t=1}^{T}x_{ith}^2\right|^{2+2\delta}\right)^{\frac{1}{2+2\delta}}\right]^{2+2\delta} \\ &\leq T^{-(2+2\delta)}\left[\sum_{h=1}^{k}\sum_{t=1}^{T}\left(E\left|x_{ith}^4\right|^{1+\delta}\right)^{\frac{1}{2+2\delta}}\right]^{2+2\delta}\leq k^{2+2\delta}\Delta<\infty \qquad \text{(C.5)} \end{aligned}$$

and $E\left\|\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\otimes\frac{\mathbf{X}_i'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|^{1+\delta}=\left(E\left\|\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\right\|^{2+2\delta}E\left\|\frac{\mathbf{X}_i'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|^{2+2\delta}\right)^{1/2}\leq\Delta$ and by Lemma A.1 $N^{-1}\sum_{i=1}^{N}\left(\frac{\mathbf{X}_i'\mathbf{X}_i}{T}\otimes\frac{\mathbf{X}_i'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right)=o_p(1)$ and together with $\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}=O_p\left(1/\sqrt{NT}\right)$, $vec(\mathbf{D}_3)=o_p\left(N^{-1/2}T^{-1}\right)$. In a similar manner, $vec(\mathbf{D}_2)=o_p\left(N^{-1/2}T^{-1}\right)$. $\|\mathbf{D}_4\|\leq N^{-1}\sum_{i=1}^{N}\left\|T^{-1}\mathbf{X}_i'\mathbf{X}_i\right\|\left\|\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right\|^2=O_p\left(N^{-1}T^{-1}\right)$. $E\|\mathbf{D}_1\|^{1+\delta}=E\left\|T^{-1}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{X}_i\right\|^{1+\delta}\leq E\left\|T^{-1/2}\mathbf{X}_i'\boldsymbol{\varepsilon}_i\right\|^{2+2\delta}=O(1)$ by (C.1), and we apply Lemma A.1 to conclude $p\lim_{N,T\to\infty}N^{-1}\sum_{i=1}^{N}\left(\hat{\mathbf{B}}_{i,T}-\mathbf{B}_{i,T}\right)=\mathbf{0}$ as required. ■

Proof of Lemma A.7. First $E\left(T^{-1}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i\right)=\mathbf{0}$ and $Var\left(T^{-1}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i\right)=\mathbf{C}_{iT}$.

$$\begin{aligned} E\left\|T^{-1}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i\right\|^{2+2\delta} &\leq T^{-(2+2\delta)}E\left|\sum_{h=1}^{k}\left|\sum_{\ell=1}^{k}\sum_{t=1}^{T}x_{ith}x_{it\ell}\eta_{i\ell}\right|^2\right|^{1+\delta} \\ &\leq T^{-(2+2\delta)}\left|\sum_{h=1}^{k}\left(E\left|\sum_{\ell=1}^{k}\sum_{t=1}^{T}x_{ith}x_{it\ell}\eta_{i\ell}\right|^{2+2\delta}\right)^{\frac{1}{1+\delta}}\right|^{1+\delta} \end{aligned}$$

but as

$$E\left|\sum_{\ell=1}^{k}\sum_{t=1}^{T}x_{ith}x_{it\ell}\eta_{i\ell}\right|^{2+2\delta}\leq\left[\sum_{\ell=1}^{k}\sum_{t=1}^{T}\left(E\left|x_{ith}x_{it\ell}\eta_{i\ell}\right|^{2+2\delta}\right)^{\frac{1}{2+2\delta}}\right]^{2+2\delta}$$

$$\begin{aligned} E\left|x_{ith}x_{it\ell}\eta_{i\ell}\right|^{2+2\delta} &\leq \left(E\left|x_{ith}x_{it\ell}\right|^{4+4\delta}E\left|\eta_{i\ell}\right|^{4+4\delta}\right)^{1/2} \\ &\leq \left(E\left|x_{ith}\right|^{8+8\delta}E\left|x_{it\ell}\right|^{8+8\delta}\right)^{1/4}\left(E\left|\eta_{i\ell}\right|^{4+4\delta}\right)^{1/2} \\ &\leq \Delta \end{aligned}$$

we have

$$\begin{aligned} E\left\|T^{-1}\mathbf{X}_i'\mathbf{X}_i\boldsymbol{\eta}_i\right\|^{2+2\delta} &\leq T^{-(2+2\delta)}\left|\sum_{h=1}^{k}\left(\left[\sum_{\ell=1}^{k}\sum_{t=1}^{T}\left(E\left|x_{ith}x_{it\ell}\eta_{i\ell}\right|^{2+2\delta}\right)^{\frac{1}{2+2\delta}}\right]^{2+2\delta}\right)^{\frac{1}{1+\delta}}\right|^{1+\delta} \\ &\leq T^{-(2+2\delta)}\left|\sum_{h=1}^{k}\left(\left[\sum_{\ell=1}^{k}\sum_{t=1}^{T}(\Delta)^{\frac{1}{2+2\delta}}\right]^{2+2\delta}\right)^{\frac{1}{1+\delta}}\right|^{1+\delta} \\ &\leq T^{-(2+2\delta)}k^{1+\delta}\Delta\left[kT\right]^{2+2\delta}=O(1) \qquad \text{(C.6)} \end{aligned}$$

Applying Lemma A.2 the required result follows. ■

C2: Proof of Lemmas for Section 3

Proof of Lemma B.1. We only discuss the proof for $\ell=1$, since for $\ell=2,...,k+1$, the proof is identical to that for Lemma 1 in Bai and Ng (2002). For (i), recalling that $\mathbf{G}_i^0=\left(\boldsymbol{\Gamma}_i^{0\prime}\boldsymbol{\beta}_i+\boldsymbol{\lambda}_i^0,\boldsymbol{\Gamma}_i^{0\prime}\right)$, $\mathbf{E}_i=$

$(\mathbf{V}_i\boldsymbol{\beta}_i+\boldsymbol{\varepsilon}_i,\mathbf{V}_i)$, denoting $\rho_\ell(t,s)=\sigma_{\ell N}(s,t)/[\sigma_{\ell N}(t,t)\,\sigma_{\ell N}(s,s)]^{1/2}$ with $\sigma_{\ell N}(s,t)=N^{-1}\sum_{i=1}^N E(e_{it\ell}e_{is\ell})$ and noting that $\sigma_{\ell N}(t,t)\leq\Delta$

$$\begin{aligned}T^{-1}\sum_{s=1}^T\sum_{t=1}^T[\sigma_{\ell N}(s,t)]^2 &= T^{-1}\sum_{s=1}^T\sum_{t=1}^T\sigma_{\ell N}(t,t)\,\sigma_{\ell N}(s,s)\,[\rho_\ell(t,s)]^2\\ &\leq \Delta T^{-1}\sum_{s=1}^T\sum_{t=1}^T|\sigma_{\ell N}(t,t)\,\sigma_{\ell N}(s,s)|^{1/2}\,|\rho_\ell(t,s)|\\ &\leq \Delta T^{-1}\sum_{s=1}^T\sum_{t=1}^T|\sigma_{\ell N}(s,t)|\leq\Delta^2\end{aligned}$$

by Assumptions B1 and B2. For (iii), noting that $e_{it1}=\mathbf{v}_{it}'\boldsymbol{\beta}_i+\varepsilon_{it}$ and $\mathbf{g}_{i1}^0=\Gamma_i^{0\prime}\boldsymbol{\beta}_i+\boldsymbol{\lambda}_i^0$,

$$\begin{aligned}E\left(\left\|\frac{1}{\sqrt{NT}}\sum_{t=1}^T\sum_{i=1}^N e_{it1}\mathbf{g}_{i1}^0\right\|^2\right) &= E\left(\frac{1}{NT}\sum_t\sum_s\sum_i\sum_j e_{it1}e_{js1}\mathbf{g}_{i1}^{0\prime}\mathbf{g}_{j1}^0\right)\\ &= \frac{1}{NT}\sum_t\sum_s\sum_i E\{[\boldsymbol{\beta}_i'\mathbf{v}_{it}\mathbf{v}_{is}'\boldsymbol{\beta}_i+\varepsilon_{it}\varepsilon_{is}]\,[\boldsymbol{\beta}_i'\Gamma_i^0\Gamma_i^{0\prime}\boldsymbol{\beta}_i+\boldsymbol{\lambda}_i^{0\prime}\boldsymbol{\lambda}_i^0]\}\\ &= \frac{1}{NT}\sum_t\sum_s\sum_i E\left[\left(\boldsymbol{\beta}_i'\mathbf{v}_{it}\mathbf{v}_{is}'\boldsymbol{\beta}_i\right)\left(\boldsymbol{\beta}_i'\Gamma_i^0\Gamma_i^{0\prime}\boldsymbol{\beta}_i\right)\right]+E\left(\boldsymbol{\beta}_i'\mathbf{v}_{it}\mathbf{v}_{is}'\boldsymbol{\beta}_i\right)E\left(\boldsymbol{\lambda}_i^{0\prime}\boldsymbol{\lambda}_i^0\right)\\ &\quad +E(\varepsilon_{it}\varepsilon_{is})\,E\left(\boldsymbol{\beta}_i'\Gamma_i^0\Gamma_i^{0\prime}\boldsymbol{\beta}_i\right)+E(\varepsilon_{it}\varepsilon_{is})\,E\left(\boldsymbol{\lambda}_i^{0\prime}\boldsymbol{\lambda}_i^0\right)\\ &= a_1+a_2+a_3+a_4.\end{aligned}$$

as

$$\left(\boldsymbol{\beta}_i'\mathbf{v}_{it}\mathbf{v}_{is}'\boldsymbol{\beta}_i\right)\left(\boldsymbol{\beta}_i'\Gamma_i^0\Gamma_i^{0\prime}\boldsymbol{\beta}_i\right)=vec\left(\Gamma_i^0\Gamma_i^{0\prime}\right)'\left[\boldsymbol{\beta}_i\boldsymbol{\beta}_i'\otimes\boldsymbol{\beta}_i\boldsymbol{\beta}_i'\right]vec\left(\mathbf{v}_{it}\mathbf{v}_{is}'\right)$$

$$\begin{aligned}a_1 &= \frac{1}{NT}\sum_t\sum_s\sum_i vec\left(E\left(\Gamma_i^0\Gamma_i^{0\prime}\right)\right)'E\left[\boldsymbol{\beta}_i\boldsymbol{\beta}_i'\otimes\boldsymbol{\beta}_i\boldsymbol{\beta}_i'\right]vec\left(E\left(\mathbf{v}_{it}\mathbf{v}_{is}'\right)\right)\\ &\leq \frac{1}{N}\sum_i\left\|\frac{1}{T}\sum_t\sum_s E\left(\mathbf{v}_{it}\mathbf{v}_{is}'\right)\right\|\Delta\left\|E\left(\Gamma_i^0\Gamma_i^{0\prime}\right)\right\|\leq\Delta.\end{aligned}$$

Similarly

$$a_2=\frac{1}{N}\sum_i tr\left[E\left(\boldsymbol{\beta}_i\boldsymbol{\beta}_i'\right)\left(\frac{1}{T}\sum_t\sum_s E\left(\mathbf{v}_{it}\mathbf{v}_{is}'\right)\right)\right]E\left(\boldsymbol{\lambda}_i^{0\prime}\boldsymbol{\lambda}_i^0\right)\leq\Delta,$$

$a_3\leq\Delta$ and $a_4\leq\Delta$. (ii) is shown in a similar manner. ■

Proof of Lemma B.2. Following the discussion in Bai (2009), p.1266, we write

$$\hat{\mathbf{F}}\boldsymbol{\Xi}_{NT}=\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^N\mathbf{F}^0\boldsymbol{g}_{\ell i}^0\mathbf{e}_{\ell i}'\hat{\mathbf{F}}+\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^N\mathbf{e}_{\ell i}\boldsymbol{g}_{\ell i}^{0\prime}\mathbf{F}^{0\prime}\hat{\mathbf{F}} \tag{C.7}$$

$$\begin{aligned}&+\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^N\mathbf{e}_{\ell i}\mathbf{e}_{\ell i}'\hat{\mathbf{F}}+\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^N\mathbf{F}^0\boldsymbol{g}_{\ell i}^0\boldsymbol{g}_{\ell i}^{0\prime}\mathbf{F}^{0\prime}\hat{\mathbf{F}}\\ &=\mathbf{A}_1+\mathbf{A}_2+\mathbf{A}_3+\mathbf{A}_4.\end{aligned} \tag{C.8}$$

Observing that $\mathbf{A}_4=\mathbf{F}^0\boldsymbol{\Upsilon}_N^0\boldsymbol{\Lambda}_{0\hat{F}}$, we have

$$\hat{\mathbf{F}}\boldsymbol{\Xi}_{NT}-\mathbf{F}^0\boldsymbol{\Upsilon}_N^0\boldsymbol{\Lambda}_{0\hat{F}}=\mathbf{A}_1+\mathbf{A}_2+\mathbf{A}_3. \tag{C.9}$$

Post-multiplying the above equation by $\hat{\mathbf{Q}}=\left(\boldsymbol{\Upsilon}_N^0\boldsymbol{\Lambda}_{0\hat{F}}\right)^{-1}$ yields

$$\hat{\mathbf{F}}\boldsymbol{\Xi}_{NT}\hat{\mathbf{Q}}-\mathbf{F}^0=(\mathbf{A}_1+\mathbf{A}_2+\mathbf{A}_3)\,\hat{\mathbf{Q}}. \tag{C.10}$$

Let

$$\mathbf{R} = \left(\Xi_{NT}\hat{\mathbf{Q}}\right)^{-1}, \tag{C.11}$$

where Ξ_{NT} is assumed to be invertible (the invertibility of Ξ_{NT} is proved in Bai 2009, p.1267) so that

$$\begin{aligned}\hat{\mathbf{F}}\mathbf{R}^{-1} - \mathbf{F}^0 &= \frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\mathbf{F}^0\boldsymbol{g}^0_{\ell i}\mathbf{e}'_{\ell i}\hat{\mathbf{F}}\hat{\mathbf{Q}} + \frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\mathbf{e}_{\ell i}\boldsymbol{g}^{0\prime}_{\ell i}\mathbf{F}^{0\prime}\hat{\mathbf{F}}\hat{\mathbf{Q}} \\ &+ \frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\mathbf{e}_{\ell i}\mathbf{e}'_{\ell i}\hat{\mathbf{F}}\hat{\mathbf{Q}}, \end{aligned} \tag{C.12}$$

as required. ■

Proof of Lemma B.3. By Lemma B.2 we have

$$T^{-1/2}\left\|\hat{\mathbf{F}}\mathbf{R}^{-1} - \mathbf{F}^0\right\| \leq T^{-1/2}\left(\|\mathbf{A}_1\| + \|\mathbf{A}_2\| + \|\mathbf{A}_3\|\right)\left\|\hat{\mathbf{Q}}\right\|,$$

where $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{A}_3$ are defined in (C.7).

$$\begin{aligned} T^{-1/2}\|\mathbf{A}_1\| &= \frac{1}{\sqrt{T}}\left\|\frac{1}{NT}\sum_{i=1}^{N}\mathbf{F}^0\boldsymbol{G}^0_i\mathbf{E}'_i\hat{\mathbf{F}}\right\| \\ &\leq \frac{1}{\sqrt{N}}\left\|\frac{\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{G}^0_i\mathbf{E}'_i\right\|\left\|\frac{\hat{\mathbf{F}}}{\sqrt{T}}\right\| \\ &= O_p\left(\frac{1}{\sqrt{N}}\right) \end{aligned}$$

by Lemma B.1 (iii). In a similar manner,

$$\begin{aligned} T^{-1/2}\|\mathbf{A}_2\| &= \frac{1}{\sqrt{T}}\left\|\frac{1}{NT}\sum_{i=1}^{N}\mathbf{E}_i\boldsymbol{G}^{0\prime}_i\mathbf{F}^{0\prime}\hat{\mathbf{F}}\right\| \\ &\leq \frac{1}{\sqrt{N}}\left\|\frac{\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{G}^0_i\mathbf{E}'_i\right\|\left\|\frac{\hat{\mathbf{F}}}{\sqrt{T}}\right\| \\ &= O_p\left(\frac{1}{\sqrt{N}}\right), \end{aligned}$$

$$\begin{aligned} T^{-1/2}\mathbf{A}_3 &= T^{-1/2}\frac{1}{T}\bar{\Sigma}_{EE}\hat{\mathbf{F}} + T^{-1/2}\frac{1}{T}\left(\Sigma_{EE} - \bar{\Sigma}_{EE}\right)\hat{\mathbf{F}} \\ &= \mathbf{A}_{31} + \mathbf{A}_{32} \end{aligned}$$

where

$$\Sigma_{EE} = \frac{1}{N}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\mathbf{e}_{i\ell}\mathbf{e}'_{i\ell},\ \bar{\Sigma}_{EE} = \frac{1}{N}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}E\left(\mathbf{e}_{i\ell}\mathbf{e}'_{i\ell}\right). \tag{C.13}$$

$$\|\mathbf{A}_{32}\| \leq \frac{1}{\sqrt{NT}}\left\|\frac{\sqrt{N}\left(\Sigma_{EE} - \bar{\Sigma}_{EE}\right)}{\sqrt{T}}\right\|\left\|\frac{\hat{\mathbf{F}}}{\sqrt{T}}\right\| = O_p\left(\frac{1}{\sqrt{NT}}\right)$$

by Assumption B1&B2, and by Lemma B.7,

$$\|\mathbf{A}_{31}\| \leq T^{-1}\mu_{\max}\left(\bar{\Sigma}_{EE}\right)\left\|\frac{\hat{\mathbf{F}}}{\sqrt{T}}\right\| = O_p\left(\frac{1}{T}\right).$$

Noting that $\left\|\hat{\mathbf{Q}}\right\| \leq \Delta$,

$$T^{-1/2}\left\|\hat{\mathbf{F}}\mathbf{R} - \mathbf{F}^0\right\| = O_p\left(\frac{1}{\sqrt{T}}\right) + O_p\left(\frac{1}{\sqrt{N}}\right) = O_p\left(\delta_{NT}^{-1}\right)$$

as required. ■

Proof of Lemma B.4. Substituting $\mathbf{F}^0 = \left(\mathbf{F}^0 - \hat{\mathbf{F}}\mathbf{R}^{-1}\right) + \hat{\mathbf{F}}\mathbf{R}^{-1}$ to $\hat{\mathbf{G}}_i = T^{-1}\hat{\mathbf{F}}'\mathbf{Z}_i^* = T^{-1}\hat{\mathbf{F}}'\mathbf{F}^0\mathbf{G}_i^0 + T^{-1}\hat{\mathbf{F}}'\mathbf{E}_i$ gives

$$\begin{aligned} N^{-1/2}\left\|\sum_{i=1}^N \hat{\mathbf{G}}_i - \mathbf{R}^{-1}\mathbf{G}_i^0\right\| &\leq N^{-1/2}\left\|T^{-1}\hat{\mathbf{F}}'\left[\left(\mathbf{F}^0 - \hat{\mathbf{F}}\mathbf{R}^{-1}\right)\right]\sum_{i=1}^N \mathbf{G}_i^0 + \sum_{i=1}^N T^{-1}\hat{\mathbf{F}}'\mathbf{E}_i\right\| \\ &\leq \sqrt{r}\left\|T^{-1/2}\left(\mathbf{F}^0 - \hat{\mathbf{F}}\mathbf{R}^{-1}\right)\right\| \\ &\quad + \frac{1}{\sqrt{T}}\left\|N^{-1/2}\sum_{i=1}^N T^{-1/2}\mathbf{F}^{0\prime}\mathbf{E}_i\right\| + \frac{1}{\sqrt{T}}\left\|N^{-1/2}\sum_{i=1}^N T^{-1/2}\left(\hat{\mathbf{F}} - \mathbf{F}^0\mathbf{R}\right)'\mathbf{E}_i\right\| \\ &= O_p\left(\delta_{NT}^{-1}\right). \end{aligned}$$

■

Proof of Lemmas B.5-B.9 is obtained in line with the discussions in Bai (2009) under Assumptions B1-B5, which are omitted. See derivations therein.

Proof of Lemma B.10. $\left\|\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{X}_i^*\right\| \leq \left\|\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right\|\frac{1}{N}\sum_{i=1}^N\left\|\frac{\mathbf{X}_i^*}{\sqrt{T}}\right\|^2 = O_p\left(\delta_{NT}^{-1}\right).$ ■

Noting that $\mathbf{X}_i^* = \mathbf{F}^0\boldsymbol{\Gamma}_i^{0\prime} + \mathbf{V}_i$ and $\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{u}_i$, $\mathbf{u}_i = \mathbf{X}_i\boldsymbol{\eta}_i + \boldsymbol{\epsilon}_{f\varepsilon,i}$, $\boldsymbol{\epsilon}_{f\varepsilon,i} = \mathbf{F}^0\boldsymbol{\lambda}_i^0 + \boldsymbol{\varepsilon}_i$, we have

$$\begin{aligned} \sum_{i=1}^N \hat{\mathbf{X}}'\mathbf{u}_i &= \sum_{i=1}^N \mathbf{X}^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{u}_i \\ &= \sum_{i=1}^N \mathbf{X}^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*\boldsymbol{\eta}_i + \sum_{i=1}^N \boldsymbol{\Gamma}_i^0\mathbf{F}^{0\prime}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} + \sum_{i=1}^N \mathbf{V}_i'\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i}. \end{aligned}$$

Proof of Lemma B.11.

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{X}_i^{*\prime}\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*\boldsymbol{\eta}_i = \frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{X}_i^{*\prime}\mathbf{M}_{\mathbf{F}}\mathbf{X}_i^*\boldsymbol{\eta}_i + \frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{X}_i^*\boldsymbol{\eta}_i$$

$$\begin{aligned} \frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{X}_i^{*\prime}\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{X}_i^*\boldsymbol{\eta}_i &= \frac{1}{\sqrt{NT}}\sum_{i=1}^N \left(\mathbf{F}\boldsymbol{\Gamma}_i' + \mathbf{V}_i\right)'\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\left(\mathbf{F}\boldsymbol{\Gamma}_i' + \mathbf{V}_i\right)\boldsymbol{\eta}_i \\ &= \frac{1}{\sqrt{NT}}\sum_{i=1}^N \boldsymbol{\Gamma}_i\mathbf{F}'\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{F}\boldsymbol{\Gamma}_i'\boldsymbol{\eta}_i \\ &\quad + \frac{1}{\sqrt{NT}}\sum_{i=1}^N \boldsymbol{\Gamma}_i\mathbf{F}'\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{V}_i\boldsymbol{\eta}_i \\ &\quad + \frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{V}_i'\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{F}\boldsymbol{\Gamma}_i'\boldsymbol{\eta}_i \\ &\quad + \frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{V}_i'\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{V}_i\boldsymbol{\eta}_i. \\ &= \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3 + \mathbf{a}_4. \end{aligned}$$

$$\begin{aligned} \frac{1}{\sqrt{NT}}\sum_{i=1}^N \boldsymbol{\Gamma}_i\mathbf{F}'\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{F}\boldsymbol{\Gamma}_i'\boldsymbol{\eta}_i &= \frac{1}{\sqrt{NT}}\sum_{i=1}^N \left(\boldsymbol{\eta}_i'\boldsymbol{\Gamma}_i' \otimes \boldsymbol{\Gamma}_i\right) vec\left(\mathbf{F}'\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\mathbf{F}\right) \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^N \left(\boldsymbol{\eta}_i'\boldsymbol{\Gamma}_i' \otimes \boldsymbol{\Gamma}_i\right)\left(\frac{\mathbf{F}'\otimes\mathbf{F}'}{T}\right) vec\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right) \end{aligned}$$

Since $(vec(\mathbf{A}))' vec(\mathbf{A}) = tr(\mathbf{A}'\mathbf{A})$ and $\|vec(\mathbf{A})\| = \|\mathbf{A}\|$, and $tr(\mathbf{A} \otimes \mathbf{B}) = tr(\mathbf{A}) tr(\mathbf{B})$ and $(\mathbf{A} \otimes \mathbf{B})' (\mathbf{A} \otimes \mathbf{B}) = (\mathbf{A}' \otimes \mathbf{B}')(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{A}'\mathbf{A} \otimes \mathbf{B}'\mathbf{B})$

$$\begin{aligned}
\|\mathbf{a}_1\| &\le \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (\boldsymbol{\eta}_i' \boldsymbol{\Gamma}_i' \otimes \boldsymbol{\Gamma}_i) \right\| \left\| \frac{\mathbf{F}' \otimes \mathbf{F}'}{T} \right\| \left\| vec\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right) \right\| \\
&= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (\boldsymbol{\eta}_i' \boldsymbol{\Gamma}_i' \otimes \boldsymbol{\Gamma}_i) \right\| \left\| \frac{\mathbf{F}'\mathbf{F}}{T} \right\|^2 \left\| \mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}} \right\| \\
&= O_p\left(\delta_{NT}^{-1}\right).
\end{aligned}$$

as $\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (\boldsymbol{\eta}_i' \boldsymbol{\Gamma}_i' \otimes \boldsymbol{\Gamma}_i) \right\| = O_p(1)$ and $\left\| \mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}} \right\| = O_p\left(\delta_{NT}^{-1}\right)$.

$$\begin{aligned}
\mathbf{a}_2 &= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_i \mathbf{F}' \left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right) \mathbf{V}_i \boldsymbol{\eta}_i = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\boldsymbol{\eta}_i' \frac{\mathbf{V}_i'}{\sqrt{T}} \otimes \boldsymbol{\Gamma}_i \right) vec\left(\mathbf{F}' \left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)\right) \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\boldsymbol{\eta}_i' \frac{\mathbf{V}_i'}{\sqrt{T}} \otimes \boldsymbol{\Gamma}_i \right) \left(\frac{\mathbf{I}_T \otimes \mathbf{F}'}{\sqrt{T}} \right) vec\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right)
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{a}_2\| &\le \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\boldsymbol{\eta}_i' \frac{\mathbf{V}_i'}{\sqrt{T}} \otimes \boldsymbol{\Gamma}_i \right) \right\| \left\| \frac{\mathbf{F}'}{\sqrt{T}} \right\| \left\| \mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}} \right\| \\
&= O_p\left(\delta_{NT}^{-1}\right).
\end{aligned}$$

Next,

$$\begin{aligned}
\|\mathbf{a}_3\| &= \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i' \left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right) \mathbf{F} \boldsymbol{\Gamma}_i' \boldsymbol{\eta}_i \right\| = \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} (\boldsymbol{\eta}_i' \boldsymbol{\Gamma}_i' \otimes \mathbf{V}_i') \, vec\left(\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right) \mathbf{F}\right) \right\| \\
&= \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} (\boldsymbol{\eta}_i' \boldsymbol{\Gamma}_i' \otimes \mathbf{V}_i') \left(\frac{\mathbf{F}' \otimes \mathbf{I_T}}{\sqrt{T}} \right) vec\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right) \right\| \\
&\le \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\eta}_i' \boldsymbol{\Gamma}_i' \otimes \frac{\mathbf{V}_i'}{\sqrt{T}} \right\| \left\| \frac{\mathbf{F}'}{\sqrt{T}} \right\| \left\| \mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}} \right\| = O_p\left(\delta_{NT}^{-1}\right).
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{a}_4\| &= \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i' \left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right) \mathbf{V}_i \boldsymbol{\eta}_i \right\| = \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{\boldsymbol{\eta}_i' \mathbf{V}_i' \otimes \mathbf{V}_i'}{T} \right) vec\left(\mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}}\right) \right\| \\
&\le \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{\boldsymbol{\eta}_i' \mathbf{V}_i' \otimes \mathbf{V}_i'}{T} \right) \right\| \left\| \mathbf{M}_{\hat{\mathbf{F}}} - \mathbf{M}_{\mathbf{F}} \right\| = O_p\left(\delta_{NT}^{-1}\right).
\end{aligned}$$

Using the above results, the required result follows. ∎

Proof of Lemma B.12.

$$\begin{aligned}
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_i^0 \mathbf{F}^{0\prime} \mathbf{M}_{\hat{\mathbf{F}}} \boldsymbol{\epsilon}_{f\varepsilon i} &= -\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_i^0 \left(\hat{\mathbf{F}} \mathbf{R}^{-1} - \mathbf{F}^0 \right)' \mathbf{M}_{\hat{\mathbf{F}}} \boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= -\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_i^0 \left(\frac{1}{NT} \sum_{\ell=1}^{k+1} \sum_{i=1}^{N} \hat{\mathbf{Q}}' \hat{\mathbf{F}}' \mathbf{e}_{\ell i} \boldsymbol{g}_{\ell i}^{0\prime} \mathbf{F}^{0\prime} \right) \mathbf{M}_{\hat{\mathbf{F}}} \boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad -\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_i^0 \left(\frac{1}{NT} \sum_{\ell=1}^{k+1} \sum_{i=1}^{N} \hat{\mathbf{Q}}' \hat{\mathbf{F}}' \mathbf{F}^{\mathbf{0}} \boldsymbol{g}_{\ell \mathbf{i}}^{\mathbf{0}} \mathbf{e}_{\ell i}' \right) \mathbf{M}_{\hat{\mathbf{F}}} \boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad -\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_i^0 \left(\frac{1}{NT} \sum_{\ell=1}^{k+1} \sum_{i=1}^{N} \hat{\mathbf{Q}} \hat{\mathbf{F}} \mathbf{e}_{\ell i} \mathbf{e}_{\ell i}' \right) \mathbf{M}_{\hat{\mathbf{F}}} \boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= -\mathbf{b}_{11} - \mathbf{b}_{12} - \mathbf{b}_{13}
\end{aligned}$$

$$
\begin{aligned}
\mathbf{b}_{11} &= \sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_i^0\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\hat{\mathbf{Q}}'\hat{\mathbf{F}}'\mathbf{e}_{\ell i}\boldsymbol{g}_{\ell i}^{0\prime}\left(\mathbf{R}'\right)^{-1}\mathbf{R}'\mathbf{F}^{0\prime}\right)\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= -\sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i}
\end{aligned}
$$

with

$$
\mathbf{A}_{mNT} = \frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\hat{\mathbf{F}}'\mathbf{e}_{\ell i}\boldsymbol{g}_{\ell i}^{0\prime}\left(\mathbf{R}'\right)^{-1}
$$

but

$$
\begin{aligned}
\left\|\mathbf{A}_{mNT}\right\| &\leq \left\|\frac{1}{\sqrt{T}}\frac{1}{N}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\mathbf{R}\frac{\mathbf{F}^{0\prime}\mathbf{e}_{\ell i}\boldsymbol{g}_{\ell i}^{0\prime}}{\sqrt{T}}\left(\mathbf{R}'\right)^{-1}\right\| + \left\|\frac{1}{N}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\frac{\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{e}_{\ell i}}{T}\boldsymbol{g}_{\ell i}^{0\prime}\left(\mathbf{R}'\right)^{-1}\right\| \\
&\leq \frac{1}{\sqrt{T}}\frac{1}{N}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\left\|\frac{\mathbf{F}^{0\prime}\mathbf{e}_{\ell i}\boldsymbol{g}_{\ell i}^{0\prime}}{\sqrt{T}}\right\| + \frac{1}{N}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\left\|\frac{\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{e}_{\ell i}}{T}\right\|\left\|\boldsymbol{g}_{\ell i}^{0}\right\|\left\|\left(\mathbf{R}'\right)^{-1}\right\| \\
&= O_p\left(T^{-1/2}\right) + O_p\left(\delta_{NT}^{-2}\right).
\end{aligned}
$$

Next,

$$
\begin{aligned}
\mathbf{b}_{11} &= \sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\left(\mathbf{I}_T-\frac{\hat{\mathbf{F}}\hat{\mathbf{F}}'}{T}\right)\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad -\sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\frac{\hat{\mathbf{F}}\hat{\mathbf{F}}'}{T}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{111}-\mathbf{b}_{112}
\end{aligned}
$$

$$
\begin{aligned}
\left\|\mathbf{b}_{111}\right\| &\leq \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{F}^0\right\|\left\|\boldsymbol{\lambda}_i^0\right\| \\
&\quad +\sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\boldsymbol{\varepsilon}_i\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-2}\right)\left[O_p\left(T^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
\left\|\mathbf{b}_{112}\right\| &\leq \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\hat{\mathbf{F}}\right\|\left\|\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right\|\left\|\boldsymbol{\lambda}_i^0\right\| \\
&\quad +\sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\hat{\mathbf{F}}\right\|\left\|\frac{\hat{\mathbf{F}}'}{\sqrt{T}}\right\|\left\|\frac{\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-2}\right)\left[O_p\left(T^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)\right]
\end{aligned}
$$

$$\begin{aligned}
\mathbf{b}_{111} &= \sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\boldsymbol{\epsilon}_{f\varepsilon,i}+\sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{F}^0\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{1111}+\mathbf{b}_{1112}
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{b}_{1111}\| &\le \sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\boldsymbol{\epsilon}_{f\varepsilon,i}\right\| \\
&\le \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\boldsymbol{\epsilon}_{f\varepsilon,i}\right\| \\
&\le \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{F}^0\right\|\left\|\boldsymbol{\lambda}_i^0\right\| \\
&\quad +\sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\boldsymbol{\varepsilon}_i\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-2}\right)\left[O_p\left(T^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)\right].
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{b}_{1112}\| &\le \sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{F}\mathbf{F}'\left(\mathbf{F}\boldsymbol{\lambda}_i+\boldsymbol{\varepsilon}_i\right)\right\| \\
&\le \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{F}^0\right\|\left\|\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right\|\left\|\boldsymbol{\lambda}_i^0\right\| \\
&\quad +\sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{F}\right\|\left\|\frac{\mathbf{F}'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-2}\right)\left[O_p\left(T^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)\right]
\end{aligned}$$

$$\begin{aligned}
\mathbf{b}_{112} &= \sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\mathbf{R}'\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad +\sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad +\sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\left[T^{-1}\mathbf{F}\mathbf{R}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad +\sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\left\{T^{-1}\mathbf{F}\left[\mathbf{R}\mathbf{R}'-\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right]\mathbf{F}'\right\}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{1121}+\mathbf{b}_{1122}+\mathbf{b}_{1123}+\mathbf{b}_{1124}
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{b}_{1121}\| &= \left\|\sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\mathbf{R}'\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i}\right\| \\
&\le \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\right\|^2\left\|\mathbf{R}'\right\|\left\|\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right\|\left\|\boldsymbol{\lambda}_i\right\| \\
&\quad +\sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\right\|^2\left\|\mathbf{R}'\right\|\left\|\frac{\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-2}\right)\left[O_p\left(T^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)\right]
\end{aligned}$$

$$
\begin{aligned}
\|\mathbf{b}_{1122}\| &= \left\|\sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i}\right\| \\
&\leq \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|T^{-1}\left\|\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right\|^2T^{-1}\left\|\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{F}^0\right\|\left\|\boldsymbol{\lambda}_i^0\right\| \\
&\quad+\sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|T^{-1}\left\|\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right\|^2T^{-1}\left\|\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\boldsymbol{\varepsilon}_i\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right)\left[O_p\left(T^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)\right]
\end{aligned}
$$

$$
\|\mathbf{b}_{1123}\| = \sqrt{NT}O_p\left(\delta_{NT}^{-2}\right)\left[O_p\left(T^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)\right]
$$

$$
\begin{aligned}
\|\mathbf{b}_{1124}\| &\leq \left\|\sqrt{NT}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\hat{\mathbf{Q}}'\mathbf{A}_{mNT}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\left\{T^{-1}\mathbf{F}\left[\mathbf{R}\mathbf{R}'-\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right]\mathbf{F}'\right\}\left(\mathbf{F}^0\boldsymbol{\lambda}_i+\boldsymbol{\varepsilon}_i\right)\right\| \\
&\leq \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{F}\right\|\left\|\mathbf{R}\mathbf{R}'-\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right\|\left\|\frac{\mathbf{F}'\mathbf{F}^0}{T}\right\|\left\|\boldsymbol{\lambda}_i\right\| \\
&\quad+\sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{A}_{mNT}\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{F}\right\|\left\|\mathbf{R}\mathbf{R}'-\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right\|\left\|\frac{\mathbf{F}'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right)\left[O_p\left(T^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)\right].
\end{aligned}
$$

To conclude, $\mathbf{b}_{11} = \sqrt{NT}O_p\left(\delta_{NT}^{-2}\right)\left[O_p\left(T^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)\right]$.

Noting $\hat{\mathbf{Q}} = \left(\boldsymbol{\Upsilon}_N^0\boldsymbol{\Lambda}_{0\hat{F}}\right)^{-1} = \boldsymbol{\Lambda}_{0\hat{F}}^{-1}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}$, we have

$$
\begin{aligned}
\mathbf{b}_{12} &= \frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{\ell=1}^{k+1}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{\ell j}^0\mathbf{e}_{\ell j}'\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{\ell=1}^{k+1}\boldsymbol{\Gamma}_j^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{\ell i}^0\boldsymbol{\epsilon}_{f\varepsilon,j}'\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{e}_{\ell i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\left(\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_j^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{\ell i}^0\boldsymbol{\epsilon}_{f\varepsilon,j}'\right)\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{e}_{\ell i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{e}_{\ell i}
\end{aligned}
$$

$$
\mathbf{H}_{\ell i} = \frac{1}{N}\sum_{j=1}^{N}\boldsymbol{\epsilon}_{f\varepsilon,j}\boldsymbol{g}_{\ell i}^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\Gamma}_j^{0\prime}.
$$

$$
\begin{aligned}
\mathbf{b}_{12} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'\mathbf{M}_{\mathbf{F}^0}\mathbf{e}_{\ell i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\mathbf{e}_{\ell i} \\
&= \mathbf{b}_{121}+\mathbf{b}_{122}
\end{aligned}
$$

$$\begin{aligned}
\mathbf{b}_{121} &= \frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_i^0\left(\mathbf{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\mathbf{E}_j'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_i^0\left(\mathbf{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\boldsymbol{\varepsilon}_i}{T} \\
&\quad +\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_i^0\left(\mathbf{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\mathbf{F}\mathbf{F}'\boldsymbol{\varepsilon}_i}{T} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_i^0\left(\mathbf{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\boldsymbol{\varepsilon}_i}{T}+o_p(1).
\end{aligned}$$

$$\begin{aligned}
\mathbf{b}_{122} &= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\mathbf{R}'\mathbf{F}^{0\prime}\right]\mathbf{e}_{\ell i}-\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\right]\mathbf{e}_{\ell i} \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'\left[T^{-1}\mathbf{F}\mathbf{R}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\right]\mathbf{e}_{\ell i}-\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'\left\{T^{-1}\mathbf{F}^0\left[\mathbf{R}\mathbf{R}'-\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\right]\mathbf{F}^{0\prime}\right\}\mathbf{e}_{\ell i} \\
&= \mathbf{b}_{1221}+\mathbf{b}_{1222}+\mathbf{b}_{1223}+\mathbf{b}_{1224}
\end{aligned}$$

$$\begin{aligned}
\left\|\mathbf{b}_{1221}\right\| &\leq \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\mathbf{R}'\mathbf{F}^{0\prime}\mathbf{e}_{\ell i} \\
&\leq \sqrt{N}\sum_{\ell=1}^{k+1}\left[\frac{1}{N}\sum_{j=1}^{N}\left\|\mathbf{\Gamma}_j^0\left(\mathbf{\Upsilon}_N^0\right)^{-1}\right\|\left\|\boldsymbol{\epsilon}_{f\varepsilon,j}'T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\right\|\right]\left[\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{g}_{\ell i}^0\right\|\left\|\mathbf{R}'\right\|\left\|\frac{\mathbf{F}^{0\prime}\mathbf{e}_{\ell i}}{\sqrt{T}}\right\|\right] \\
&= \sqrt{N}O_p\left(\delta_{NT}^{-2}\right)
\end{aligned}$$

$$\begin{aligned}
\left\|\mathbf{b}_{1222}\right\| &\leq \sqrt{NT}\sum_{\ell=1}^{k+1}\frac{1}{N}\sum_{j=1}^{N}\left\|\mathbf{\Gamma}_j^0\left(\mathbf{\Upsilon}_N^0\right)^{-1}\right\|\left\|\boldsymbol{\epsilon}_{f\varepsilon,j}'T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{g}_{\ell i}^0\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{e}_{\ell i}\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right)
\end{aligned}$$

$$\left\|\mathbf{b}_{1224}\right\|\leq\sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\left\|\frac{\mathbf{H}_{\ell i}'\mathbf{F}^0}{T}\right\|\left\|\mathbf{R}\mathbf{R}'-\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\right\|\left\|\frac{\mathbf{F}^{0\prime}\mathbf{e}_{\ell i}}{\sqrt{T}}\right\|=\sqrt{N}O_p\left(\delta_{NT}^{-2}\right)$$

since $\left\|\frac{\mathbf{H}_{\ell i}'\mathbf{F}^0}{T}\right\|=O_p(1)$.

$$\begin{aligned}
\mathbf{b}_{1223} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'\left[T^{-1}\mathbf{F}\mathbf{R}\mathbf{R}'\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^0\right)'\right]\mathbf{e}_{\ell i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'\left[T^{-1}\mathbf{F}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^0\right)'\right]\mathbf{e}_{\ell i} \\
&\quad +\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}_{\ell i}'\left[T^{-1}\mathbf{F}\left(\mathbf{R}\mathbf{R}'-\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\right)\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^0\right)'\right]\mathbf{e}_{\ell i} \\
&= \mathbf{b}_{12231}+\mathbf{b}_{12232}
\end{aligned}$$

$$\begin{aligned}
\left\|\mathbf{b}_{12232}\right\| &\leq \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\left\|\frac{\mathbf{H}_{\ell i}'\mathbf{F}^0}{T}\right\|\left\|\mathbf{R}\mathbf{R}'-\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\right\|\left\|\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^0\right)'\mathbf{e}_{\ell i}\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right)
\end{aligned}$$

$$\begin{aligned}
\mathbf{b}_{12231} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}'_{\ell i}T^{-1}\mathbf{F}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\hat{\mathbf{F}}'\mathbf{e}_{\ell\mathbf{j}}\boldsymbol{g}^{\mathbf{0}\prime}_{\ell\mathbf{j}}\mathbf{F}^{0^{\prime}}\mathbf{e}_{\ell i} \\
&+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}'_{\ell i}T^{-1}\mathbf{F}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\hat{\mathbf{F}}'\mathbf{F}^{\mathbf{0}}\boldsymbol{g}^{\mathbf{0}}_{\ell\mathbf{j}}\mathbf{e}_{\ell j}^{\prime}\mathbf{e}_{\ell i} \\
&+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}'_{\ell i}T^{-1}\mathbf{F}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\hat{\mathbf{F}}'\mathbf{e}_{\ell j}\mathbf{e}'_{\ell j}\mathbf{e}_{\ell i} \\
&= \mathbf{b}_{122311}+\mathbf{b}_{122312}+\mathbf{b}_{122313}
\end{aligned}$$

$$\begin{aligned}
\mathbf{b}_{122311} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}'_{\ell i}T^{-1}\mathbf{F}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\mathbf{R}'\mathbf{F}^{0\prime}\mathbf{e}_{\ell\mathbf{j}}\boldsymbol{g}^{\mathbf{0}\prime}_{\ell\mathbf{j}}\mathbf{F}^{0^{\prime}}\mathbf{e}_{\ell i} \\
&+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}'_{\ell i}T^{-1}\mathbf{F}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\mathbf{e}_{\ell\mathbf{j}}\boldsymbol{g}^{\mathbf{0}\prime}_{\ell\mathbf{j}}\mathbf{F}^{0\prime}\mathbf{e}_{\ell i} \\
&= \mathbf{b}_{1223111}+\mathbf{b}_{1223112}
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{b}_{1223111}\| &\leq \frac{1}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\left\|\frac{\mathbf{H}'_{\ell i}\mathbf{F}}{T}\right\|\left\|\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}^{0\prime}\mathbf{e}_{\ell i}}{T}\right\|\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\mathbf{R}'\frac{\mathbf{F}^{0\prime}\mathbf{e}_{\ell\mathbf{j}}}{\sqrt{T}}\boldsymbol{g}^{\mathbf{0}\prime}_{\ell\mathbf{j}}\right\| \\
&= O_p\left(T^{-1/2}\right)
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{b}_{1223112}\| &\leq \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\left\|\frac{\mathbf{H}'_{\ell i}\mathbf{F}}{T}\right\|\left\|\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}^{0\prime}\mathbf{e}_{\ell i}}{\sqrt{T}}\right\|\frac{1}{N}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\left\|\hat{\mathbf{Q}}'\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\mathbf{e}_{\ell\mathbf{j}}\right\|\left\|\boldsymbol{g}^{\mathbf{0}\prime}_{\ell\mathbf{j}}\right\| \\
&= \sqrt{N}O_p\left(\delta_{NT}^{-2}\right).
\end{aligned}$$

Noting $\hat{\mathbf{Q}}=\left(\boldsymbol{\Upsilon}_N^0\boldsymbol{\Lambda}_{0\hat{F}}\right)^{-1}=\boldsymbol{\Lambda}_{0\hat{F}}^{-1}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}$

$$\mathbf{b}_{122312}=\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\frac{\mathbf{H}'_{\ell i}\mathbf{F}}{T}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\sum_{h=1}^{k+1}\sum_{j=1}^{N}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}^{0}_{hj}\frac{\mathbf{e}'_{hj}\mathbf{e}_{\ell i}}{T} \tag{C.14}$$

$$\begin{aligned}
\mathbf{b}_{122313} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}'_{\ell i}T^{-1}\mathbf{F}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\frac{1}{NT}\sum_{s=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\mathbf{R}'\mathbf{F}^{0\prime}\mathbf{e}_{sj}\mathbf{e}'_{sj}\mathbf{e}_{\ell i} \\
&+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\mathbf{H}'_{\ell i}T^{-1}\mathbf{F}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\frac{1}{NT}\sum_{s=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\mathbf{e}_{sj}\mathbf{e}'_{sj}\mathbf{e}_{\ell i} \\
&= \mathbf{b}_{1223131}+\mathbf{b}_{1223132}
\end{aligned}$$

$$\begin{aligned}
\mathbf{b}_{1223131} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\frac{\mathbf{H}'_{\ell i}\mathbf{F}}{T}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\frac{1}{\sqrt{T}}\hat{\mathbf{Q}}'\mathbf{R}'\frac{\mathbf{F}^{0\prime}\bar{\Sigma}_{EE}\mathbf{e}_{\ell i}}{\sqrt{T}} \\
&+\frac{1}{T}\frac{1}{N}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\frac{\mathbf{H}'_{\ell i}\mathbf{F}}{T}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^{0}}{T}\right)^{-1}\hat{\mathbf{Q}}'\mathbf{R}'\frac{\mathbf{F}^{0\prime}\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\mathbf{e}_{\ell i}}{\sqrt{T}} \\
&= \mathbf{b}_{12231311}+\mathbf{b}_{12231312}
\end{aligned}$$

$$
\begin{aligned}
\left\|\mathbf{b}_{12231311}\right\| &= \left\|\frac{1}{\sqrt{T}}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_j^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{\ell i}^0\frac{\boldsymbol{\epsilon}_{f\varepsilon,j}'\mathbf{F}}{T}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{Q}}'\mathbf{R}'\frac{\mathbf{F}^{0\prime}\bar{\Sigma}_{EE}\mathbf{e}_{\ell i}}{\sqrt{T}}\right\| \\
&= \left\|\frac{1}{\sqrt{T}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_j^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{g}_{\ell i}^0\frac{\mathbf{e}_{\ell i}'\bar{\Sigma}_{EE}\mathbf{F}^0}{T}\mathbf{R}\hat{\mathbf{Q}}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\frac{\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,j}}{T}\right\| \\
&\leq \frac{1}{\sqrt{T}}\sum_{\ell=1}^{k+1}\frac{1}{N}\sum_{j=1}^{N}\left\|\boldsymbol{\Gamma}_j^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\right\|\left\|\frac{\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,j}}{T}\right\|\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{g}_{\ell i}^0\frac{\mathbf{e}_{\ell i}'\bar{\Sigma}_{EE}\mathbf{F}^0}{T}\right\|\left\|\mathbf{R}\right\|\left\|\hat{\mathbf{Q}}\right\|\left\|\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\right\| \\
&= O_p\left(T^{-1/2}\right).
\end{aligned}
$$

$$
\begin{aligned}
\left\|\mathbf{b}_{12231312}\right\| &\leq \frac{1}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\left\|\frac{\mathbf{H}_{\ell i}'\mathbf{F}}{T}\right\|\left\|\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{R}'\right\|\left\|\frac{\mathbf{F}^{0\prime}\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\mathbf{e}_{\ell i}}{T}\right\| \\
&= O_p\left(T^{-1/2}\right).
\end{aligned}
$$

$$
\begin{aligned}
\left\|\mathbf{b}_{1223132}\right\| &= \left\|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_j^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{\ell i}^0\frac{\boldsymbol{\epsilon}_{f\varepsilon,j}'\mathbf{F}^0}{T}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\frac{1}{NT}\sum_{s=1}^{k+1}\sum_{h=1}^{N}\hat{\mathbf{Q}}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{e}_{sh}\left(\mathbf{e}_{sh}'\mathbf{e}_{\ell i}\right)\right\| \\
&= \left\|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{\ell=1}^{k+1}\frac{1}{NT}\sum_{s=1}^{k+1}\sum_{h=1}^{N}\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_j^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{\ell i}^0\left(\mathbf{e}_{\ell i}'\mathbf{e}_{sh}\right)\frac{\boldsymbol{\epsilon}_{f\varepsilon,j}'\mathbf{F}^0}{T}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{Q}}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{e}_{sh}\right\| \\
&\leq \sqrt{T}\left(\frac{1}{N}\sum_{j=1}^{N}\left\|\boldsymbol{\Gamma}_j^0\right\|\left\|\frac{\boldsymbol{\epsilon}_{f\varepsilon,j}'\mathbf{F}^0}{T}\right\|\right)\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\frac{\boldsymbol{g}_{\ell i}^0\mathbf{e}_{\ell i}'}{\sqrt{T}}\right\|\left(\frac{1}{N}\sum_{s=1}^{k+1}\sum_{h=1}^{N}\left\|\frac{\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{e}_{sh}}{T}\right\|\left\|\frac{\mathbf{e}_{sh}}{\sqrt{T}}\right\|\right) \\
&\quad\times\left\|\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\right\|\left\|\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\right\|\left\|\hat{\mathbf{Q}}'\right\| \\
&= \sqrt{T}O_p\left(\delta_{NT}^{-2}\right).
\end{aligned}
$$

Noting $\hat{\mathbf{Q}}=\left(\boldsymbol{\Upsilon}_N^0\boldsymbol{\Lambda}_{0\hat{F}}\right)^{-1}=\boldsymbol{\Lambda}_{0\hat{F}}^{-1}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}$

$$
\begin{aligned}
\mathbf{b}_{13} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\left(\frac{1}{T}\hat{\mathbf{Q}}'\hat{\mathbf{F}}'\mathbf{e}_{\ell i}\mathbf{e}_{\ell i}'\right)\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\frac{1}{T}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'\Sigma_{EE}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i}.
\end{aligned}
$$

but, by Lemma 7 of Norkute et al (2018),

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\left(\frac{1}{T}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'\Sigma_{EE}\right)\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\left(\frac{1}{T}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'\bar{\Sigma}_{EE}\right)\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i}+o_p(1) \\
&= \mathbf{b}_{13}^*+o_p(1).
\end{aligned}
$$

$$\begin{aligned}
\mathbf{b}^*_{13} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}^0_i\left(\frac{1}{T}\left(\boldsymbol{\Upsilon}^0_N\right)^{-1}\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'\bar{\Sigma}_{EE}\right)\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}^0_i\frac{1}{T}\left(\boldsymbol{\Upsilon}^0_N\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\mathbf{F}^{0\prime}\bar{\Sigma}_{EE}\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}^0_i\frac{1}{T}\left(\boldsymbol{\Upsilon}^0_N\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\mathbf{F}^{0\prime}\bar{\Sigma}_{EE}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}^0_i\frac{1}{T}\left(\boldsymbol{\Upsilon}^0_N\right)^{-1}\left[\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'-\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\mathbf{F}^{0\prime}\right]\bar{\Sigma}_{EE}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}^*_{131}+\mathbf{b}^*_{132}+\mathbf{b}^*_{133}
\end{aligned}$$

$$\begin{aligned}
\left\|\mathbf{b}^*_{132}\right\| &= \left\|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}^0_i\frac{1}{T}\left(\boldsymbol{\Upsilon}^0_N\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\mathbf{F}^{0\prime}\bar{\Sigma}_{EE}\left(-\mathbf{P}_{\hat{\mathbf{F}}}+\mathbf{P}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{f\varepsilon,i}\right\| \\
&\le \sqrt{\frac{N}{T}}\left\|\left(\boldsymbol{\Upsilon}^0_N\right)^{-1}\right\|\left\|\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}^{0\prime}}{\sqrt{T}}\right\|\mu_{\max}\left(\bar{\Sigma}_{EE}\right)\left\|\mathbf{P}_{\hat{\mathbf{F}}}-\mathbf{P}_{\mathbf{F}^0}\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}^0_i\right\|\left\|\frac{\boldsymbol{\epsilon}_{f\varepsilon,i}}{\sqrt{T}}\right\| \\
&= \sqrt{\frac{N}{T}}O_p\left(\delta^{-1}_{NT}\right).
\end{aligned}$$

Noting

$$\begin{aligned}
\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'-\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\mathbf{F}^{0\prime} &= \left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'-\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\mathbf{F}^{0\prime} \\
&= \left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'\mathbf{M}_{\mathbf{F}^0}
\end{aligned}$$

$$\begin{aligned}
\mathbf{b}^*_{133} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}^0_i\frac{1}{T}\left(\boldsymbol{\Upsilon}^0_N\right)^{-1}\left[\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'-\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\mathbf{F}^{0\prime}\right]\bar{\Sigma}_{EE}\mathbf{M}_{\hat{\mathbf{F}}}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}^0_i\frac{1}{T}\left(\boldsymbol{\Upsilon}^0_N\right)^{-1}\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'\mathbf{M}_{\mathbf{F}^0}\bar{\Sigma}_{EE}\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}^0_i\frac{1}{T}\left(\boldsymbol{\Upsilon}^0_N\right)^{-1}\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\hat{\mathbf{F}}'\mathbf{M}_{\mathbf{F}^0}\bar{\Sigma}_{EE}\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}^*_{1321}+\mathbf{b}^*_{1322}
\end{aligned}$$

Noting

$$\frac{1}{\sqrt{T}}\left\|\mathbf{M}_{\mathbf{F}^0}\hat{\mathbf{F}}\right\|=O_p\left(\delta^{-1}_{NT}\right)$$

as $\frac{1}{\sqrt{T}}\left\|\mathbf{M}_{\mathbf{F}^0}\hat{\mathbf{F}}\right\|=\frac{1}{\sqrt{T}}\left\|\left(\mathbf{M}_{\mathbf{F}^0}-\mathbf{M}_{\hat{\mathbf{F}}}\right)\hat{\mathbf{F}}\right\|=\frac{1}{\sqrt{T}}\left\|\left(\mathbf{P}_{\mathbf{F}^0}-\mathbf{P}_{\hat{\mathbf{F}}}\right)\hat{\mathbf{F}}\right\|\le\left\|\mathbf{P}_{\mathbf{F}^0}-\mathbf{P}_{\hat{\mathbf{F}}}\right\|\left\|\frac{\hat{\mathbf{F}}}{\sqrt{T}}\right\|=O_p\left(\delta^{-1}_{NT}\right)$,

$$\frac{1}{\sqrt{T}}\left\|\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\epsilon}_{f\varepsilon,i}\right\|=O_p\left(1\right),$$

$$\begin{aligned}
\left\|\mathbf{b}^*_{1321}\right\| &\le \sqrt{\frac{N}{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}^0_i\right\|\left\|\left(\boldsymbol{\Upsilon}^0_N\right)^{-1}\right\|\left\|\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^0}{T}\right)^{-1}\right\|\frac{1}{\sqrt{T}}\left\|\hat{\mathbf{F}}'\mathbf{M}_{\mathbf{F}^0}\right\|\mu_{\max}\left(\bar{\Sigma}_{EE}\right)\frac{1}{\sqrt{T}}\left\|\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\epsilon}_{f\varepsilon,i}\right\| \\
&= O_p\left(\delta^{-1}_{NT}\right)
\end{aligned}$$

$$\begin{aligned}
\left\|\mathbf{b}_{1322}^{*}\right\| &\leq \sqrt{\frac{N}{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{i}^{0}\right\|\left\|\left(\mathbf{\Upsilon}_{N}^{0}\right)^{-1}\right\|\left\|\left(\frac{\hat{\mathbf{F}}'\mathbf{F}^{0}}{T}\right)^{-1}\right\|\frac{1}{\sqrt{T}}\left\|\hat{\mathbf{F}}'\mathbf{M}_{\mathbf{F}^{0}}\right\|\mu_{\max}\left(\bar{\Sigma}_{EE}\right)\left\|\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^{0}}\right\|\left\|\frac{\boldsymbol{\epsilon}_{f\varepsilon,i}}{\sqrt{T}}\right\| \\
&= O_{p}\left(\delta_{NT}^{-2}\right).
\end{aligned}$$

so $\mathbf{b}_{11}=o_{p}(1)$, $\mathbf{b}_{13}=o_{p}(1)$, and

$$\mathbf{b}_{13}=o_{p}(1). \tag{C.15}$$

This suggest the transformation of regressors:

$$\tilde{\mathbf{X}}_{i}^{*}=\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_{i}^{*}-\mathbf{M}_{\hat{\mathbf{F}}}\hat{\mathbf{E}}_{j}\hat{\boldsymbol{G}}_{j}^{0\prime}\left(\mathbf{\Upsilon}_{N}^{0}\right)^{-1}\hat{\mathbf{\Gamma}}_{i}^{0}.$$

■

Proof of Lemma B.13.

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i}'\mathbf{M}_{\mathbf{F}^{0}}\boldsymbol{\epsilon}_{f\varepsilon,i}+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i}'\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^{0}}\right)\boldsymbol{\epsilon}_{f\varepsilon,i}$$

but

$$\begin{aligned}
\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i}'\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^{0}}\right)\boldsymbol{\epsilon}_{f\varepsilon,i} &= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i}'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)\mathbf{R}'\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i}'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i}'\left[T^{-1}\mathbf{F}^{0}\mathbf{R}\left(\hat{\mathbf{F}}-\mathbf{F}^{0}\mathbf{R}\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i}'\left\{T^{-1}\mathbf{F}\left[\mathbf{R}\mathbf{R}'-\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right]\mathbf{F}'\right\}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= -\mathbf{b}_{21}-\mathbf{b}_{22}-\mathbf{b}_{23}-\mathbf{b}_{24}.
\end{aligned}$$

$$\begin{aligned}
\mathbf{b}_{21} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i}'\left[T^{-1}\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^{0}\right)\mathbf{R}\mathbf{R}'\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i}'\left[T^{-1}\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^{0}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^{0}\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad +\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i}'\left[T^{-1}\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^{0}\right)\left[\mathbf{R}\mathbf{R}'-\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^{0}\right)^{-1}\right]\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{211}+\mathbf{b}_{212}.
\end{aligned}$$

$$\left\|\mathbf{b}_{212}\right\|\leq\sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}T^{-1}\left\|\mathbf{V}_{i}'\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^{0}\right)\right\|\left\|\mathbf{R}\mathbf{R}'-\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^{0}\right)^{-1}\right\|\left\|\frac{\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i}}{T}\right\|=\sqrt{NT}O_{p}\left(\delta_{NT}^{-4}\right).$$

$$
\begin{aligned}
\mathbf{b}_{211} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^0\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i}\\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\mathbf{F}^0\boldsymbol{g}_{\ell j}^0\mathbf{e}_{\ell j}'\hat{\mathbf{F}}\hat{\mathbf{Q}}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i}\\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\mathbf{e}_{\ell j}\boldsymbol{g}_{\ell j}^{0\prime}\mathbf{F}^{0\prime}\hat{\mathbf{F}}\hat{\mathbf{Q}}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i}\\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\mathbf{e}_{\ell j}\mathbf{e}_{\ell j}'\hat{\mathbf{F}}\hat{\mathbf{Q}}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i}\\
&= \mathbf{b}_{2111}+\mathbf{b}_{2112}+\mathbf{b}_{2113}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{b}_{2111} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\left[T^{-1}\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^0\mathbf{e}_{\ell j}'\right)\mathbf{F}^0\mathbf{R}\hat{\mathbf{Q}}\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i}\\
&\quad+\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\left[T^{-1}\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^0\mathbf{e}_{\ell j}'\right)\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\hat{\mathbf{Q}}\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i}\\
&= \mathbf{b}_{21111}+\mathbf{b}_{21112}
\end{aligned}
$$

$$
\begin{aligned}
\left\|\mathbf{b}_{21111}\right\| &= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\left[T^{-1}\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^0\mathbf{e}_{\ell j}'\right)\mathbf{F}^0\mathbf{R}\hat{\mathbf{Q}}\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\left(\mathbf{F}^0\boldsymbol{\lambda}_i+\boldsymbol{\varepsilon}_i\right)\right\|\\
&\le \frac{1}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^0\frac{\mathbf{e}_{\ell j}'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\mathbf{R}\right\|\left\|\hat{\mathbf{Q}}\right\|\left\|\boldsymbol{\lambda}_i\right\|\\
&\quad+\frac{1}{T}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^0\frac{\mathbf{e}_{\ell j}'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\mathbf{R}\right\|\left\|\hat{\mathbf{Q}}\right\|\left\|\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\right\|\left\|\frac{\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|\\
&= O_p\left(T^{-1/2}\right).
\end{aligned}
$$

as $\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^0\frac{\mathbf{e}_{\ell j}'\mathbf{F}^0}{\sqrt{T}}\right\|=O_p\left(1\right)$. Similarly,

$$
\begin{aligned}
\left\|\mathbf{b}_{21112}\right\| &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\left[T^{-1}\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^0\mathbf{e}_{\ell j}'\right)\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\hat{\mathbf{Q}}\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i}\\
&\le \frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^0\frac{\mathbf{e}_{\ell j}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)}{T}\right\|\left\|\hat{\mathbf{Q}}\right\|\left\|\boldsymbol{\lambda}_i\right\|\\
&\quad+\frac{1}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^0\frac{\mathbf{e}_{\ell j}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)}{T}\right\|\left\|\hat{\mathbf{Q}}\right\|\left\|\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\right\|\left\|\frac{\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|\\
&= O_p\left(\delta_{NT}^{-2}\right).
\end{aligned}
$$

Thus, $\mathbf{b}_{2111} = O_p\left(\delta_{NT}^{-2}\right) + O_p\left(T^{-1/2}\right)$. Noting $\hat{\mathbf{Q}} = \left(\boldsymbol{\Upsilon}_N^0 \boldsymbol{\Lambda}_{0\hat{F}}\right)^{-1} = \boldsymbol{\Lambda}_{0\hat{F}}^{-1}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}$

$$\begin{aligned}
\mathbf{b}_{2112} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(\frac{1}{N}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\mathbf{e}_{\ell j}\boldsymbol{g}_{\ell j}^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{E}_j}{T}\mathbf{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\frac{\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i}}{T} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{E}_j}{T}\mathbf{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i^0 \\
&\quad+\sqrt{\frac{1}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{E}_j}{T}\mathbf{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\frac{\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{E}_j}{T}\mathbf{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{\lambda}_i^0 + o_p(1). \qquad \text{(C.16)}
\end{aligned}$$

$$\begin{aligned}
\mathbf{b}_{2113} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\mathbf{e}_{\ell j}\mathbf{e}_{\ell j}'\hat{\mathbf{F}}\hat{\mathbf{Q}}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\mathbf{e}_{\ell j}\mathbf{e}_{\ell j}'\mathbf{F}^0\mathbf{R}\hat{\mathbf{Q}}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\mathbf{e}_{\ell j}\mathbf{e}_{\ell j}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\hat{\mathbf{Q}}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{21131} + \mathbf{b}_{21132}
\end{aligned}$$

Recalling $\Sigma_{EE} = N^{-1}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\mathbf{e}_{\ell i}\mathbf{e}_{\ell i}' = N^{-1}\sum_{i=1}^{N}\mathbf{E}_i'\mathbf{E}_i$ and $\bar{\Sigma}_{EE} = N^{-1}\sum_{i=1}^{N}E\left(\mathbf{E}_i'\mathbf{E}_i\right)$,

$$\begin{aligned}
\mathbf{b}_{21132} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(T^{-1}\bar{\Sigma}_{EE}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\hat{\mathbf{Q}}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(T^{-1}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\hat{\mathbf{Q}}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{211321} + \mathbf{b}_{211322}
\end{aligned}$$

$$\begin{aligned}
\left\|\mathbf{b}_{211321}\right\| &\leq \sqrt{\frac{N}{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'}{\sqrt{T}}\right\|\mu_{\max}\left(\bar{\Sigma}_{EE}\right)\left\|\frac{\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)}{\sqrt{T}}\right\|\left\|\hat{\mathbf{Q}}\right\|\left\|\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i}}{T}\right\| \\
&= O_p\left(\delta_{NT}^{-1}\right)
\end{aligned}$$

$$\begin{aligned}
\left\|\mathbf{b}_{211322}\right\| &\leq \frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'}{\sqrt{T}}\right\|\left\|\frac{\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)}{\sqrt{T}}\right\|\left\|\frac{\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)}{\sqrt{T}}\right\|\left\|\hat{\mathbf{Q}}\right\|\left\|\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i}}{T}\right\| \\
&= O_p\left(\delta_{NT}^{-1}\right)
\end{aligned}$$

$$\begin{aligned}
\mathbf{b}_{21131} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(T^{-1}\bar{\Sigma}_{EE}\mathbf{F}^0\mathbf{R}\hat{\mathbf{Q}}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(T^{-1}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\mathbf{F}^0\mathbf{R}\hat{\mathbf{Q}}\right)\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{211311}+\mathbf{b}_{211312}
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{b}_{211311}\| &\le \left\|\frac{\sqrt{N}}{T}\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\bar{\Sigma}_{EE}\mathbf{F}^0}{\sqrt{T}}\mathbf{R}\hat{\mathbf{Q}}\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}T^{-1}\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i}\right\| \\
&\le \frac{\sqrt{N}}{T}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\bar{\Sigma}_{EE}\mathbf{F}^0}{\sqrt{T}}\right\|\|\mathbf{R}\|\left\|\hat{\mathbf{Q}}\right\|\left\|\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\right\|\left\|T^{-1}\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i}\right\| \\
&= O_p\left(T^{-1/2}\right)
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{b}_{211312}\| &\le \frac{1}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'}{\sqrt{T}}\right\|\left\|\frac{\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\mathbf{R}\hat{\mathbf{Q}}\right\|\left\|\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\right\|\left\|T^{-1}\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i}\right\| \\
&= O_p\left(T^{-1/2}\right)
\end{aligned}$$

as $\left\|\frac{\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\mathbf{F}^0}{\sqrt{T}}\right\| = O_p(1)$. $\mathbf{b}_{21132}$ is $o_p\left(T^{-1/2}\right)$.

$$\begin{aligned}
\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\boldsymbol{\epsilon}_{f\varepsilon,i} &= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\mathbf{R}'\mathbf{F}^{0\prime}\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad-\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad-\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\mathbf{F}^0\mathbf{R}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad-\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left\{T^{-1}\mathbf{F}\left[\mathbf{R}\mathbf{R}'-\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right]\mathbf{F}'\right\}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= -\mathbf{b}_{21}-\mathbf{b}_{22}-\mathbf{b}_{23}-\mathbf{b}_{24}.
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{b}_{22}\| &\le \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\left\|\mathbf{V}_i'\left[T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\right]\left(\mathbf{F}^0\boldsymbol{\lambda}_i^0+\boldsymbol{\varepsilon}_i\right)\right\| \\
&\le \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|T^{-1}\mathbf{V}_i'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{F}^0\right\|\left\|\boldsymbol{\lambda}_i^0\right\| \\
&\quad+\sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|T^{-1}\mathbf{V}_i'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)\right\|\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\boldsymbol{\varepsilon}_i\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right)
\end{aligned}$$

$$
\begin{aligned}
\|\mathbf{b}_{24}\| &\leq \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\left\|\mathbf{V}_i'\left\{T^{-1}\mathbf{F}\left[\mathbf{RR}'-\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right]\mathbf{F}'\right\}\left(\mathbf{F}^0\boldsymbol{\lambda}_i^0+\boldsymbol{\varepsilon}_i\right)\right\| \\
&\leq \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}}{\sqrt{T}}\right\|\left\|\mathbf{RR}'-\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right\|\left\|\frac{\mathbf{F}'\mathbf{F}^0}{T}\right\|\left\|\boldsymbol{\lambda}_i^0\right\| \\
&\quad+\sqrt{\frac{N}{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}}{\sqrt{T}}\right\|\left\|\mathbf{RR}'-\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right\|\left\|\frac{\mathbf{F}'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= \sqrt{N}O_p\left(\delta_{NT}^{-2}\right).
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{b}_{23} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\mathbf{F}^0\mathbf{RR}'\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^0\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^0\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\mathbf{F}^0\left(\mathbf{RR}'-\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right)\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^0\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{231}+\mathbf{b}_{232}
\end{aligned}
$$

$\mathbf{b}_{232}=o_p\left(1\right)$.

$$
\begin{aligned}
\mathbf{b}_{231} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left[T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\left(\hat{\mathbf{F}}\mathbf{R}^{-1}-\mathbf{F}^0\right)'\right]\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\hat{\mathbf{F}}'\mathbf{e}_{\ell\mathbf{j}}\boldsymbol{g}_{\ell\mathbf{j}}^{\mathbf{0}\prime}\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\hat{\mathbf{F}}'\mathbf{F}^{\mathbf{0}}\boldsymbol{g}_{\ell\mathbf{j}}^{\mathbf{0}}\mathbf{e}_{\ell j}'\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\hat{\mathbf{F}}'\mathbf{e}_{\ell j}\mathbf{e}_{\ell j}'\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{2311}+\mathbf{b}_{2312}+\mathbf{b}_{2313}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{b}_{2311} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\mathbf{R}'\mathbf{F}^{0\prime}\mathbf{e}_{\ell\mathbf{j}}\boldsymbol{g}_{\ell\mathbf{j}}^{\mathbf{0}\prime}\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{e}_{\ell\mathbf{j}}\boldsymbol{g}_{\ell\mathbf{j}}^{\mathbf{0}\prime}\mathbf{F}^{0\prime}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{23111}+\mathbf{b}_{23112}
\end{aligned}
$$

$$
\begin{aligned}
\|\mathbf{b}_{23111}\| &= \left\|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\mathbf{R}'\mathbf{F}^{0\prime}\mathbf{e}_{\ell\mathbf{j}}\boldsymbol{g}_{\ell\mathbf{j}}^{\mathbf{0}\prime}\mathbf{F}^{0\prime}\left(\mathbf{F}^0\boldsymbol{\lambda}_i+\boldsymbol{\varepsilon}_i\right)\right\| \\
&\leq \frac{\sqrt{N}}{\sqrt{T}}\frac{1}{T}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right\|\left\|\frac{1}{N}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\mathbf{R}'\frac{\mathbf{F}^{0\prime}\mathbf{e}_{\ell\mathbf{j}}}{\sqrt{T}}\boldsymbol{g}_{\ell\mathbf{j}}^{\mathbf{0}\prime}\right\|\left\|\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right\| \\
&\quad+\frac{\sqrt{N}}{T}\frac{1}{T}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right\|\left\|\frac{1}{N}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\mathbf{R}'\frac{\mathbf{F}^{0\prime}\mathbf{e}_{\ell\mathbf{j}}}{\sqrt{T}}\boldsymbol{g}_{\ell\mathbf{j}}^{\mathbf{0}\prime}\right\|\left\|\frac{\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= O_p\left(T^{-1}\right).
\end{aligned}
$$

as $\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^{0}\frac{\mathbf{e}_{\ell j}'\mathbf{F}^{0}}{\sqrt{T}}\right\| = O_p(1)$. Similarly,

$$
\begin{aligned}
\|\mathbf{b}_{23112}\| &\leq \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\left\|\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{e}_{\ell\mathbf{j}}\boldsymbol{g}_{\ell\mathbf{j}}^{\mathbf{0}\prime}\mathbf{F}^{0\prime}\left(\mathbf{F}^0\boldsymbol{\lambda}_i+\boldsymbol{\varepsilon}_i\right)\right\| \\
&\leq \frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\right\|\left\|\hat{\mathbf{Q}}\right\|\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^{0}\frac{\mathbf{e}_{\ell j}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)}{T}\right\|\left\|T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right\|\|\boldsymbol{\lambda}_i\| \\
&\quad+\frac{1}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\left(T^{-1}\mathbf{F}^{0\prime}\mathbf{F}^0\right)^{-1}\right\|\left\|\hat{\mathbf{Q}}\right\|\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\boldsymbol{g}_{\ell j}^{0}\frac{\mathbf{e}_{\ell j}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)}{T}\right\|\left\|T^{-1/2}\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i\right\| \\
&= O_p\left(\delta_{NT}^{-2}\right).
\end{aligned}
$$

Thus, $\mathbf{b}_{2311} = O_p\left(\delta_{NT}^{-2}\right) + O_p\left(T^{-1}\right)$. Noting $\hat{\mathbf{Q}} = \left(\boldsymbol{\Upsilon}_N^0\boldsymbol{\Lambda}_{0\hat{F}}\right)^{-1} = \boldsymbol{\Lambda}_{0\hat{F}}^{-1}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}$

$$
\begin{aligned}
\mathbf{b}_{2312} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{\ell\mathbf{j}}^{\mathbf{0}}\mathbf{e}_{\ell j}'\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{j=1}^{N}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\mathbf{G}_j^0\mathbf{E}_j'\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{j=1}^{N}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\mathbf{G}_j^0\mathbf{E}_j'\mathbf{F}^0\boldsymbol{\lambda}_i \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{j=1}^{N}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\mathbf{G}_j^0\mathbf{E}_j'\boldsymbol{\varepsilon}_i \\
&= \frac{1}{\sqrt{N}T^{3/2}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\mathbf{G}_j^0\frac{\mathbf{E}_j'\mathbf{F}^0}{\sqrt{T}}\boldsymbol{\lambda}_i \\
&\quad+\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\mathbf{G}_j^0\frac{\mathbf{E}_j'\boldsymbol{\varepsilon}_i}{T} \\
&= O_p\left(N^{-1/2}T^{-1}\right)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{b}_{2313} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\hat{\mathbf{F}}'\mathbf{e}_{\ell j}\mathbf{e}_{\ell j}'\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\mathbf{R}'\mathbf{F}^{0\prime}\mathbf{e}_{\ell j}\mathbf{e}_{\ell j}'\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\frac{1}{NT}\sum_{\ell=1}^{k+1}\sum_{j=1}^{N}\hat{\mathbf{Q}}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{e}_{\ell j}\mathbf{e}_{\ell j}'\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{23131}+\mathbf{b}_{23132}
\end{aligned}
$$

Recalling $\Sigma_{EE} = N^{-1}\sum_{\ell=1}^{k+1}\sum_{i=1}^{N}\mathbf{e}_{\ell i}\mathbf{e}_{\ell i}' = N^{-1}\sum_{i=1}^{N}\mathbf{E}_i'\mathbf{E}_i$ and $\bar{\Sigma}_{EE} = N^{-1}\sum_{i=1}^{N}E\left(\mathbf{E}_i'\mathbf{E}_i\right)$,

$$
\begin{aligned}
\mathbf{b}_{23132} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\hat{\mathbf{Q}}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'T^{-1}\bar{\Sigma}_{EE}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\hat{\mathbf{Q}}'\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'T^{-1}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\boldsymbol{\epsilon}_{f\varepsilon,i}
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{b}_{23131} &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i' T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\hat{\mathbf{Q}}'\mathbf{R}'\mathbf{F}^{0\prime}T^{-1}\bar{\Sigma}_{EE}\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&\quad+\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i' T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\hat{\mathbf{Q}}'\mathbf{R}'\mathbf{F}^{0\prime}T^{-1}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\boldsymbol{\epsilon}_{f\varepsilon,i} \\
&= \mathbf{b}_{231311}+\mathbf{b}_{231312}
\end{aligned}
$$

$$
\begin{aligned}
\left\|\mathbf{b}_{231311}\right\| &\le \left\|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i' T^{-1}\mathbf{F}^0\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\hat{\mathbf{Q}}'\mathbf{R}'\mathbf{F}^{0\prime}T^{-1}\bar{\Sigma}_{EE}\left(\mathbf{F}^0\boldsymbol{\lambda}_i+\boldsymbol{\varepsilon}_i\right)\right\| \\
&\le \frac{1}{\sqrt{T}}\sqrt{\frac{N}{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{R}'\right\|\left\|\frac{\mathbf{F}^{0\prime}\bar{\Sigma}_{EE}\mathbf{F}^0}{T}\right\|\left\|\boldsymbol{\lambda}_i\right\| \\
&\quad+\frac{1}{\sqrt{T}}\sqrt{\frac{N}{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{R}'\right\|\left\|\frac{\mathbf{F}^{0\prime}\bar{\Sigma}_{EE}\boldsymbol{\varepsilon}_i}{T}\right\| \\
&= O_p\left(T^{-1/2}\right)
\end{aligned}
$$

$$
\begin{aligned}
\left\|\mathbf{b}_{231312}\right\| &\le \frac{1}{T}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{R}'\right\|\left\|\frac{\mathbf{F}^{0\prime}}{\sqrt{T}}\right\|\left\|\frac{\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\boldsymbol{\lambda}_i\right\| \\
&\quad+\frac{1}{T}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}^0}{\sqrt{T}}\right\|\left\|\left(T^{-1}\mathbf{F}'\mathbf{F}\right)^{-1}\right\|\left\|\hat{\mathbf{Q}}'\right\|\left\|\mathbf{R}'\right\|\left\|\frac{\sqrt{N}\mathbf{F}^{0\prime}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)}{\sqrt{T}}\right\|\left\|\frac{\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= O_p\left(T^{-1}\right)
\end{aligned}
$$

as $\left\|\frac{\sqrt{N}\left(\Sigma_{EE}-\bar{\Sigma}_{EE}\right)\mathbf{F}^0}{\sqrt{T}}\right\|=O_p(1)$. $\mathbf{b}_{21132}$ is $O_p\left(T^{-1/2}\right)$. ■

Proof of Lemma B.14.

$$
\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_i}{T}=\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_i}{T}+o_p(1)
$$

$$
\begin{aligned}
\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_i}{T} &= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\boldsymbol{\varepsilon}_i}{T} \\
&\quad-\frac{1}{T}\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\mathbf{F}^0}{\sqrt{T}}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\frac{\mathbf{F}^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\boldsymbol{\varepsilon}_i}{T}+O_p\left(T^{-1}\right).
\end{aligned}
$$

since $E\left\|\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\right\|^2\le\Delta$

$$
\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\left[\frac{\mathbf{E}_j'\boldsymbol{\varepsilon}_i}{T}-E\left(\frac{\mathbf{E}_j'\boldsymbol{\varepsilon}_i}{T}\right)\right]=o_p(1).
$$

Recalling that $\mathbf{E}_i=\left(\mathbf{V}_i\boldsymbol{\beta}_i+\boldsymbol{\varepsilon}_i,\mathbf{V}_i\right)$, $E\left(\frac{\mathbf{E}_j'\boldsymbol{\varepsilon}_i}{T}\right)=0$ for $j\neq i$ and

$$
E\left(\frac{\mathbf{E}_i'\boldsymbol{\varepsilon}_i}{T}\right)=T^{-1}E\begin{pmatrix}\left(\mathbf{V}_i\boldsymbol{\beta}_i+\boldsymbol{\varepsilon}_i\right)'\boldsymbol{\varepsilon}_i\\ \mathbf{V}_i'\boldsymbol{\varepsilon}_i\end{pmatrix}=E\begin{pmatrix}\sigma_{\varepsilon i}^2\\ \mathbf{0}\end{pmatrix},
$$

giving

$$\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_j^0\frac{\mathbf{E}_j'\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\varepsilon}_i}{T}=\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{g}_{1i}^0\sigma_{\varepsilon i}^2+o_p\left(1\right).$$

■

Proof of Lemma B.15. A similar derivation in the proof of Lemma B.16 gives

$$\begin{aligned}
&\sqrt{\frac{T}{N}}\frac{1}{N}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_n^0\frac{\mathbf{E}_n'\mathbf{E}_j}{T}\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\frac{\mathbf{F}^{0\prime}\mathbf{u}_i}{T}\\
&-\sqrt{\frac{T}{N}}\frac{1}{N}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_n^0E\left(\frac{\mathbf{E}_n'\mathbf{E}_j}{T}\right)\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\frac{\mathbf{F}^{0\prime}\mathbf{u}_i}{T}\\
=&\sqrt{\frac{T}{N}}\frac{1}{N}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\boldsymbol{G}_n^0\left[\frac{\mathbf{E}_n'\mathbf{E}_j-E\left(\mathbf{E}_n'\mathbf{E}_j\right)}{T}\right]\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\Gamma_i^0\boldsymbol{\eta}_i+\boldsymbol{\lambda}_i^0\right)+o_p\left(1\right)\\
=&\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\left[\left(\Gamma_i^0\boldsymbol{\eta}_i+\boldsymbol{\lambda}_i^0\right)'\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\otimes\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\right]\\
&\times\left\{\frac{1}{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\left(\boldsymbol{G}_j^0\otimes\boldsymbol{G}_n^0\right)vec\left[\frac{\mathbf{E}_n'\mathbf{E}_j-E\left(\mathbf{E}_n'\mathbf{E}_j\right)}{T}\right]\right\}
\end{aligned}$$

but as $\frac{1}{N}\sum_{i=1}^{N}\left[\left(\Gamma_i^0\boldsymbol{\eta}_i+\boldsymbol{\lambda}_i^0\right)'\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\otimes\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\right]=O_p\left(1\right)$ and $E\left\|\boldsymbol{G}_j^0\otimes\boldsymbol{G}_n^0\right\|^2\leq\Delta$, by Assumption B1-B3, inside of curly brackets is $O_p\left(T^{-1/2}\right)$. ■

Proof of Lemma B.16.

$$\begin{aligned}
&\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{E}_j}{T}\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\frac{\mathbf{F}^{0\prime}\mathbf{u}_i}{T}\\
&-\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}E\left(\frac{\mathbf{V}_i'\mathbf{E}_j}{T}\right)\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}\mathbf{F}^0}{T}\right)^{-1}\frac{\mathbf{F}^{0\prime}\mathbf{u}_i}{T}\\
=&\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\frac{\mathbf{V}_i'\mathbf{E}_j-E\left(\mathbf{V}_i'\mathbf{E}_j\right)}{T}\right]\boldsymbol{G}_j^{0\prime}\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\Gamma_i^0\boldsymbol{\eta}_i+\boldsymbol{\lambda}_i^0\right)+o_p\left(1\right)\\
=&\sqrt{\frac{T}{N}}\left\{\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\frac{\mathbf{V}_i'\mathbf{E}_j-E\left(\mathbf{V}_i'\mathbf{E}_j\right)}{T}\right]\left[\left(\Gamma_i^0\boldsymbol{\eta}_i+\boldsymbol{\lambda}_i^0\right)'\otimes\boldsymbol{G}_j^{0\prime}\right]\right\}vec\left[\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\right]+o_p\left(1\right)
\end{aligned}$$

but as $E\left\|\left(\Gamma_i^0\boldsymbol{\eta}_i+\boldsymbol{\lambda}_i^0\right)'\otimes\boldsymbol{G}_j^{0\prime}\right\|^2\leq\Delta$, and by Assumption B1-B3, inside of curly brackets is $O_p\left(T^{-1/2}\right)$, which gives the required result. ■

Proof of Lemma B.17. Recall that

$$\begin{aligned}
\hat{\boldsymbol{\xi}}_{NT}=&-\frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\Gamma}}_i\hat{\boldsymbol{\Upsilon}}_N^{-1}\hat{\boldsymbol{g}}_{1i}\hat{\sigma}_{\varepsilon i}^2+\frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\Gamma}}_i\hat{\boldsymbol{\Upsilon}}_N^{-1}\left(\frac{1}{N}\sum_{j=1}^{N}\hat{\boldsymbol{G}}_j\hat{\boldsymbol{\Omega}}_{EE,j}\hat{\boldsymbol{G}}_j'\right)\hat{\boldsymbol{\Upsilon}}_N^{-1}\hat{\boldsymbol{\lambda}}_i-\frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\Omega}}_{VE,i}\hat{\boldsymbol{G}}_i'\hat{\boldsymbol{\Upsilon}}_N^{-1}\hat{\boldsymbol{\lambda}}_i\\
=&-\mathbf{h}_1+\mathbf{h}_2-\mathbf{h}_3, \qquad (C.17)
\end{aligned}$$

$$\begin{aligned}
\hat{\boldsymbol{\Gamma}}_i'&=\frac{\hat{\mathbf{F}}'\mathbf{X}_i^*}{T},\ \hat{\boldsymbol{\Upsilon}}_N=N^{-1}\sum_{i=1}^{N}\hat{\boldsymbol{G}}_i\hat{\boldsymbol{G}}_i',\ \hat{\boldsymbol{G}}_i=\frac{\hat{\mathbf{F}}'\mathbf{Z}_i^*}{T},\ \hat{\boldsymbol{g}}_{1i}=\frac{\hat{\mathbf{F}}'\mathbf{y}_i^*}{T}\\
\hat{\sigma}_{\varepsilon i}^2&=\frac{\hat{\mathbf{u}}_i'\mathbf{M}_{\hat{\mathbf{F}}}\hat{\mathbf{u}}_i}{T},\ ,\ \hat{\boldsymbol{\Omega}}_{EE,i}=\frac{\hat{\mathbf{E}}_i'\hat{\mathbf{E}}_i}{T},\ \hat{\mathbf{E}}_i=\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{Z}_i^*\\
\hat{\boldsymbol{\lambda}}_i&=\frac{\hat{\mathbf{F}}'\hat{\mathbf{u}}_i}{T},\ \hat{\boldsymbol{\Omega}}_{VE,i}=\frac{\hat{\mathbf{V}}_i'\hat{\mathbf{E}}_i}{T},\ \hat{\mathbf{V}}_i=\mathbf{M}_{\hat{\mathbf{F}}}\mathbf{X}_i^*.
\end{aligned}$$

$$
\begin{aligned}
\mathbf{h}_1 &= \frac{1}{N}\sum_{i=1}^{N}\mathbf{R}_\Gamma^{-1}\boldsymbol{\Gamma}_i^0\hat{\boldsymbol{\Upsilon}}_N^{-1}\boldsymbol{g}_{1i}^0\sigma_{\varepsilon i}^2+\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\boldsymbol{\Gamma}}_i-\mathbf{R}_\Gamma^{-1}\boldsymbol{\Gamma}_i^0\right)\hat{\boldsymbol{\Upsilon}}_N^{-1}\hat{\boldsymbol{g}}_{1i}\hat{\sigma}_{\varepsilon i}^2\\
&+\frac{1}{N}\sum_{i=1}^{N}\mathbf{R}_\Gamma^{-1}\boldsymbol{\Gamma}_i^0\hat{\boldsymbol{\Upsilon}}_N^{-1}\hat{\boldsymbol{g}}_{1i}\left(\hat{\sigma}_{\varepsilon i}^2-\sigma_{\varepsilon i}^2\right)+\frac{1}{N}\sum_{i=1}^{N}\mathbf{R}_\Gamma^{-1}\boldsymbol{\Gamma}_i^0\hat{\boldsymbol{\Upsilon}}_N^{-1}\left(\hat{\boldsymbol{g}}_{1i}-\mathbf{R}_{\mathbf{g}_1}^{-1}\boldsymbol{g}_{1i}^0\right)\sigma_{\varepsilon i}^2\\
&= \mathbf{a}_0+\mathbf{a}_1+\mathbf{a}_2+\mathbf{a}_3.
\end{aligned}
$$

$$
\begin{aligned}
\hat{\sigma}_{\varepsilon i}^2 &= \frac{\left[\mathbf{u}_i-\mathbf{X}_i^*\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\right]'\mathbf{M}_{\mathbf{F}^0}\left[\mathbf{u}_i-\mathbf{X}_i^*\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\right]}{T}\\
&+\frac{\left[\mathbf{u}_i-\mathbf{X}_i^*\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\right]'\left(\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right)\left[\mathbf{u}_i-\mathbf{X}_i^*\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\right]}{T}
\end{aligned}
$$

but the norm of the second term is bounded by $\left\|\mathbf{M}_{\hat{\mathbf{F}}}-\mathbf{M}_{\mathbf{F}^0}\right\|\left(\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\|^4\left\|\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right\|^4+\left\|\frac{\mathbf{X}_i^*}{\sqrt{T}}\right\|^4\left\|\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right\|^4\right)=o_p(1)$ uniformly over i as $E\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\|^4\le\Delta$ and $E\left\|\frac{\mathbf{X}_i^*}{\sqrt{T}}\right\|^4\le\Delta$. The first term is

$$
\begin{aligned}
&\frac{\mathbf{u}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{u}_i}{T}+\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)'\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i}{T}\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\\
&+\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)'\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{u}_i}{T}+\frac{\mathbf{u}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i}{T}\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\\
=\;&\mathbf{b}_0+\mathbf{b}_1+\mathbf{b}_2+\mathbf{b}_3
\end{aligned}
$$

$\|\mathbf{b}_2\|$ and $\|\mathbf{b}_3\|$ uniformly $O_p\left((NT)^{-1/2}\right)$ as $E\left\|\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{u}_i}{T}\right\|^2\le\Delta$ and $\|\mathbf{b}_1\|$ is uniformly $O_p\left((NT)^{-1}\right)$ as $E\left\|\frac{\mathbf{V}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{V}_i}{T}\right\|^2\le\Delta$. Finally

$$
\frac{\mathbf{u}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{u}_i}{T}-E\left(\frac{\mathbf{u}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{u}_i}{T}\right)=O_p\left(T^{-1/2}\right)
$$

uniformly over i. Thus, $\hat{\sigma}_{\varepsilon i}^2=O_p(1)$ uniformly over i. In sum,

$$
\|\mathbf{a}_1\|\le\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\Gamma}}_i-\mathbf{R}_\Gamma^{-1}\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\boldsymbol{\Upsilon}}_N^{-1}\right\|\sqrt{r}\left\|\frac{\mathbf{y}_i^*}{\sqrt{T}}\right\|\hat{\sigma}_{\varepsilon i}^2=\frac{1}{\sqrt{N}}O_p\left(\delta_{NT}^{-1}\right).
$$

Next, using the above result,

$$
\|\mathbf{a}_2\|\le\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{R}_\Gamma^{-1}\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\boldsymbol{\Upsilon}}_N^{-1}\right\|\sqrt{r}\left\|\frac{\mathbf{y}_i^*}{\sqrt{T}}\right\|\left|\hat{\sigma}_{\varepsilon i}^2-\sigma_{\varepsilon i}^2\right|=O_p\left(T^{-1/2}\right).
$$

and

$$
\|\mathbf{a}_3\|\le\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{R}_\Gamma^{-1}\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\boldsymbol{\Upsilon}}_N^{-1}\right\|\left\|\hat{\boldsymbol{g}}_{1i}-\mathbf{R}_{\mathbf{g}_1}^{-1}\boldsymbol{g}_{1i}^0\right\|\sigma_{\varepsilon i}^2=\frac{1}{\sqrt{N}}O_p\left(\delta_{NT}^{-1}\right).
$$

Finally, since $\left\|\hat{\boldsymbol{\Upsilon}}_N-\boldsymbol{\Upsilon}_N^0\right\|=o_p(1)$ by Lemma B.4, we have

$$
\mathbf{h}_1=\frac{1}{N}\sum_{i=1}^{N}\mathbf{R}_\Gamma^{-1}\boldsymbol{\Gamma}_i^0\boldsymbol{\Upsilon}_N^{0-1}\boldsymbol{g}_{1i}^0E\left(\frac{\mathbf{u}_i'\mathbf{M}_{\mathbf{F}^0}\mathbf{u}_i}{T}\right)+o_p(1) \tag{C.18}
$$

Next, denoting $\hat{\boldsymbol{\Upsilon}}_N^*=\frac{1}{N}\sum_{j=1}^{N}\hat{\boldsymbol{G}}_j\hat{\boldsymbol{\Omega}}_{EE,j}\hat{\boldsymbol{G}}_j'$

$$
\begin{aligned}
\mathbf{h}_2 &= \frac{1}{N}\sum_{i=1}^{N}\mathbf{R}_\Gamma^{-1}\boldsymbol{\Gamma}_i^0\hat{\boldsymbol{\Upsilon}}_N^{-1}\hat{\boldsymbol{\Upsilon}}_N^*\hat{\boldsymbol{\Upsilon}}_N^{-1}\mathbf{R}_\lambda^{-1}\boldsymbol{\lambda}_i^0\\
&+\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\boldsymbol{\Gamma}}_i-\mathbf{R}_\Gamma^{-1}\boldsymbol{\Gamma}_i^0\right)\hat{\boldsymbol{\Upsilon}}_N^{-1}\hat{\boldsymbol{\Upsilon}}_N^*\hat{\boldsymbol{\Upsilon}}_N^{-1}\hat{\boldsymbol{\lambda}}_i\\
&= \mathbf{c}_0+\mathbf{c}_1.
\end{aligned}
$$

$$\|\mathbf{c}_1\| \leq \frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\Gamma}}_i - \mathbf{R}_{\Gamma}^{-1}\boldsymbol{\Gamma}_i^0\right\|\left\|\hat{\boldsymbol{\Upsilon}}_N^{-1}\hat{\boldsymbol{\Upsilon}}_N^{*}\hat{\boldsymbol{\Upsilon}}_N^{-1}\right\|\left\|\hat{\boldsymbol{\lambda}}_i\right\| = \frac{1}{\sqrt{N}}O_p\left(\delta_{NT}^{-1}\right)$$

as

$$\begin{aligned}
\left\|\hat{\boldsymbol{\lambda}}_i\right\| &= \left\|\frac{\hat{\mathbf{F}}'\hat{\mathbf{u}}_i}{T}\right\| = \left\|T^{-1}\mathbf{R}'\mathbf{F}^{0\prime}\left[\mathbf{u}_i - \mathbf{X}_i^{*}\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\right] + T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\left[\mathbf{u}_i - \mathbf{X}_i^{*}\left(\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right)\right]\right\| \\
&\leq \|\mathbf{R}'\|\left\|\frac{\mathbf{F}^{0\prime}\mathbf{u}_i}{T}\right\| + \|\mathbf{R}'\|\left\|\frac{\mathbf{F}^{0\prime}\mathbf{X}_i^{*}}{T}\right\|\left\|\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right\| + \left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{u}_i\right\| + \left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{X}_i^{*}\right\|\left\|\hat{\boldsymbol{\beta}}_{PC}-\boldsymbol{\beta}\right\| \\
&= O_p(1)
\end{aligned}$$

uniformly over i, as $E\left\|\frac{\mathbf{F}^{0\prime}\mathbf{u}_i}{T}\right\|^2 \leq \Delta$, $E\left\|\frac{\mathbf{F}^{0\prime}\mathbf{X}_i^{*}}{T}\right\|^2 \leq \Delta$ and $\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{u}_i\right\| = O_p\left(\delta_{NT}^{-2}\right)$ and $\left\|T^{-1}\left(\hat{\mathbf{F}}-\mathbf{F}^0\mathbf{R}\right)'\mathbf{X}_i^{*}\right\| = O_p\left(\delta_{NT}^{-2}\right)$ by Lemmas, and a similar discussion leads to

$$\|\mathbf{c}_2\| = \frac{1}{\sqrt{N}}O_p\left(\delta_{NT}^{-1}\right).$$

Note that the above implies that

$$\hat{\boldsymbol{\lambda}}_i = \frac{\hat{\mathbf{F}}'\hat{\mathbf{u}}_i}{T} = \mathbf{R}\frac{\mathbf{F}^{0\prime}\mathbf{u}_i}{T} + o_p(1).$$

Finally, $\hat{\boldsymbol{\Upsilon}}_N - \boldsymbol{\Upsilon}_N^0 = o_p(1)$ and noting that $\frac{1}{N}\sum_{j=1}^{N} vec\left(\hat{\mathbf{G}}_j\hat{\boldsymbol{\Omega}}_{EE,j}\hat{\mathbf{G}}_j'\right) = \frac{1}{N}\sum_{j=1}^{N}\left(\hat{\mathbf{G}}_j\otimes\hat{\mathbf{G}}_j\right) vec\left(\hat{\boldsymbol{\Omega}}_{EE,j}\right)$,

$$\begin{aligned}
\frac{1}{N}\sum_{j=1}^{N}\left(\hat{\mathbf{G}}_j\otimes\hat{\mathbf{G}}_j\right) vec\left(\hat{\boldsymbol{\Omega}}_{EE,j}\right) &= \frac{1}{N}\sum_{j=1}^{N}\left(\mathbf{G}_j^0\otimes\mathbf{G}_j^0\right) vec\left(\bar{\boldsymbol{\Omega}}_{EE,j}\right) \\
&\quad + \frac{1}{N}\sum_{j=1}^{N}\left(\hat{\mathbf{G}}_j\otimes\hat{\mathbf{G}}_j\right) vec\left(\hat{\boldsymbol{\Omega}}_{EE,j} - \bar{\boldsymbol{\Omega}}_{EE,j}\right) \\
&\quad + \frac{1}{N}\sum_{j=1}^{N}\left[\left(\hat{\mathbf{G}}_j\otimes\hat{\mathbf{G}}_j\right) - \left(\mathbf{G}_j^0\otimes\mathbf{G}_j^0\right)\right] vec\left(\bar{\boldsymbol{\Omega}}_{EE,j}\right) \\
&= \mathbf{d}_0 + \mathbf{d}_1 + \mathbf{d}_2
\end{aligned}$$

but

$$\|\mathbf{d}_1\| \leq \frac{1}{N}\sum_{j=1}^{N}\left\|\hat{\mathbf{G}}_j\otimes\hat{\mathbf{G}}_j\right\|\left(\left\|\hat{\boldsymbol{\Omega}}_{EE,j} - \boldsymbol{\Omega}_{EE,j}\right\| + \left\|\boldsymbol{\Omega}_{EE,j} - \bar{\boldsymbol{\Omega}}_{EE,j}\right\|\right) = o_p(1)$$

with $\boldsymbol{\Omega}_{EE,j} = T^{-1}\mathbf{E}_j'\mathbf{E}_j$, and

$$\|\mathbf{d}_2\| \leq \frac{1}{N}\sum_{j=1}^{N}\left\|\left(\hat{\mathbf{G}}_j\otimes\hat{\mathbf{G}}_j\right) - \left(\mathbf{G}_j^0\otimes\mathbf{G}_j^0\right)\right\|\left\|\bar{\boldsymbol{\Omega}}_{EE,j}\right\| = o_p(1)$$

by Lemma B.4, concluding that $\frac{1}{N}\sum_{j=1}^{N}\left(\hat{\mathbf{G}}_j\otimes\hat{\mathbf{G}}_j\right) vec\left(\hat{\boldsymbol{\Omega}}_{EE,j}\right) = \frac{1}{N}\sum_{j=1}^{N}\left(\mathbf{G}_j^0\otimes\mathbf{G}_j^0\right) vec\left(\bar{\boldsymbol{\Omega}}_{EE,j}\right) + o_p(1)$ and

$$\mathbf{h}_2 = \frac{1}{N}\sum_{i=1}^{N}\mathbf{R}_{\Gamma}^{-1}\boldsymbol{\Gamma}_i^0\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\left(\frac{1}{N}\sum_{j=1}^{N}\mathbf{G}_j^0\bar{\boldsymbol{\Omega}}_{EE,j}\mathbf{G}_j^{0\prime}\right)\left(\boldsymbol{\Upsilon}_N^0\right)^{-1}\frac{\mathbf{F}^{0\prime}\mathbf{u}_i}{T} + o_p(1). \tag{C.19}$$

$$\begin{aligned}
\mathbf{h}_3 &= \frac{1}{N}\sum_{i=1}^{N}\bar{\mathbf{\Omega}}_{VE,i}\mathbf{G}_i^{0\prime}\mathbf{R}^{-1\prime}\hat{\mathbf{\Upsilon}}_N^{-1}\boldsymbol{\lambda}_i^0 \\
&+\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\mathbf{\Omega}}_{VE,i}-\bar{\mathbf{\Omega}}_{VE,i}\right)\hat{\boldsymbol{G}}_i^{\prime}\hat{\mathbf{\Upsilon}}_N^{-1}\hat{\boldsymbol{\lambda}}_i \\
&+\frac{1}{N}\sum_{i=1}^{N}\bar{\mathbf{\Omega}}_{VE,i}\left(\hat{\boldsymbol{G}}_i-\mathbf{R}^{-1}\boldsymbol{G}_i^0\right)^{\prime}\hat{\mathbf{\Upsilon}}_N^{-1}\hat{\boldsymbol{\lambda}}_i \\
&= \mathbf{h}_{31}+\mathbf{h}_{32}+\mathbf{h}_{33}.
\end{aligned}$$

$$\begin{aligned}
\mathbf{h}_{32} &= \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\mathbf{\Omega}}_{VE,i}-\mathbf{\Omega}_{VE,i}\right)\hat{\boldsymbol{G}}_i^{\prime}\hat{\mathbf{\Upsilon}}_N^{-1}\hat{\boldsymbol{\lambda}}_i+\frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{\Omega}_{VE,i}-\bar{\mathbf{\Omega}}_{VE,i}\right)\hat{\boldsymbol{G}}_i^{\prime}\hat{\mathbf{\Upsilon}}_N^{-1}\hat{\boldsymbol{\lambda}}_i \\
&= \mathbf{j}_1+\mathbf{j}_2
\end{aligned}$$

$$\begin{aligned}
\mathbf{j}_1 &= \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\mathbf{\Omega}}_{VE,i}-\mathbf{\Omega}_{VE,i}\right)\mathbf{G}_i^{0\prime}\left(\mathbf{R}^{-1}\right)^{\prime}\hat{\mathbf{\Upsilon}}_N^{-1}\hat{\boldsymbol{\lambda}}_i \\
&+\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\mathbf{\Omega}}_{VE,i}-\mathbf{\Omega}_{VE,i}\right)\left(\hat{\boldsymbol{G}}_i-\mathbf{R}^{-1}\mathbf{G}_i^0\right)^{\prime}\hat{\mathbf{\Upsilon}}_N^{-1}\hat{\boldsymbol{\lambda}}_i \\
&= \mathbf{g}_1+\mathbf{g}_2
\end{aligned}$$

$$\begin{aligned}
\|\mathbf{g}_2\| &\leq \left(\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\mathbf{\Omega}}_{VE,i}-\mathbf{\Omega}_{VE,i}\right\|^4\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{\lambda}}_i\right\|^4\right)^{1/4}\left[\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{G}}_i-\mathbf{R}^{-1}\mathbf{G}_i^0\right\|^2\right]^{1/2}\left\|\hat{\mathbf{\Upsilon}}_N^{-1}\right\| \\
&= o_p\left(\delta_{NT}^{-1}\right)
\end{aligned}$$

as $\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{G}}_i-\mathbf{R}^{-1}\mathbf{G}_i^0\right\|^2 = O_p\left(\delta_{NT}^{-1}\right)$ and $\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\mathbf{\Omega}}_{VE,i}-\mathbf{\Omega}_{VE,i}\right\|^4 = O_p\left(\delta_{NT}^{-1}\right)$. In a similar manner, it is easily shown that $\mathbf{g}_1 = o_p(1)$. and also $\mathbf{j}_2 = o_p(1)$, thus, $\mathbf{h}_{32} = o_p(1)$. As $\left\|\bar{\mathbf{\Omega}}_{VE,i}\right\|^4 \leq \Delta$ and $E\left\|\hat{\boldsymbol{\lambda}}_i\right\|^4 \leq \Delta$, we have

$$\begin{aligned}
\|\mathbf{h}_{33}\| &\leq \frac{1}{N}\sum_{i=1}^{N}\left\|\bar{\mathbf{\Omega}}_{VE,i}\right\|\left\|\hat{\boldsymbol{G}}_i-\mathbf{R}^{-1}\boldsymbol{G}_i^0\right\|\hat{\mathbf{\Upsilon}}_N^{-1}\left\|\hat{\boldsymbol{\lambda}}_i\right\| \\
&\leq \left(\frac{1}{N}\sum_{i=1}^{N}\left\|\hat{\boldsymbol{G}}_i-\mathbf{R}^{-1}\boldsymbol{G}_i^0\right\|^2\right)^{1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\bar{\mathbf{\Omega}}_{VE,i}\right\|^2\left\|\hat{\boldsymbol{\lambda}}_i\right\|^2\right)^{1/2}\left\|\hat{\mathbf{\Upsilon}}_N^{-1}\right\| \\
&= O_p\left(\delta_{NT}^{-1}\right).
\end{aligned}$$

Finally *as* $\hat{\mathbf{\Upsilon}}_N-\mathbf{\Upsilon}_N^0 = o_p(1)$,

$$\mathbf{h}_3 = \frac{1}{N}\sum_{i=1}^{N}\bar{\mathbf{\Omega}}_{VE,i}\boldsymbol{G}_i^{0\prime}\mathbf{R}^{-1}\left(\mathbf{\Upsilon}_N^0\right)^{-1}\frac{\mathbf{F}^{0\prime}\mathbf{u}_i}{T}+o_p(1). \tag{C.20}$$

Using the above results, $\hat{\boldsymbol{\xi}}_{NT}-\boldsymbol{\xi}_{NT}^{\dagger}\rightarrow_p \mathbf{0}$ under slope heterogeneity and homogeneity. ■

Appendix D: Monte Carlo Supplements

Bias corrected estimator of Bai (2009) and our proposed covariance estimator

The estimator $\left(\hat{\boldsymbol{\beta}}_{Bai}, \hat{\mathbf{F}}_b\right)$ is the solution of the set of nonlinear equations

$$\hat{\boldsymbol{\beta}}_{Bai} = \left(\sum_{i=1}^{N} \mathbf{X}_i^{*\prime} \mathbf{M}_{\hat{F}_b} \mathbf{X}_i^*\right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i^{*\prime} \mathbf{M}_{\hat{F}_b} \mathbf{y}_i^* \text{ and } \left[\frac{1}{NT}\sum_{i=1}^{N}\left(\mathbf{y}_i^* - \mathbf{X}_i^* \hat{\boldsymbol{\beta}}_{bai}\right)\left(\mathbf{y}_i^* - \mathbf{X}_i^* \hat{\boldsymbol{\beta}}_{bai}\right)'\right] \hat{\mathbf{F}}_b = \hat{\mathbf{F}}_b \boldsymbol{\Phi}_{NT} \tag{C.1}$$

where $\boldsymbol{\Phi}_{NT}$ is a diagonal matrix that consists r largest eigenvalues of the above matrix in the brackets, arranged in descending order, with the restriction $\hat{\mathbf{F}}_b'\hat{\mathbf{F}}_b/T = \mathbf{I}_r$ and $\hat{\boldsymbol{\Lambda}}_b'\hat{\boldsymbol{\Lambda}}_b$ is diagonal, where $\hat{\boldsymbol{\Lambda}}_b = \left(\hat{\boldsymbol{\lambda}}_1, \hat{\boldsymbol{\lambda}}_2, ..., \hat{\boldsymbol{\lambda}}_N\right)'$, $\hat{\boldsymbol{\lambda}}_i = T^{-1}\hat{\mathbf{F}}_b'\left(\mathbf{y}_i^* - \mathbf{X}_i^*\hat{\boldsymbol{\beta}}_b\right)$. The solution $\left(\hat{\boldsymbol{\beta}}_b, \hat{\mathbf{F}}_b\right)$ is obtained by iteration until convergence. For the computation of variance estimator of $\hat{\boldsymbol{\beta}}_b$ and the bias correction, the following transformed $\mathbf{X}_i^*$ is used:

$$\tilde{\mathbf{X}}_{bi} = \mathbf{M}_{\hat{F}_b}\mathbf{X}_i^* - N^{-1}\sum_{j=1}^{N}\hat{a}_{ij}\mathbf{M}_{\hat{\mathbf{F}}_\mathbf{b}}\mathbf{X}_j^*,\ \hat{a}_{ij} = \hat{\boldsymbol{\lambda}}_i'\left(\hat{\boldsymbol{\Lambda}}_b'\hat{\boldsymbol{\Lambda}}_b/N\right)^{-1}\hat{\boldsymbol{\lambda}}_j. \tag{C.2}$$

The HAC variance estimator of $\hat{\boldsymbol{\beta}}_{Bai}$ is given by

$$\hat{\mathbf{V}}_b = \left(\sum_{i=1}^{N}\tilde{\mathbf{X}}_{bi}'\tilde{\mathbf{X}}_{bi}\right)^{-1}\sum_{i=1}^{N}\tilde{\mathbf{X}}_{bi}'\hat{\mathbf{u}}_{bi}\hat{\mathbf{u}}_{bi}'\tilde{\mathbf{X}}_{bi}\left(\sum_{i=1}^{N}\tilde{\mathbf{X}}_{bi}'\tilde{\mathbf{X}}_{bi}\right)^{-1} \tag{C.3}$$

with $\hat{\mathbf{u}}_{bi} = \mathbf{M}_{\hat{\mathbf{F}}_b}\left(\mathbf{y}_i^* - \mathbf{X}_i^*\hat{\boldsymbol{\beta}}_b\right) = \left(\hat{u}_{bi1}, ..., \hat{u}_{biT}\right)'$, and the bias corrected estimator is

$$\tilde{\boldsymbol{\beta}}_{Bai} = \hat{\boldsymbol{\beta}}_{Bai} - \frac{1}{N}\hat{\mathcal{B}} - \frac{1}{T}\hat{\mathcal{C}}, \tag{C.4}$$

where

$$\hat{\mathcal{B}} = -\tilde{\mathbf{A}}_{bNT}^{-1}\frac{1}{N}\sum_{i=1}^{N}\frac{\left(\mathbf{X}_i^* - N^{-1}\sum_{j=1}^{N}\hat{a}_{ij}\mathbf{X}_j^*\right)\hat{\mathbf{F}}_b}{T}\left(\frac{\hat{\boldsymbol{\Lambda}}_b'\hat{\boldsymbol{\Lambda}}_b}{N}\right)^{-1}\hat{\boldsymbol{\lambda}}_i\hat{\hat{\sigma}}_i^2 \tag{C.5}$$

$$\hat{\mathcal{C}} = -\tilde{\mathbf{A}}_{bNT}^{-1}\frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{X}_i^{*\prime}\widehat{\mathbf{M}_{\mathbf{F}_b}\boldsymbol{\Omega}}_b\mathbf{F}_b}{T}\left(\frac{\hat{\boldsymbol{\Lambda}}_b'\hat{\boldsymbol{\Lambda}}_b}{N}\right)^{-1}\hat{\boldsymbol{\lambda}}_i \tag{C.6}$$

$\hat{\hat{\sigma}}_i^2 = T^{-1}\hat{\mathbf{u}}_{bi}'\hat{\mathbf{u}}_{bi}$, $\tilde{\mathbf{A}}_{bNT} = \frac{1}{NT}\sum_{i=1}^{N}\tilde{\mathbf{X}}_{bi}'\tilde{\mathbf{X}}_{bi}$, letting $\mathbf{M}_{\hat{F}_b}\mathbf{X}_i^* = \hat{\mathbf{X}}_{b,i}$, then denote

$$\frac{\mathbf{X}_i^{*\prime}\widehat{\mathbf{M}_{\mathbf{F}^0}\boldsymbol{\Omega}}_b\mathbf{F}^0}{T} = \frac{1}{TN}\sum_{j=1}^{N}\left[\sum_{t=1}^{T}\hat{u}_{b,jt}^2\hat{\mathbf{x}}_{b,it}\hat{\mathbf{f}}_{b,t}' + \sum_{s=1}^{S}\sum_{t=s+1}^{T}\left(1 - \frac{s}{S+1}\right)\hat{u}_{b,jt}\hat{u}_{b,jt-s}\left(\hat{\mathbf{x}}_{b,it}\hat{\mathbf{f}}_{b,t-s}' + \hat{\mathbf{x}}_{b,it-s}\hat{\mathbf{f}}_{b,t}'\right)\right]. \tag{C.7}$$

We have chosen $S = \lfloor T^{1/4} \rfloor$.