Handling Distinct Correlated Effects with CCE

Stauskas, Ovidijus and De Vos, Ignace

BI Norwegian Business School, Department of Economics, Vrije Universiteit Amsterdam, Department of Economics and Data Science, Tinbergen Institute

17 May 2024

Handling Distinct Correlated Effects with CCE

Ovidijus Stauskas[1] and Ignace De Vos[2,3]

[1]BI Norwegian Business School, Department of Economics
[2]VU Amsterdam, Department of Econometrics and Data Science
[3]Tinbergen Institute

May 21, 2024

Abstract

The Common Correlated Effects (CCE) approach by Pesaran (2006) is a popular method for estimating panel data models with interactive effects. Due to its simplicity, i.e. unobserved common factors are approximated with cross-section averages of the observables, the estimator is highly flexible and lends itself to a wide range of applications. Despite such flexibility, however, properties of CCE estimators are typically only examined under the restrictive assumption that all the observed variables load on the same set of factors, which ensures joint identification of the factor space. In this paper, we take a different perspective, and explore the empirically relevant case where the dependent and explanatory variables are driven by distinct but correlated factors. Hence, we consider the case of *Distinct Correlated Effects*. Such settings can be argued to be relevant for practice, for instance in studies linking economic growth to climatic variables. In so doing, we consider panel dimensions such that $TN^{-1} \to \tau < \infty$ as $(N, T) \to \infty$, which is known to induce an asymptotic bias for the pooled CCE estimator even under the usual common factor assumption. We subsequently develop a robust boostrap-based toolbox that enables asymptotically valid inference in both homogeneous and heterogeneous panels, without requiring knowledge about whether factors are distinct or common.

JEL classification: C33, C38, C15
Keywords: panel data, bootstrap, interactive effects, CCE, factors, information criterion

1 Introduction

Consider the interactive effects model for unit $i = 1, \ldots, N$ and period $t = 1, \ldots, T$, where $y_{i,t} \in \mathbb{R}$, $\mathbf{x}_{i,t} \in \mathbb{R}^k$ and $\varepsilon_{i,t}$ is a mean zero, weakly dependent idiosyncratic innovation:

$$y_{i,t} = \boldsymbol{\beta}'\mathbf{x}_{i,t} + e_{i,t}, \qquad e_{i,t} = \boldsymbol{\gamma}_i'\mathbf{f}_t + \varepsilon_{i,t}, \tag{1.1}$$

Equation (1.1) defines a multi-factor error structure, where the panel units exhibit "strong" cross-section dependence (see e.g. Chudik et al., 2011) due to common unobserved factors $\mathbf{f}_t \in \mathbb{R}^m$ to which they respond with heterogeneous intensities (loadings) $\boldsymbol{\gamma}_i \in \mathbb{R}^m$. Interactive effects come natural in macroeconomic applications with panel data where both N and T are large. For instance, $\mathbf{f}_t$ may represent the unobserved global technological progress, where $\boldsymbol{\gamma}_i$ is the local absorption intensity (see e.g. Eberhardt and Teal, 2011). For micro applications, see for instance Westerlund et al. (2019).

In practice, $\mathbf{f}_t$ is typically correlated with $\mathbf{x}_{i,t}$. Pesaran (2006), and many subsequent studies, allow for this possibilty by explicitly letting the regressors be driven by the same factors, $\mathbf{f}_t$:

$$\mathbf{x}_{i,t} = \boldsymbol{\Gamma}_i'\mathbf{f}_t + \mathbf{v}_{i,t} \tag{1.2}$$

where $\boldsymbol{\Gamma}_i \in \mathbb{R}^{m \times k}$ is the loading matrix and $\mathbf{v}_{i,t} \in \mathbb{R}^k$ is the vector of idiosyncratic innovations. Model (1.1) - (1.2) then exhibits not only strong cross-section dependence, but also endogeneity, thus making it essential to control for $\mathbf{f}_t$ in the estimation of $\boldsymbol{\beta}$. Under the assumption that factors are common and the matrix of average loadings $\overline{\mathbf{C}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{C}_i$, with $\mathbf{C}_i = [\boldsymbol{\gamma}_i + \boldsymbol{\Gamma}_i\boldsymbol{\beta}]$, has at least rank m, this is easy to achieve with the Common Correlated Effects (CCE) approach of Pesaran (2006), which estimates the factor space with the cross-section averages (CAs) of the observables $\widehat{\mathbf{f}}_t = \bar{\mathbf{z}}_t = \frac{1}{N}\sum_{i=1}^{N}\mathbf{z}_{i,t}$, where $\mathbf{z}_{i,t} = [y_{i,t}, \mathbf{x}_{i,t}']' \in \mathbb{R}^{k+1}$. The latter is then added as a regressor to (1.1), which is in turn estimated by Least Squares (LS). The resulting estimator is consistent as $N \to \infty$ and exhibits excellent small sample performance (see e.g. Westerlund and Urbain, 2015). It has accordingly been applied and extended to various more general settings, such as structural break modelling or unit root testing (see e.g. Karavias et al., 2023, and Norkutė and Westerlund, 2021).

In the standard CCE model, the set of CAs $\bar{\mathbf{z}}_t$ are sufficiently informative for the m unobserved factors when $rk(\overline{\mathbf{C}}) = m$, and $\widehat{\mathbf{f}}_t$ is then consistent for (the space spanned by) $\mathbf{f}_t$, which is sufficient for consistency. This assumption can be verified with the procedure in De Vos et al. (2024). Adding $\widehat{\mathbf{f}}_t$ as observed regressors and estimating the resulting model with LS then yields consistent estimates of $\boldsymbol{\beta}$ as $N \to \infty$, for T fixed or growing (see Westerlund et al., 2019). Asymptotic normal inference ensues provided $TN^{-1} \to 0$, whereas if $TN^{-1} \to \tau < \infty$, a bias-correction is unavoidably needed due to the accumulation of factor estimation error (see e.g. Westerlund and Urbain, 2015). This is, however, not straightforward, as the specific structure (functional form) of the bias depends on whether the number of employed averages (g) exceeds or matches the number of factors (m), because in the former case $g - m$ CAs are redundant and produce nuisance parameters (see Karabiyik et al., 2017). This makes the resulting asymptotic bias exceedingly difficult to remedy with analytical corrections, as not all bias-components (or the functional form) are known or consistently estimable, unless in specific and restrictive settings. De Vos and Stauskas (2024) therefore provide a consistent bootstrap correction to sidestep the issue and remedy the bias problem without knowledge of g and m.

Notwithstanding, the common assumption in most theoretical work so far, is that all the observed variables in (1.1)-(1.2) are driven by the same factors, thus enabling a straightforward joint estimation of the factor space. In this paper, we challenge this assumption, and investigate properties of the CCE approach when $y_{i,t}$ and $\mathbf{x}_{i,t}$ may be driven by distinct but correlated factors. We refer to this setting as *Distinct Correlated Effects* (DCE), as opposed to the standard *Common Correlated Effects* (CCE) assumption. The setting is easily seen to be empirically relevant, as it is not always reasonable to expect factors to be common over all observables. Consider for example a regression of economic growth on climatic variables in the spirit of Dell et al. (2012). Unobserved factors underlying the climatic regressors (e.g. global climate patterns and trends) are likely to be distinct from those directly affecting economic growth (e.g. technological progress, productivity, business cycles, crises, according to economic theory). The two sets of factors are, however, likely to be correlated (climatic hardship drives technological innovation), thus consistency still requires the unobserved factor space in either $y_{i,t}$ or the regressors to be controlled for. The asymptotic behavior of the CCE estimators in this case is, however, largely unknown, so we relax in this paper the common factor assumption, and establish properties and solutions for CCE estimation in practice.

To make the above discussion a little more precise, we depart from (1.1)-(1.2) by following Cui et al. (2022) or De Vos and Stauskas (2024), and let

$$y_{i,t} = \boldsymbol{\beta}'\mathbf{x}_{i,t} + \boldsymbol{\gamma}_i'\mathbf{f}_{\mathbf{y},t} + \varepsilon_{i,t}, \tag{1.3}$$

$$\mathbf{x}_{i,t} = \boldsymbol{\Gamma}_i'\mathbf{f}_{\mathbf{x},t} + \mathbf{v}_{i,t} \tag{1.4}$$

such that $\mathbf{f}_{\mathbf{y},t} \in \mathbb{R}^{m_y}$ and $\mathbf{f}_{\mathbf{x},t} \in \mathbb{R}^{m_x}$ denote respectively the m_y factors affecting the regressand, and the m_x factors affecting the regressors. The total number of factors is then $m = m_y + m_x$ and gathered in $\mathbf{f}_t = [\mathbf{f}_{\mathbf{y},t}', \mathbf{f}_{\mathbf{x},t}']'$. We also explicitly allow that $\mathbb{C}ov(\mathbf{f}_{\mathbf{y},t}, \mathbf{f}_{\mathbf{x},t}) \neq \mathbf{0}_{m_y \times m_x}$, such that factors can be correlated. We will focus in particular on the case where $\mathbf{f}_{\mathbf{y},t} \cap \mathbf{f}_{\mathbf{x},t} = \varnothing$, as it is the most extreme/challenging case for CCE, thereby making the conclusions most relevant for practice. That is, solutions for the former will also allow consistent inference when $\mathbf{f}_{\mathbf{x},t} \subseteq \mathbf{f}_{\mathbf{y},t}$. It is now easy to see that in the Distinct Correlated Effects setting, the full factor space $\mathbf{f}_t$

is not generally estimable by the CAs. That is, if $rk(\overline{\boldsymbol{\Gamma}}) = m_x$, then $\overline{\mathbf{x}}_t$ is consistent for the space spanned by $\mathbf{f}_{\mathbf{x},t}$. However, since $\mathbf{f}_{\mathbf{y},t}$ loads on $y_{i,t}$ only, taking cross-section averages of (1.3) and rearranging implies that the estimating equation for $\mathbf{f}_{\mathbf{y},t}$ would in principle be:

$$\overline{\boldsymbol{\gamma}}'\mathbf{f}_{\mathbf{y},t} = \overline{y}_t - \boldsymbol{\beta}'\overline{\mathbf{x}}_t - \overline{\varepsilon}_t,$$
$$\mathbf{f}_{\mathbf{y},t} = (\overline{\boldsymbol{\gamma}}\overline{\boldsymbol{\gamma}}')^{-1}\overline{\boldsymbol{\gamma}}(\overline{y}_t - \boldsymbol{\beta}'\overline{\mathbf{x}}_t) + O_p(N^{-1/2})$$

since $\overline{\varepsilon}_t = O_p(N^{-1/2})$ under our assumptions. Yet, since $\overline{\boldsymbol{\gamma}} \in \mathbb{R}^{m_y \times 1}$, we have that $rk(\overline{\boldsymbol{\gamma}}) \leq 1$, so that the inverse $(\overline{\boldsymbol{\gamma}}\overline{\boldsymbol{\gamma}}')^{-1}$ does not exist when $m_y > 1$. In effect, the rank condition is not generally satisfied for $\mathbf{f}_{\mathbf{y},t}$ in this distinct factor setting, and the latter factors cannot be estimated. $\mathbf{f}_{\mathbf{y},t}$ is thus only estimable with CCE in the unlikely case that $m_y = 1$, or $\mathbf{f}_{\mathbf{y},t} = \mathbf{f}_{\mathbf{x},t} = \mathbf{f}_t$ (common factors). Since neither of these special cases is also easy to verify in practice, the properties of CCE need to be verified and a generally robust approach is needed for inference. This is the objective of the current paper.

Several versions of the distinct factor case have been considered in the literature with clear advantages and drawbacks. For example, Bai (2009) or Moon and Weidner (2015) assume no particular model (or factor space) for $\mathbf{x}_{i,t}$, and focus on estimating the factors in $e_{i,t}$ from (1.1) only, with Principal Components (PC). While flexible, this approach relies on a non-linear optimization problem, therefore convergence issues may arise (see e.g. Jiang et al., 2021). For CCE estimators, Juodis (2022) considers $\mathbf{f}_t = [\mathbf{f}'_{1,t}, \mathbf{f}'_{2,t}]'$ that drives $\mathbf{x}_{i,t}$, while $y_{i,t}$ loads on $\mathbf{f}_{1,t}$ only, which can be nested in (1.1) - (1.2). Here, $\mathbf{f}_{2,t}$ is not estimable from the CAs since its average loading has zero rank, therefore the problem differs from ours. The setup closest to ours is discussed in Cui et al. (2022), who aim to produce an unbiased estimator of $\boldsymbol{\beta}$ with the Two Stage Instrumental Variable (2SIV) approach. Specifically, $\mathbf{f}_{\mathbf{x},t}$ is estimated with PC, and $\mathbf{x}_{i,t}$ is purged of their effect thus "de-correlating" it with $e_{i,t}$ in (1.1) and ensuring consistency (see their Proposition 3.1). Next, PC is applied to the first stage residuals $y_{i,t} - \widehat{\boldsymbol{\beta}}'\mathbf{x}_{i,t}$ to extract $\mathbf{f}_{\mathbf{y},t}$. This leads to the second stage, where $\mathbf{f}_{\mathbf{y},t}$ is asymptotically purged ensuring an asymptotically standard normal inference.

The CCE estimators are very commonly applied in practice, and it turns out the latter strategy is partially feasible for CCE estimators too for solving the problem of an uninformative $\overline{y}_t$. As the rank of $\overline{\boldsymbol{\Gamma}}$ is m_x by assumption, it validates estimating $\mathbf{f}_{\mathbf{x},t}$ with $\overline{\mathbf{x}}_t$ and performing the (first stage) de-correlation step to make CCE consistent. We will also follow this approach. The second stage purge, however, is not generally feasible with CCE, but also not necessary for consistency. The consequences are that $\mathbf{f}_{\mathbf{y},t}$ will remain in the residuals of the model, which in turn necessitates the bootstrap for valid inference. To a limited degree, this route was already taken in Proposition 1 of De Vos and Stauskas (2024) to illustrate the possibilities of the panel cross-section (CS) bootstrap scheme by Kapetanios (2008). The key finding is that

$\mathbf{f}_{\mathbf{y},t}$ renders the asymptotic distribution of CCE non-standard if $m_x < g$, because the excess CAs have a non-trivial effect (see a similar finding in Juodis, 2022). They also demonstrate that the variance and the bias of the asymptotic distribution depend on the unknown $\mathbb{C}ov(\mathbf{f}_{\mathbf{x},t}, \mathbf{f}_{\mathbf{y},t})$, which renders both the bias and variance analytically inestimable. In turn, they establish conditions under which the CS bootstrap is able to replicate this distribution. As a result, this re-enables estimation of the asymptotic variance and remedies bias under the usual $TN^{-1} \to \tau < \infty$ asymptotics, in the spirit of Gonçalves and Perron (2014) or Djogbenou et al. (2015). However, the analysis of distinct factors in De Vos and Stauskas (2024) is restricted to homogeneous $\boldsymbol{\beta}$ estimated with the pooled CCE (CCEP) estimator and with covariance stationary $\mathbf{f}_t$, which somewhat limits the generality.

The contribution of the current study is thus the development of the CCE methodology in the Distinct Correlated Effects setting. This involves extending the CCE methodology to handle distinct factors in heterogeneous panels for CCEP and Mean Group (CCEMG) estimators, while establishing an inferential bootstrap toolbox that is possibly also robust to deviations from stationarity. The key result is that while the standard asymptotic tools and variance estimators may fail depending on slope heterogeneity, the asymptotic distributions and biases, if present, can in each case be captured by the proposed CS bootstrap tools. This leads to a powerful outcome: asymptotically valid (bootstrap-aided) inference can ensue under uninformative $\overline{y}_t$, under homogeneity or heterogeneity, so long as the rank of $\overline{\mathbf{\Gamma}}$ is m_x. The latter, we note, can be verified with De Vos et al. (2024). This significantly boosts applicability of the CCE methods.

This paper is organized as follows: Section 2 presents our assumptions, the details on CCEP and CCEMG and explains the CS bootstrap scheme. In Section 3, we derive the asymptotic distribution of both estimators in the original and bootstrap samples and discuss inference. Monte Carlo evidence and a comparison to 2SIV approach by Cui et al. (2022) are provided in Section 4. We use the following notation: $\text{rk}(\mathbf{A})$, $\det(\mathbf{A})$ and $\text{tr}(\mathbf{A})$ denote respectively the rank, determinant, and trace of an arbitrary matrix $\mathbf{A}$, while $vec(\mathbf{A})$ vectorizes $\mathbf{A}$ by stacking its columns on top of each other. $\|\mathbf{A}\| = \sqrt{\text{tr}(\mathbf{A}'\mathbf{A})}$ is the Frobenius (Euclidean) norm, while '$\to_d$' stands for convergence in distribution. By $\text{diag}(\mathbf{A}, \mathbf{B})$, we represent a matrix with $\mathbf{A}$ and $\mathbf{B}$ as diagonal blocks. The symbols $\to_{p^*}$ ($\to_p$) and $\to_{d^*}$ ($\to_d$) represent convergence in probability and convergence in distribution with respect to the induced (generic) probability measure.

2 Econometric Setup

2.1 Assumptions and Estimation

Consider model (1.3) - (1.4) in time-stacked notation for $i = 1, \dots, N$:

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{F}_\mathbf{y}\gamma_i + \varepsilon_i, \tag{2.1}$$

$$\mathbf{X}_i = \mathbf{F}_\mathbf{x}\boldsymbol{\Gamma}_i + \mathbf{V}_i, \tag{2.2}$$

such that the set of observables is $\mathbf{Z}_i = [\mathbf{y}_i, \mathbf{X}_i]$, where $\mathbf{y}_i = [y_{i,1}, \dots, y_{i,T}]' \in \mathbb{R}^{T\times 1}$, $\mathbf{X}_i = [\mathbf{x}_{i,1}, \dots, \mathbf{x}_{i,T}]' \in \mathbb{R}^{T\times k}$, $\mathbf{V}_i = [\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,T}]' \in \mathbb{R}^{T\times k}$ and $\boldsymbol{\varepsilon}_i = [\varepsilon_{i,1}, \dots, \varepsilon_{i,T}]' \in \mathbb{R}^{T\times 1}$. Let also $\mathbf{F} = [\mathbf{F}_\mathbf{y}, \mathbf{F}_\mathbf{x}] \in \mathbb{R}^{T\times m}$, with $\mathbf{F}_\mathbf{y} = [\mathbf{f}_{\mathbf{y},1}, \dots, \mathbf{f}_{\mathbf{y},T}]' \in \mathbb{R}^{T\times m_y}$ and $\mathbf{F}_\mathbf{x} = [\mathbf{f}_{\mathbf{x},1}, \dots, \mathbf{f}_{\mathbf{x},T}]' \in \mathbb{R}^{T\times m_x}$. Since we focus on the likely case with $m_y > 1$, $\overline{\mathbf{y}}$ is not generally sufficiently informative to estimate the full $\mathbf{F}_\mathbf{y}$. Therefore, we propose to instead de-correlate the regressors with $\mathbf{F}_{\mathbf{y},t}$ by projecting out the estimated $\mathbf{F}_\mathbf{x}$. The latter factor space can in the Distinct Correlated Effects setting be estimated with $\overline{\mathbf{X}}$, since averaging (2.2) over units gives:

$$\widehat{\mathbf{F}}_\mathbf{x} = \overline{\mathbf{X}} = \mathbf{F}_\mathbf{x}\overline{\boldsymbol{\Gamma}} + \overline{\mathbf{V}}, \tag{2.3}$$

which implies, assuming $\text{rk}(\overline{\boldsymbol{\Gamma}}) = m_x$, that

$$\mathbf{F}_\mathbf{x} = (\widehat{\mathbf{F}}_\mathbf{x} - \overline{\mathbf{V}})\overline{\boldsymbol{\Gamma}}^+, \tag{2.4}$$

where $\overline{\boldsymbol{\Gamma}}^+$ is the MP inverse of $\overline{\boldsymbol{\Gamma}}$. While the above estimator is consistent for $\mathbf{F}_\mathbf{x}$, we note that this does not necessarily require all the CAs in $\overline{\mathbf{X}}$. As such, we accomodate the use of subsets of the CA (or IC selection of averages) by defining $\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} \in \mathbb{R}^{T\times g}$ as a selection of g averages from $\overline{\mathbf{X}}$ by the $k \times g$ selector $\mathbf{q}_{\dot{\mathbf{x}}}$:

$$\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} = \overline{\mathbf{X}}\mathbf{q}_{\dot{\mathbf{x}}} = \mathbf{F}_\mathbf{x}\overline{\boldsymbol{\Gamma}}\mathbf{q}_{\dot{\mathbf{x}}} + \overline{\mathbf{V}}\mathbf{q}_{\dot{\mathbf{x}}}$$

For a given subset, the corresponding rank condition needed for consistency is then $rk(\overline{\boldsymbol{\Gamma}}\mathbf{q}_{\dot{\mathbf{x}}}) = m_x$, as in Assumption 4 below. This implies that the chosen subset is sufficiently informative on $\mathbf{F}_\mathbf{x}$.

For the analysis, we work under the following assumptions:

Assumption 1 *(Idiosyncratic errors) $\varepsilon_{i,t}$ and $\mathbf{v}_{i,t}$ are stationary variables, independent across i with $\mathbb{E}(\varepsilon_{i,t}) = 0$, $\mathbb{E}(\mathbf{v}_{i,t}) = \mathbf{0}_{k\times 1}$, $\sigma_i^2 = \mathbb{E}(\varepsilon_{i,t}^2)$, $\boldsymbol{\Sigma}_i = \mathbb{E}(\mathbf{v}_{i,t}\mathbf{v}_{i,t}')$, $\boldsymbol{\Omega}_i = \mathbb{E}(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i')$, with $\boldsymbol{\Omega}_i, \boldsymbol{\Sigma}_i$ positive definite and $\mathbb{E}(\varepsilon_{i,t}^6) < \infty$, $\mathbb{E}(\|\mathbf{v}_{i,t}\|^6) < \infty$ for all i and t. Additionally, let $\tilde{\mathbf{u}}_{i,t} = (\varepsilon_{i,t}, \mathbf{v}_{i,t}')'$. Then*

$$\frac{1}{T^3}\sum_{t=1}^T\sum_{q=1}^T\sum_{r=1}^T\sum_{s=1}^T\|\mathbb{E}(\tilde{\mathbf{u}}_{i,t}\tilde{\mathbf{u}}_{i,q}'\tilde{\mathbf{u}}_{i,r}\tilde{\mathbf{u}}_{i,s}')\| = O(1), \quad \frac{1}{T}\sum_{t=1}^T\sum_{s=1}^T\|\mathbb{E}(\tilde{\mathbf{u}}_{i,t}\tilde{\mathbf{u}}_{i,s}')\| = O(1)$$

as $T \to \infty$, whereas $\frac{1}{N}\sum_{i=1}^N \sigma_i^2 \to \sigma^2 < \infty$ and $\frac{1}{N}\sum_{i=1}^N \boldsymbol{\Sigma}_i \to \boldsymbol{\Sigma} < \infty$ as $N \to \infty$.

Assumption 2 *(Distinct factors) Let* $\mathbf{f}_t = (\mathbf{f}'_{\mathbf{y}}, \mathbf{f}'_{\mathbf{x}})'$ *be covariance stationary with* $\mathbb{E}(\|\mathbf{f}_t\|^4) < \infty$*, absolute summable autocovariances and* $T^{-1}\mathbf{F}'\mathbf{F} \to^p \mathbf{\Sigma_F}$ *as* $T \to \infty$*, such that*

$$\mathbf{\Sigma_F} = \begin{bmatrix} \mathbf{\Sigma}_{\mathbf{F}_y} & \mathbf{\Sigma}'_{\mathbf{F}_{\mathbf{x,y}}} \\ \mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x,y}}} & \mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}} \end{bmatrix}$$

with $\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x},y}} = \text{plim}_{T\to\infty} T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{y}}$ *denoting the covariance between* $\mathbf{F_x}$ *and* $\mathbf{F_y}$*. Also* $\mathbf{\Sigma}_{\mathbf{F_x}}$ *and and* $\mathbf{\Sigma}_{\mathbf{F_y}}$ *are positive definite.*

Assumption 3 *(Factor loadings) The factor loadings are given by*

$$\begin{aligned} \boldsymbol{\gamma}_i &= \boldsymbol{\gamma} + \boldsymbol{\eta}_{\gamma,i} \qquad \boldsymbol{\eta}_{\gamma,i} \sim IID(\mathbf{0}_{m_y\times 1}, \mathbf{\Omega}_\gamma) \\ \mathbf{\Gamma}_i &= \mathbf{\Gamma} + \boldsymbol{\eta}_{\Gamma,i} \qquad vec(\boldsymbol{\eta}_{\Gamma,i}) \sim IID(\mathbf{0}_{km_x\times 1}, \mathbf{\Omega}_\Gamma) \end{aligned}$$

where $\boldsymbol{\gamma}, \mathbf{\Gamma}$ *are constant matrices,* $\mathbf{\Sigma}_{\gamma\Gamma} = \mathbb{E}(\boldsymbol{\eta}_{\gamma,i} \otimes \boldsymbol{\eta}_{\Gamma,i})$ *is a covariance matrix,* $\boldsymbol{\eta}_{\gamma,i}, \boldsymbol{\eta}_{\Gamma,i}$ *are independent across i and of the other model components, and* $\|\boldsymbol{\gamma}\|, \|\mathbf{\Gamma}\|, \|\mathbf{\Sigma}_{\gamma\Gamma}\|, \|\mathbf{\Omega}_\gamma\|, \|\mathbf{\Omega}_\Gamma\|$ *are finite.*

Assumption 4 *(Rank condition)* $\text{rk}(\overline{\mathbf{\Gamma}}\mathbf{q}_{\dot{\mathbf{x}}}) = m_x$*, with* $\mathbf{q}_{\dot{\mathbf{x}}}$ *a* $k \times g$ *selector matrix.*

Assumption 5 *(Independence)* $\mathbf{f}_t, \varepsilon_{i,s}, \mathbf{v}_{j,l}, \widetilde{\boldsymbol{\eta}}_n$ *are mutually independent for all* i, j, n, t, s, l.

Assumption 6 *(Slope heterogeneity) The slopes* $\boldsymbol{\beta}_i$ *follow*

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{v}_i, \qquad \boldsymbol{v}_i \sim IID(\mathbf{0}_{k\times 1}, \mathbf{\Omega}_{\boldsymbol{v}})$$

with $\mathbf{\Omega}_{\boldsymbol{v}}$ *a finite nonnegative definite* $k \times k$ *matrix and the* $\boldsymbol{v}_i$ *are independent of* $\mathbf{f}_t, \varepsilon_{i,s}, \mathbf{v}_{j,l}, \widetilde{\boldsymbol{\eta}}_n$ *for all* i, j, n, t, s, l.

Assumption 7 *(Identification)* $\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i} = T^{-1}\mathbf{X}'_i\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i$*, with* $\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} = \overline{\mathbf{X}}\mathbf{q}_{\dot{\mathbf{x}}}$*, is non-singular for all* N, T*, and*

$$\mathbb{E}\left(\left\|(T^{-1}\mathbf{V}'_i\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{V}_i)^{-1}\right\|^2\right) < \infty$$

also when $\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} = \mathbf{F_x}$*, where* $\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} = \mathbf{I}_T - \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}(\widehat{\mathbf{F}}'_{\dot{\mathbf{x}}}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^+\widehat{\mathbf{F}}'_{\dot{\mathbf{x}}}$.

The above assumptions are similar to those in Pesaran (2006); Karabiyik et al. (2017) or Westerlund (2018). Assumption 1, however, generalizes the aforementioned studies by allowing the idiosyncratic innovations $\mathbf{v}_{i,t}$ and $\varepsilon_{i,t}$ to be both serially correlated and heteroskedastic, unlike in e.g. Karabiyik et al. (2017). The combination of time series dependence and our $TN^{-1} \to \tau < \infty$ asymptotics also necessitates some stronger requirements, as reflected in the additional summability conditions for higher moments given in Assumption 1. Assumption 2 imposes covariance stationarity on the factors specific to the dependent and explanatory variables and is similar to the one in Cui et al. (2022). Later we relax this requirement. Assumption 3 also generalizes Pesaran (2006) by allowing the loadings to be correlated within, but not between, individuals. Next, Assumption 4 enables a flexible specification of the (CAs included in the) factor estimator through the selector matrix $\mathbf{q}_{\dot{\mathbf{x}}} \in \mathbb{R}^{k\times g}$, and thus avoids the restriction in our theory that CAs of *all* the explanatory variables are necessarily required in the CCE specifications. This corresponds to practice where some observables (e.g. dummy variables, or regressors with low (cross-section) variation) are excluded from the set of CA to enable computation and identification (see e.g. Westerlund and Petrova, 2018; De Vos and Westerlund, 2019, for examples). Assumption 6 formalizes the slope heterogeneity, while Assumption 7 is sufficient for identification of the mean $\boldsymbol{\beta}$ when the slopes are heterogeneous.

We next define the CCEP and CCEMG estimators as a function of a given dataset and specification. Letting $\mathcal{B} = \{\mathbf{Z}_i\}_{i=1}^N$ denote the observed dataset, and defining $\overline{\mathbf{Q}}_{\dot{\mathbf{x}}} = \frac{1}{N}\sum_{i=1}^N \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}$, we have respectively

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} = \widehat{\boldsymbol{\beta}}_{CCEP}(\dot{\mathbf{x}},\mathcal{B}) &= \overline{\mathbf{Q}}_{\dot{\mathbf{x}}}^{-1}\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i, \\
&= \boldsymbol{\beta} + \overline{\mathbf{Q}}_{\dot{\mathbf{x}}}^{-1}\frac{1}{NT}\sum_{i=1}^N \left(\mathbb{I}_{\boldsymbol{\nu}\neq\mathbf{0}} \times \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i + \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right) \quad (2.5)
\end{aligned}
$$

and

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} = \widehat{\boldsymbol{\beta}}_{CCEMG}(\dot{\mathbf{x}},\mathcal{B}) &= \frac{1}{NT}\sum_{i=1}^N \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&= \boldsymbol{\beta} + \mathbb{I}_{\boldsymbol{\nu}\neq\mathbf{0}} \times \overline{\boldsymbol{\nu}} + \frac{1}{NT}\sum_{i=1}^N \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{-1}(\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i), \quad (2.6)
\end{aligned}
$$

where $\overline{\boldsymbol{\nu}} = \frac{1}{N}\sum_{i=1}^N \boldsymbol{\nu}_i$, and the $\dot{\mathbf{x}}$ subscript refers to the specification of the CAs ($\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}$). $\mathbb{I}_{\boldsymbol{\nu}\neq\mathbf{0}}$ is an indicator function which equals $\mathbf{I}_k$ or $\mathbf{0}_{k\times k}$ depending on whether the slopes are heterogeneous or not.[1] The estimators of the asymptotic variance suggested by Pesaran (2006) depend similarly on the chosen averages, and are defined

[1] Note that if $\mathbf{q}_{\dot{\mathbf{x}}} = \mathbf{I}_k$, such that the whole $\overline{\mathbf{X}}$ is employed, then (2.5) simplifies by noticing that

$$
\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}(\boldsymbol{\gamma}+\boldsymbol{\eta}_{\gamma,i}) = \frac{1}{T}\overline{\mathbf{X}}'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma} + \frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} = \frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i},
$$

since then $\overline{\mathbf{X}}'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} = \mathbf{0}_{k\times T}$. We conduct our analysis for the upcoming theorems with an arbitrary $\mathbf{q}_{\dot{\mathbf{x}}}$ as long as the rank condition is satisfied to accommodate general choices.

as:

$$\widehat{\boldsymbol{\Theta}}_{CCEP,\dot{\mathbf{x}}} = \overline{\mathbf{Q}}_{\dot{\mathbf{x}}}^{-1} \left(\frac{1}{N(N-1)} \sum_{i=1}^{N} \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i} \widehat{\boldsymbol{\nu}}_i \widehat{\boldsymbol{\nu}}_i' \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i} \right) \overline{\mathbf{Q}}_{\dot{\mathbf{x}}}^{-1}, \tag{2.7}$$

$$\widehat{\boldsymbol{\Theta}}_{CCEMG,\dot{\mathbf{x}}} = \frac{1}{N(N-1)} \sum_{i=1}^{N} \widehat{\boldsymbol{\nu}}_i \widehat{\boldsymbol{\nu}}_i' \tag{2.8}$$

with $\widehat{\boldsymbol{\nu}}_i = \widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}$, for the CCEP and CCEMG estimator, respectively.

Clearly, the expansions in (2.5) and (2.6) reveal that $\mathbf{F}_{\mathbf{y}}$ non-trivially enters the asymptotic analysis of both estimators. Intuitively, since $\mathbf{F}_{\mathbf{y}}$ is not projected out (as it is non-estimable), it will affect the asymptotic distribution by altering the variance and possibly the mean (since $\mathbf{F}_{\mathbf{y}}$ is typically not mean-zero). Moreover, because m_y is unknown and likely to be bigger than 1, we also run the risk of having more factors than available CAs. In order to handle the consequences of this deviation from the standard CCE assumption, we propose the cross-section (CS) bootstrap approach established by De Vos and Stauskas (2024) for CCE estimators in $(N, T) \to \infty$ panels. We begin with a general description and outline the practical implementation of the resampling scheme.

2.2 Bootstrap Algorithm

The CS bootstrap scheme is straightforward to implement, and has the advantage that all factors in the data are automatically replicated in the bootstrap realm, without requiring a decision or knowledge about distinct vs. common factors by the researcher. Given the need to approximate the asymptotic distribution in both cases, this is an important advantage for practice. The core assumption behind the CS resampling algorithm is that $N \to \infty$ and that $\mathbf{Z}_i, \mathbf{Z}_j$ are independent for each i and $j \neq i$, *conditional* on $\sigma\{\mathbf{F}\}$. That is, the cross-section correlation in the data is due to the unobserved factors. To present the resampling scheme, recall that $\mathcal{B} = \{\mathbf{Z}_i\}_{i=1}^N$ denotes the original dataset, and let $\mathcal{B}_b^* = \{\mathbf{Z}_i^*\}_{i=1}^N$ denote bootstrap sample $b = 1, \ldots, B$, obtained as described in Algorithm 1 below. Accordingly, for $s \in \{CCEP, CCEMG\}$, we use $\widehat{\boldsymbol{\beta}}_{s,b}^* = \widehat{\boldsymbol{\beta}}_s(\dot{\mathbf{x}}, \mathcal{B}_b^*)$ to denote the estimates in bootstrap sample b following the specification $\dot{\mathbf{x}}$.

Algorithm 1: Cross-section resampling scheme.

1) Initialization: Estimate given the chosen specification $\mathbf{q}_{\dot{\mathbf{x}}}$ and estimator s the $\widehat{\boldsymbol{\beta}}_s = \widehat{\boldsymbol{\beta}}_s(\dot{\mathbf{x}}, \mathcal{B})$ based on the original sample.

2) for $b = 1 : B$ do:

   i) Generate $\mathcal{B}_b^* = \{\mathbf{Z}_i^*\}_{i=1}^N$ according to

   $$\mathbf{Z}_i^* = \mathbf{Z}_{i^*} \qquad for \qquad i = 1, \ldots, N$$

   where i^* is for each i an independent random draw from $\mathcal{I} = \{1, \ldots, N\}$.

   ii) Obtain $\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^* = \overline{\mathbf{X}}^* \mathbf{q}_{\dot{\mathbf{x}}}$ and estimate $\widehat{\boldsymbol{\beta}}_{s,b}^* = \widehat{\boldsymbol{\beta}}_s(\dot{\mathbf{x}}, \mathcal{B}_b^*)$

3) Save the results $\mathbf{B}_s^* = [\widehat{\boldsymbol{\beta}}_{s,1}^*, \ldots, \widehat{\boldsymbol{\beta}}_{s,B}^*]$ and form the following confidence interval widely used in the bootstrap literature (see Davison and Hinkley, 1997, p. 194) to test the null $\boldsymbol{\beta}_0$:

$$CI(\alpha, \widehat{\boldsymbol{\beta}}_{s,\dot{\mathbf{x}}}^*) = \left[\, 2\widehat{\boldsymbol{\beta}}_{s,\dot{\mathbf{x}}} - \boldsymbol{\theta}_{(1-\alpha/2)}^*(\mathbf{B}_s^*)\,,\, 2\widehat{\boldsymbol{\beta}}_{s,\dot{\mathbf{x}}} - \boldsymbol{\theta}_{\alpha/2}^*(\mathbf{B}_s^*)\, \right], \tag{2.9}$$

where $\boldsymbol{\theta}_\alpha^*(\cdot)$ is the empirical (row-wise) α-quantile of the obtained bootstrap distribution inside the brackets.

We refer to the Supplement for the formal representation of the resampling scheme and expressions of the estimators for asymptotic analysis. It also straightforwardly follows that a bootstrap sample $\mathcal{B}_b^*$ generated according to Algorithm 1 adheres to:

$$\mathbf{y}_i^* = \mathbf{y}_{i^*} = \mathbf{X}_{i^*}\boldsymbol{\beta} + \mathbf{F}_{\mathbf{y}}\gamma_{i^*} + \boldsymbol{\varepsilon}_{i^*} \tag{2.10}$$

$$\mathbf{X}_i^* = \mathbf{X}_{i^*} = \mathbf{F}_{\mathbf{x}}\boldsymbol{\Gamma}_{i^*} + \mathbf{V}_{i^*} \tag{2.11}$$

such that the unobserved factors $\mathbf{F}_{\mathbf{x}}$ and $\mathbf{F}_{\mathbf{y}}$ are indeed copied in the bootstrap realm, regardless of their number or the data generating process. The factor loadings and innovation matrices are similarly copied in their entirety, but implicitly permuted across units under the assumption that these matrices are cross-sectionally independent. This retains the within-unit correlations and variances of loadings and innovations, as well as their time series properties, which is crucial to capture the asymptotic distribution. It is easy to show that the estimator of the factor space in the bootstrap realm corresponds to:

$$\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^* = \frac{1}{N}\sum_{i=1}^N \mathbf{X}_i^* \mathbf{q}_{\dot{\mathbf{x}}} = \overline{\mathbf{X}}^* \mathbf{q}_{\dot{\mathbf{x}}} = (\mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}}_w + \overline{\mathbf{V}}_w)\mathbf{q}_{\dot{\mathbf{x}}}, \tag{2.12}$$

where $\overline{\boldsymbol{\Gamma}}_w = \frac{1}{N}\sum_{i=1}^N s_i \boldsymbol{\Gamma}_i$ and $\overline{\mathbf{V}}_w = \frac{1}{N}\sum_{i=1}^N s_i \mathbf{V}_i$ are unobserved bootstrap quantities, reweighted by the sampling frequencies s_i, where s_i denotes the sampling frequency of unit i in the bootstrap dataset $\mathcal{B}_b^*$, and s_i follows a multinomial distribution. The properties of s_i imply that $\overline{\mathbf{V}}_w \rightarrow_{p^*} \mathbf{0}_{T\times k}$ and $\overline{\mathbf{\Gamma}}_w \mathbf{q}_{\dot{\mathbf{x}}} \rightarrow_{p^*} \mathbf{\Gamma}\mathbf{q}_{\dot{\mathbf{x}}}$ as $N \rightarrow \infty$, and in turn $(\overline{\mathbf{\Gamma}}_w \mathbf{q}_{\dot{\mathbf{x}}})^+ \rightarrow_{p^*} (\mathbf{\Gamma}\mathbf{q}_{\dot{\mathbf{x}}})^+$. This confirms that the asymptotic information content in the cross-section averages, as determined by $(\mathbf{\Gamma}\mathbf{q}_{\dot{\mathbf{x}}})^+$, is also replicated in the bootstrap sample.

3 Asymptotic Results

In this section we will discuss the asymptotic distribution of both CCEP and CCEMG estimators in the original and bootstrap samples, based on Algorithm 1. We consider both heterogeneous and homogeneous slopes and demonstrate that as long as the condition $m_x = g$ can be met, asymptotically standard normal inference can ensue, though in some cases aid by the bootstrap is necessary. To begin with, we assume that $\mathbb{I}_{\boldsymbol{\nu}\neq \mathbf{0}} = \mathbf{0}_{k\times k}$. This case, among other results, was discussed in De Vos and Stauskas (2024). We re-state the key results in order to identify the challenges of the distinct CE case, and subsequently extend it to heterogeneous panels and discuss the possibility of non-stationary factors.

3.1 Homogeneous Slopes

Consider first the asymptotic distribution of the CCEP estimator when slopes are homogeneous:

Theorem 1. *Under Assumptions 1 - 5, we have as $(N, T) \rightarrow \infty$ such that $TN^{-1} \rightarrow \tau < \infty$ the following asymptotic representations:*

(a) If $m_x < g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \rightarrow_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \mathbf{\Sigma}^{-1}(\mathbf{\Psi} + \mathbf{\Psi}_f)\mathbf{\Sigma}^{-1}\right) + \mathbf{\Sigma}^{-1}(\sqrt{\tau}\mathbf{h}_1 + \mathbf{h}_2)$$

with $\mathbf{\Gamma}_{\ddot{\mathbf{x}}} = \mathbf{\Gamma}\mathbf{q}_{\ddot{\mathbf{x}}}$, $\mathbf{\Psi} = \lim_{N,T\rightarrow\infty} \frac{1}{N}\sum_{i=1}^N \mathbb{E}\left(T^{-1}\mathbf{V}_i'\varepsilon_i\varepsilon_i'\mathbf{V}_i\right)$, $\mathbf{h}_1 = \mathbf{h}_{1,1} + \mathbf{h}_{1,2} - \mathbf{h}_{1,3}$, *where*

$$\begin{aligned}
\mathbf{h}_{1,1} &= \mathbf{\Sigma}_{\gamma\Gamma}' vec\left((\mathbf{\Gamma}_{\ddot{\mathbf{x}}}^+)'\mathbf{q}_{\ddot{\mathbf{x}}}'\mathbf{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\mathbf{H}_{\dot{\mathbf{x}},m_x}\mathbf{\Sigma}_{\mathbf{F_x}}\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x},y}}\right),\\
\mathbf{h}_{1,2} &= \widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}\mathbf{\Gamma}'(\mathbf{\Gamma}_{\ddot{\mathbf{x}}}^+)'\mathbf{q}_{\ddot{\mathbf{x}}}'\mathbf{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\mathbf{H}_{\dot{\mathbf{x}},m_x}\mathbf{\Sigma}_{\mathbf{F_x}}^+\mathbf{\Sigma}_{\mathbf{F}_{x,y}}\gamma,\\
\mathbf{h}_{1,3} &= \widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}\mathbf{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\mathbf{H}_{\dot{\mathbf{x}},m_x}\mathbf{\Sigma}_{\mathbf{F_x}}^+\mathbf{\Sigma}_{\mathbf{F}_{x,y}}\gamma,
\end{aligned}$$

and $\mathbf{T}_{\dot{\mathbf{x}}}$ is a $g\times g$ partitioning matrix such that $\mathbf{\Gamma}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}} = [\mathbf{\Gamma}_{\ddot{\mathbf{x}},m_x}, \mathbf{\Gamma}_{\ddot{\mathbf{x}},-m_x}]$, where $\mathbf{\Gamma}_{\ddot{\mathbf{x}},m_x}$ is an $m_x\times m_x$ full rank matrix, $\mathbf{\Gamma}_{\ddot{\mathbf{x}},-m_x}$ is $m_x\times(g-m_x)$, and $\mathbf{H}_{\dot{\mathbf{x}},m_x} = [\mathbf{\Gamma}_{\ddot{\mathbf{x}},m_x}^{-1}, \mathbf{0}_{m_x\times(g-m_x)}]'$. Moreover, $\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}} = \mathrm{diag}\left([\mathbf{1}_{(\overline{\mathbf{X}}_1\notin\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})}, \mathbf{1}_{(\overline{\mathbf{X}}_2\notin\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})}, \ldots, \mathbf{1}_{(\overline{\mathbf{X}}_k\notin\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})}]\right)$. The definition of $\mathbf{\Psi}_f$ and $\mathbf{h}_2$ are provided in the Supplement.

(b) If $m_x = g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Psi} + \widetilde{\boldsymbol{\Psi}}_f)\boldsymbol{\Sigma}^{-1}\right) + \sqrt{\tau}\boldsymbol{\Sigma}^{-1}\widetilde{\mathbf{h}}_1,$$

where $\widetilde{\mathbf{h}}_1 = \widetilde{\mathbf{h}}_{1,1} + \widetilde{\mathbf{h}}_{1,2} - \widetilde{\mathbf{h}}_{1,3}$, *where*

$$\widetilde{\mathbf{h}}_{1,1} = \boldsymbol{\Sigma}'_{\gamma\Gamma} vec\left((\boldsymbol{\Gamma}^{+}_{\ddot{\mathbf{x}}})'\mathbf{q}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}(\boldsymbol{\Gamma}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F_x}}\boldsymbol{\Gamma}_{\ddot{\mathbf{x}}})^{+}\boldsymbol{\Gamma}_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},y}}\right),$$
$$\widetilde{\mathbf{h}}_{1,2} = \widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}'(\boldsymbol{\Gamma}^{+}_{\ddot{\mathbf{x}}})'\mathbf{q}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}(\boldsymbol{\Gamma}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F_x}}\boldsymbol{\Gamma}_{\ddot{\mathbf{x}}})^{+}\boldsymbol{\Gamma}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{x,y}}\gamma,$$
$$\widetilde{\mathbf{h}}_{1,3} = \widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}(\boldsymbol{\Gamma}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F_x}}\boldsymbol{\Gamma}_{\ddot{\mathbf{x}}})^{+}\boldsymbol{\Gamma}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{x,y}}\gamma.$$

The definition of $\widetilde{\boldsymbol{\Psi}}_f$ is provided in the Supplement.

Theorem 1 (a) and (b) confirm our prediction that the presence of the unaccounted $\mathbf{F_y}$ affects both the mean and the variance of the asymptotic distribution of the CCEP estimator as $TN^{-1} \to \tau < \infty$. In particular, the asymptotic variance is affected by $\mathbf{F_y}$ irrespective of the relative expansion rate of N and T. Asymptotic bias similarly is a function of the remaining factors due to the presence of $\boldsymbol{\Sigma}_{\mathbf{F}_{x,y}}$, the covariance between the **y**- and **x**-specific factors. We similarly find that the asymptotic distribution also depends on the difference between g (the number of CA used) and m_x (the number of **x**-specific factors). Importantly, in part (a), we have $g > m_x$, in which case the distribution features $\mathbf{h}_2$, a stochastic term which does not converge to the normal distribution, thereby making the overall distribution non-standard and invalidating standard normal inference. The guilty term is mainly driven by the interaction of two components: the error part of the $g - m_x$ redundant CAs, and the covariance between $\mathbf{F_x}$ and $\mathbf{F_y}$. Unlike the other deterministic bias components, $\mathbf{h}_2$ cannot be eliminated even if $TN^{-1} \to 0$. In addition, the asymptotic variance estimator in (2.7) is also inconsistent due to the presence of $\mathbf{F_y}$ in the model residuals. As shown by De Vos and Stauskas (2024) (see Proposition 3), the analytical variance estimators are only consistent in the common factor case $\mathbf{F_x} = \mathbf{F_y} = \mathbf{F}$, otherwise $\boldsymbol{\Psi}_f$ is not captured. Part (b), on the other hand, shows that if we have exactly $m_x = g$, then the distribution does not contain terms that impede asymptotic normality as such. Nevertheless, the bias $\widetilde{\mathbf{h}}_1$ still depends on $\boldsymbol{\Sigma}_{\mathbf{F}_{x,y}}$. This means that the bias cannot be estimated and corrected as in Westerlund and Urbain (2013), even under $m_x = g$, because $\mathbf{F_y}$ is neither observed nor estimable. In addition, we have similarly to Part (a) that the variance estimator in (2.7) is inconsistent.

The key conclusion from the above theory is thus that standard asymptotic inference with CCEP cannot be trusted when factors are distinct, even if $TN^{-1} \to 0$. Analytical variance estimators are inconsistent (regardless of g and m), asymptotic bias is not estimable with standard approaches, and the asymptotic distribution features a non-normal component when $g > m_x$. The cross-section bootstrap, however, does enable valid inference in this setting, as we show next in Theorem 2.

Theorem 2. *Under Assumptions 1 - 5 we have as $(N, T) \to \infty$ such that $TN^{-1} \to \tau < \infty$ the following asymptotic representations:*

(a) If $m_x < g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \to_{d^*} \mathcal{N}(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Psi} + \boldsymbol{\Psi}_f)\boldsymbol{\Sigma}^{-1}) + \boldsymbol{\Sigma}^{-1}(\sqrt{\tau}\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}^+)$$

where $\mathbf{h}^+ = 2(\mathbf{h}_2^* - \mathbf{h}_2)$ with the definition of $\mathbf{h}_2^*$ provided in the Supplement. The remaining quantities are as defined in Theorem 1.

(b) If $m_x = g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \to_{d^*} \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Psi} + \widetilde{\boldsymbol{\Psi}}_f)\boldsymbol{\Sigma}^{-1}\right) + \sqrt{\tau}\boldsymbol{\Sigma}^{-1}\widetilde{\mathbf{h}}_1,$$

with definitions as in Theorem 1 (b), and we have under the same conditions:

$$\sup_{x\in\mathbb{R}^{k\times 1}} \left| \mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \le x] - \mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \le x] \right| \to_p 0,$$

where the inequalities should be interpreted coordinate-wise.

It is evident for the $m_x < g$ case in part (a) that while the asymptotic variance is replicated in the bootstrap realm, the bias is not due to the presence of an extra noise term $\mathbf{h}_2^+$. The latter represents a distortion of the stochastic term $\mathbf{h}_2$ in the bootstrap realm caused by the moments of resampling weights s_i. The bootstrap is thus not consistent when g exceeds m_x. However, if $m_x = g$, the original sample and bootstrap distributions coincide due to the fact that there are no excess CAs. The bootstrap is thus consistent in this case. In practice, this implies that bootstrap-aided inference in the distinct factor case requires verification of $g = m_x$. In order to asymptotically guarantee this condition, we follow De Vos and Stauskas (2024) and employ the following Information Criterion (IC) adapted from Margaritella and Westerlund (2023):

$$IC(M_{\dot{\mathbf{x}}}) = \log\left(\det(\overline{\mathbf{Q}}_{\dot{\mathbf{x}}})\right) + g \cdot k \cdot p_{N,T}, \tag{3.1}$$

where $M_{\dot{\mathbf{x}}}$ is a combination of column indices of $\overline{\mathbf{X}}$, and $\mathbf{q}_{\dot{\mathbf{x}}}$ picks the corresponding $g = \text{cols}(\mathbf{q}_{\dot{\mathbf{x}}})$ averages in practice as before. Let accordingly $M_{\dot{\mathbf{x}},0}$ denote the set of averages from $\overline{\mathbf{X}}$ such that $\text{rk}(\boldsymbol{\Gamma}\mathbf{q}_{\dot{\mathbf{x}}}) = m_x$, $\text{cols}(\mathbf{q}_{\dot{\mathbf{x}}}) = m_x$, and $p_{N,T}$ is a penalty term in function of the panel dimensions N, T, such that $p_{N,T} \to 0$. This leads to the following selector for the CAs such that $m_x = g$, which should be implemented in Step 1 of Algorithm 1:

$$\widehat{M}_{\dot{\mathbf{x}}} = \underset{M_{\dot{\mathbf{x}}} \subseteq \overline{M}_{\dot{\mathbf{x}}}}{\arg\min}\, IC(M_{\dot{\mathbf{x}}}), \tag{3.2}$$

where $\overline{M}_{\dot{\mathbf{x}}}$ denotes the index set of all possible combinations of CAs. Provided that $(N,T) \to \infty$ such that $p_{N,T}C_{N,T}^2 \to \infty$ where $C_{N,T} = \min\{\sqrt{N}, \sqrt{T}\}$, we have that

$$\mathbb{P}(\widehat{M}_{\dot{\mathbf{x}}} = M_{\dot{\mathbf{x}},0}) \to 1 \quad \textit{and} \quad \mathbb{P}(g = m_x) \to 1.$$

This condition on the penalty is satisfied by several suggestions made by Bai and Ng (2002), among others. For instance, $p_{N,T} = \frac{N+T}{NT}\log(C_{N,T}^2)$ showcases the best small sample performance provided that T is sufficiently large, which is a suitable option as we consider $TN^{-1} \to \tau < \infty$. Importantly, $M_{\dot{\mathbf{x}},0}$ does not have to be unique as the selected set of CAs will be the one with the most informative loadings $\overline{\mathbf{\Gamma}}\mathbf{q}_{\dot{\mathbf{x}}}$ (see the characterisation of such a set in Proposition 3 of De Vos and Stauskas, 2024).[2] Note that the rank condition in Assumption 4, which ensures that the selection exercise is feasible in the first place, can be checked with the methodology of De Vos et al. (2024). In summary, the consistency of (3.1) asymptotically guarantees that the conditions in part (b) of Theorem 2 are met, so that the asymptotic bias and the variance can be estimated consistently by the means of the CS bootstrap, and asymptotically unbiased inference can ensue with (2.9).

3.2 Heterogeneous Slopes

We now consider the case of heterogeneous slopes by letting $\mathbb{I}_{\boldsymbol{\nu}\neq\mathbf{0}} = \mathbf{I}_k$ and explore both the CCEP and CCEMG estimator.

Theorem 3. *Under Assumptions 1 - 7, as* $(N,T) \to \infty$

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \mathbf{\Sigma}^{-1}\mathbf{\Psi}_{\nu}\mathbf{\Sigma}^{-1}\right),$$

where $\mathbf{\Sigma} = \mathrm{plim}_{N\to\infty}\frac{1}{NT}\sum_{i=1}^N \mathbf{V}_i'\mathbf{V}_i$ *and* $\mathbf{\Psi}_{\nu} = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N \mathbf{\Sigma}_i\mathbf{\Omega}_{\nu}\mathbf{\Sigma}_i$.

Theorem 3 reveals that the CCEP estimator remains $\sqrt{N}$-consistent, unbiased, and asymptotically normal in the distinct factor case under heterogeneity, irrespective of the relative expansion rate of N and T. The theorem also puts forward two striking and somewhat counter-intuitive results, which are major deviations from the homogeneous setup. The first is that the CCEP estimator is asymptotically normal and unbiased irrespective of whether $m_x < g$ or $m_x = g$. Moreover, $\mathbf{F}_{\mathbf{y}}$ no longer affects the asymptotic variance. This result coincides with the findings for $\mathbf{F}_{\mathbf{x}} = \mathbf{F}_{\mathbf{y}} = \mathbf{F}$ of Stauskas (2022) (with non-stationary factors) and the heterogeneous slopes analysis (with stationary factors) of De Vos and Stauskas (2024). To the best of our knowledge, Theorem 3 is the first to highlight robustness of the CCEP estimator to distinct factors in heterogeneous panels. The intuition behind this result is as follows. First,

[2] In the original paper of Margaritella and Westerlund (2023), that set minimizes the mean squared error $\widehat{\sigma}_{\dot{\mathbf{x}}}^2 = \frac{1}{NT}\sum_{i=1}^N \widehat{\boldsymbol{v}}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\widehat{\boldsymbol{v}}_i$, with $\widehat{\boldsymbol{v}}_i = \mathbf{y}_i - \mathbf{X}_i\widehat{\boldsymbol{\beta}}_{\mathbf{z}}$.

the slope heterogeneity $\boldsymbol{\nu}_i$ dominates the asymptotic distribution through

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i\boldsymbol{\nu}_i + o_p(1), \tag{3.3}$$

which obeys the standard Central Limit Theorem (CLT), while the terms driven by $\mathbf{F_y}$ and the idiosyncratic error $\boldsymbol{\varepsilon}_i$ in (2.5) are of a lower order. The $\mathbf{F_y}$ term in particular is also of lower order because $\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i$ is asymptotically uncorrelated with $\mathbf{F_y}$, since $\mathbf{F_x}$ is projected out. Therefore,

$$\left\|\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F_y}\gamma_i\right\| = o_p(1). \tag{3.4}$$

In effect, the influence of $\mathbf{F_y}$ is asymptotically negligible so long as $\mathbf{F_x}$ can be consistently estimated, as implied by Assumption 4. We next turn to the CCEMG estimator.

Theorem 4. *Under Assumptions 1 - 7, as* $(N,T)\to\infty$ *such that* $TN^{-1}\to\tau>0$

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}-\boldsymbol{\beta}) \to_d \mathcal{N}\left(\mathbf{0}_{k\times 1},\boldsymbol{\Omega}_{\boldsymbol{\nu}}\right),$$

where $\boldsymbol{\Omega}_{\boldsymbol{\nu}} = \mathbb{E}(\boldsymbol{\nu}_i\boldsymbol{\nu}_i')$.

Similarly to Theorem 3, the main takeaway is that the CCEMG estimator is asymptotically normal and unbiased, with its variance unaffected by the presence of $\mathbf{F_y}$. This result also holds irrespective of $m_x < g$ or $m_x = g$.[3] The rationale behind this outcome is the same as behind Theorem 3, meaning that the slope heterogeneity is the slowest decaying term:

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}-\boldsymbol{\beta}) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\nu}_i + o_p(1). \tag{3.5}$$

This result is new in the CCE literature, and it also extends Theorem 4.1 in Cui et al. (2022) in the PC context, because in the latter study (3.5) holds only when m_x is known. Particularly, their two-stage procedure can now be reduced to the first stage estimation of $\mathbf{F_x}$ only, where the dominance of $\boldsymbol{\nu}_i$ relegates the effect $\mathbf{F_y}$ to the idiosyncratic components.[4] This is also the main message of our Theorem 4 for CCE

[3]Note that the requirement of $TN^{-1}\to\tau<\infty$ is only a sufficient condition to asymptotically eliminate the accumulated errors. While it is suitable under our N,T configurations, it may not be necessary as in Theorem 3.

[4]Note that according to (2.6) and Theorem 4, under homogeneous $\boldsymbol{\beta}$, we have $\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}-\boldsymbol{\beta}) = o_p(1)$. This means that we can always consistently estimate the homogeneous $\boldsymbol{\beta}$ by CCEMG , but inference should be based on $\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}-\boldsymbol{\beta})$, as suggested by Theorem 1 and 2. We skip such analysis in the interest of space.

estimation.

Theorems 3 and 4 suggest that the variance estimators in (2.7) - (2.8) should be consistent, unlike in the homogeneous case. This is confirmed by Theorem 5.

Theorem 5. *Under Assumptions 1 - 7, as* $(N, T) \to \infty$

$$(a)\ N\widehat{\mathbf{\Theta}}_{CCEP,\dot{\mathbf{x}}} \to_p \mathbf{\Sigma}^{-1}\mathbf{\Psi}_{\nu}\mathbf{\Sigma}^{-1},$$

$$(b)\ N\widehat{\mathbf{\Theta}}_{CCEMG,\dot{\mathbf{x}}} \to_p \mathbf{\Omega}_{\nu}.$$

Clearly, bootstrap inference is not required with the CCE approach if it is known a priori that slopes are heterogeneous. Additionally, we do not need to take into consideration whether $m_x = g$ or $m_x < g$, which is a major convenience. However, as it is often unknown whether factors are distinct or slopes are heterogeneous, the most suitable approach would be one robust to each setting, and which does not require discrimination between the two cases. Indeed, as we can rely on the CS bootstrap so long as $m_x = g$ is guaranteed in the homogeneous slopes case, it is natural to attempt the same in heterogeneous panels. It is especially innocuous, since the asymptotic properties of CCEP and CCEMG are invariant to whether $m_x = g$ or $m_x < g$, according to Theorems 3 and 4. In Theorem 6 below, we thus establish bootstrap consistency for both estimators when slopes are heterogeneous.

Theorem 6. *Under Assumptions 1 - 7, as* $(N, T) \to \infty$ *such that* $TN^{-1} \to \tau > 0$,

$$(a)\ \sup_{x \in \mathbb{R}^{k\times 1}} \left| \mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \leq x] - \mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \leq x] \right| \to_p 0,$$

$$(b)\ \sup_{x \in \mathbb{R}^{k\times 1}} \left| \mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}) \leq x] - \mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \leq x] \right| \to_p 0,$$

where inequalities are to be interpreted coordinate wise.

The main practical implication of Theorems 2 and 6 is that with bootstrap inference, researchers do not need to differentiate between homogeneous and heterogeneous panels, nor whether $\mathbf{y}_i$ and $\mathbf{X}_i$ are driven by common or distinct factors. That is, the same confidence intervals and bias-adjustments apply in either setting under Assumption 4. Even if the bootstrap is not strictly necessary in heterogeneous panels, Theorem 6 shows that its application is innocuous. Theorem 7 in the supplementary material also provides the bootstrap world equivalent of Theorem 5, for completeness, and thereby establishes also the validity of the bootstrap-t interval in this setting.

Remark 1. *Note that the cross-section independence of* $\mathbf{V}_i$ *and* $\boldsymbol{\varepsilon}_i$ *is not required if it is known that slopes are heterogeneous. This is because the asymptotic distribution does not feature these variance components, so their dependence structure does not need to be* *replicated in the bootstrap realm. We can therefore relax this assumption along the lines of Pesaran and Tosetti (2011) by requiring instead* $\mathbf{U}_t = (\mathbf{M}_N \otimes \mathbf{I}_{k+1})\boldsymbol{\xi}_t$*, where* $\mathbf{U}_t \in \mathbb{R}^{N(k+1)\times 1}$ *is a cross-section stack of* $\mathbf{u}_{i,t}$ *and* $\boldsymbol{\xi}_t$ *obeys the time-dependence requirements of Assumption 1. Here,* $\mathbf{M}_N$ *is an* $N \times N$ *"network matrix" with bounded row and column norms.*

3.3 Distinct Correlated Effects with General Unknown Processes

It is known that CCE estimators are able to accommodate a wide variety of data generating processes of the factors without sacrificing asymptotic normal inference or the rate of consistency. This includes factors that are (mixtures of) integrated processes or deterministic trends, as demonstrated by Westerlund (2018) for the homogeneous case, or Stauskas (2022) in the heterogeneous setting. Both studies, however, examined properties under the common factor assumption. It is therefore natural to wonder whether this generality of CCE translates to the distinct factor setting. Let accordingly $\mathbf{F}$ be such that $\mathbf{D}_T^{-1}\mathbf{F}'\mathbf{F}\mathbf{D}_T^{-1} \Rightarrow \boldsymbol{\Sigma}_\mathbf{F}$ is asymptotically full-rank, and $\mathbf{D}_{T,F,k+1}^{-1} vec(\mathbf{U}_i'\mathbf{F})$ converges weakly as $T \to \infty$, where $\mathbf{D}_{T,F,k+1} = (\mathbf{D}_{T,F} \otimes \mathbf{I}_{k+1})$ for some normalization matrix $\mathbf{D}_T = \text{diag}(\mathbf{D}_{T,\mathbf{y}}, \mathbf{D}_{T,\mathbf{x}})$, such that $\mathbf{D}_{T,\mathbf{a}} = \text{diag}(T^{p_{\mathbf{a},1}}, \ldots, T^{p_{\mathbf{a},m_\mathbf{a}}})$, $\mathbf{a} \in \{\mathbf{x}, \mathbf{y}\}$ and $p_{\mathbf{a},j} \geq 1/2$. Here, "$\otimes$" and "$\Rightarrow$" represent Kronecker product and weak convergence, respectively. Proposition 1 below formulates conditions under which asymptotically normal inference ensues for the CCEP and CCEMG estimators with general unknown factors.

Proposition 1. *Under Assumptions 1 - 7 for* $m_x < g$ *as* $(N,T) \to \infty$ *with* $TN^{-1} \to \tau > 0$*, plus a covariance stationary* $\mathbf{F}_\mathbf{y}$ *with absolute summable autocovariances, we have the following asymptotic representations:*

(a) *(heterogeneous* $\boldsymbol{\beta}$*)* $\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) = \boldsymbol{\Sigma}^{-1}\dfrac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i\boldsymbol{\nu}_i + o_p(1),$

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\nu}_i + o_p(1).$$

If, in addition, $\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},\mathbf{y}}}$ *is deterministic, then*

(b) *(homogeneous* $\boldsymbol{\beta}$*)* $\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta})$

$$= \boldsymbol{\Sigma}^{-1}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[\frac{\mathbf{V}_i'\boldsymbol{\varepsilon}_i}{\sqrt{T}} + \boldsymbol{\Theta}_i\mathbf{D}_{T,F,k}^{-1}\text{vec}(\mathbf{V}_i'\mathbf{F})\right] + \sqrt{\tau}\mathbf{h}_1(\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{xy}}}) + \mathbf{h}_2\right) + o_p(1),$$

where $\boldsymbol{\Theta}_i$ *is a random matrix in function of the factor loadings. Also,* $\mathbf{h}_1$ *and* $\mathbf{h}_2$ *are equivalents of the respective terms in Theorem 1.*

Part (a) reveals that the findings in Westerlund (2018) or Stauskas (2022) carry over to the DCE setting so long as the general **x**-factors are asymptotically projected out. Therefore, the conclusions of Theorems 3 and 4 apply. The restriction of covariance stationary $\mathbf{F}_{\mathbf{y}}$, which remains in the residuals, is then sufficient to preserve the same rate of consistency and the asymptotic normal distribution. Indeed, the terms akin to (3.4) remain asymptotically negligible (see the Supplement for details). Since pooling causes bias under homogeneous $\boldsymbol{\beta}$, part (b) requires another restriction to keep the bias terms $\mathbf{h}_1$ non-random and avoid a non-standard asymptotic distribution. This is because the expansions depend on the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},\mathbf{y}}}$, similarly to Theorem 1. The latter is deterministic if $\mathbf{F}_{\mathbf{x}}$ contains at most deterministic trends or moderately integrated processes (see e.g. Magdalinos and Phillips, 2009). For instance, if $\mathbf{f}_{\mathbf{x},t} = (1, t, t^2, \ldots, t^{m_x-1})'$ with $\mathbf{D}_{T,\mathbf{x}} = \text{diag}(T^{1/2}, T^{3/2}, \ldots, T^{m_x-1/2})$ and $\mathbf{D}_{T,\mathbf{y}} = \sqrt{T}\mathbf{I}_{m_y}$, it can be shown that

$$\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},\mathbf{y}}} = \operatorname*{plim}_{T\to\infty} \mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\mathbf{D}_{T,\mathbf{y}}^{-1} = \int_{s=0}^{1} (\mathbf{s} \times \boldsymbol{\mu}_{\mathbf{f}_{\mathbf{y}}}')ds$$

as $T \to \infty$, where $\mathbf{s} = (1, s, s^2, \ldots, s^{m_x-1})'$ and $\boldsymbol{\mu}_{\mathbf{f}_{\mathbf{y}}} = \mathbb{E}(\mathbf{f}_{\mathbf{y},t})$ (see the supplementary material). Clearly, this restriction is not needed if $TN^{-1} \to 0$. Under these conditions, the first component in the brackets is asymptotically normal. Similarly to Theorem 1, we have that $\mathbf{h}_2$ is not a normal variate, but that it is absent when $m_x = g$.

If $\mathbf{F}_{\mathbf{y}}$ is not restricted to be covariance stationary, however, further restrictions are needed. Since it remains unobserved, $\mathbf{F}_{\mathbf{y}}$ will generally dominate the asymptotic distribution and alter the rate of consistency, implying that we need to analyze $\sqrt{N}\mathbf{D}_{T,\mathbf{y}}(\widehat{\boldsymbol{\beta}}_{CCEP,\ddot{\mathbf{x}}} - \boldsymbol{\beta})$. While it is appealing to keep m_y unrestricted, we must now impose that $m_y \leq k$. Otherwise, some **y**-specific factors will not be stabilized and the distribution may diverge. For these reasons, we leave exploration of an unrestricted $\mathbf{F}_{\mathbf{y}}$ for future research.

Remark 2. *Even under covariance stationary $\mathbf{F}_{\mathbf{x}}$, we can allow the whole common component to be non-stationary by means of breaking loadings. Suppose that at time t^* the loadings change from $\boldsymbol{\Gamma}_{1,i}$ to $\boldsymbol{\Gamma}_{2,i}$. Let $\boldsymbol{\Gamma}_{i,t} = \mathbb{I}(t < t^*)\boldsymbol{\Gamma}_{1,i} + \mathbb{I}(t \geq t^*)\boldsymbol{\Gamma}_{2,i}$ be the resulting time-varying version of $\boldsymbol{\Gamma}_i$ with $\mathbb{I}(A)$ being the indicator function for the event A taking the value one when A is true and zero otherwise. This means that the common component of $\mathbf{x}_{i,t}$ can be written as*

$$\boldsymbol{\Gamma}_{i,t}'\mathbf{f}_{\mathbf{x},t} = \mathbb{I}(t < t^*)\boldsymbol{\Gamma}_{1,i}'\mathbf{f}_{\mathbf{x},t} + \mathbb{I}(t \geq t^*)\boldsymbol{\Gamma}_{2,i}'\mathbf{f}_{\mathbf{x},t} = \boldsymbol{\Xi}_i'\mathbf{g}_t, \tag{3.6}$$

where $\boldsymbol{\Xi}_i = [\boldsymbol{\Gamma}_{1,i}', \boldsymbol{\Gamma}_{2,i}']' \in \mathbb{R}^{2m_x \times k}$ and $\mathbf{g}_t = [\mathbb{I}(t < t^*)\mathbf{f}_{\mathbf{x},t}', \mathbb{I}(t \geq t^*)\mathbf{f}_{\mathbf{x},t}']' \in \mathbb{R}^{2m_x}$, similarly to Breitung and Eickmeier (2011). Hence, the model with breaking loadings can be written equivalently as a model without break but with $2m_x$ factors. Following Assumption 4, if $g > 2m_x$, $\overline{\boldsymbol{\Xi}}\mathbf{q}_{\ddot{\mathbf{x}}} = [\overline{\boldsymbol{\Xi}}_{\ddot{\mathbf{x}},2m_x}, \overline{\boldsymbol{\Xi}}_{\ddot{\mathbf{x}},-2m_x}]$, where $\overline{\boldsymbol{\Xi}}_{\ddot{\mathbf{x}},2m_x}$ and $\overline{\boldsymbol{\Xi}}_{\ddot{\mathbf{x}},-2m_x}$ are $2m_x \times 2m_x$ and $2m_x \times (g - 2m_x)$, respectively, whereas if $2m_x = g$, then $\overline{\boldsymbol{\Xi}}\mathbf{q}_{\ddot{\mathbf{x}}} = \overline{\boldsymbol{\Xi}}_{\ddot{\mathbf{x}},2m_x}$ and in any case $\text{rk}(\overline{\boldsymbol{\Xi}}\mathbf{q}_{\ddot{\mathbf{x}}}) = 2m_x$.

4 Monte Carlo Simulations

In this section, we verify our theoretical predictions with a simulation study. To that end, we utilize a data generating process similar to De Vos and Stauskas (2024). In particular, we let the time varying unobservables follow:

$$\begin{aligned}
\mathbf{f}_{a,t} &= \rho \mathbf{f}_{a,t-1} + \sqrt{1-\rho^2}\boldsymbol{\nu}_t^f, \qquad \boldsymbol{\nu}_t^f \sim \mathcal{N}(\mathbf{0}_{m_a\times 1}, \mathbf{I}_{m_a}/m_a), \qquad a \in \{\mathbf{x},\mathbf{y}\} \\
\varepsilon_{i,t} &= \rho \varepsilon_{i,t-1} + \sqrt{1-\rho^2}\nu_{i,t}^{\varepsilon}, \qquad \nu_{i,t}^{\varepsilon} \sim \mathcal{N}(0, \sigma_i^2) \\
\mathbf{v}_{i,t} &= \rho \mathbf{v}_{i,t-1} + \sqrt{1-\rho^2}\boldsymbol{\nu}_{i,t}^x, \qquad \boldsymbol{\nu}_{i,t}^x \sim \mathcal{N}(\mathbf{0}_{k\times 1}, \sigma_{x,i}^2 \mathbf{I}_k)
\end{aligned}$$

where each variable is initiated at 0 and the first 50 periods are discarded as a burn-in to neutralize initial conditions. We set the autocorrelation parameter to $\rho = 0.8$ for all experiments in accordance with the high serial correlation that is typically encountered in practice. We set $k = 3$ and $m_y = m_x = 2$ to let distinct $\mathbf{F_y}$ and $\mathbf{F_x}$ drive $\mathbf{y}_i$ and $\mathbf{X}_i$, respectively. With $m_x < k$ and $m_y > 1$, the design reflects the setting of interest in this paper, where the rank condition on $\overline{\mathbf{C}}$ fails and $\mathbf{F_y}$ is inestimable with the CA. Hence, only $\text{rk}(\overline{\mathbf{\Gamma}}) = m_x$ applies, so that only $\mathbf{F_x}$ is estimable. Moreover, we induce a correlation of $\rho_f = corr(\mathbf{F_y}, \mathbf{F_x}) \in (0.3, 0.7)$ between the two factor sets. We thus consider both low and high dependence in the factors. To illustrate also robustness to heteroskedasticity, variances are drawn from $\sigma_i^2 \sim \sigma^2 + (\chi_1^2 - 1)$ and $\sigma_{x,i}^2 \sim \sigma_x^2 + (\chi_1^2 - 1)$ respectively, with $\sigma_x^2 = 2$ and $\sigma^2 = 1$ for all experiments.

To generate loadings, we let $\widetilde{\mathbf{C}} = [\gamma_i, \mathbf{\Gamma}_i] = \widetilde{\mathbf{C}} + \widetilde{\boldsymbol{\eta}}_i \boldsymbol{\iota}_{1+k}'$, with $\widetilde{\boldsymbol{\eta}}_i \sim \mathcal{N}(\mathbf{0}_{m\times 1}, \sigma_\eta^2 \mathbf{I}_m)$. This implies that loadings are perfectly correlated within individuals. Because we only estimate $\mathbf{F_x}$ from the CAs, we also regulate their informativeness through the population mean $\mathbf{\Gamma}$, as controlled through the parameter $d = \det(\mathbf{\Gamma}\mathbf{\Gamma}')$. For the latter, we generate given an upper bound d^u the entries in $\mathbf{\Gamma}$ independently from $\mathcal{U}[0,2]$ such that $d^u - 0.1 \leq d \leq d^u$. The obtained $\mathbf{\Gamma}$ is then fixed over Monte Carlo replications and sample sizes. We take $d^u = 10$ as our baseline scenario with a standard information content, and study the impact of a less informative setting by lowering d^u to 5.[5] Slopes are generated as

$$\boldsymbol{\beta}_i = \beta \boldsymbol{\iota}_{k\times 1} + \boldsymbol{v}_i, \qquad \textit{with} \qquad v_{i,\ell} \sim (\chi_1^2 - 1)\sqrt{\sigma_v^2/2} \qquad \text{for} \qquad \ell = 1, \ldots, k$$

where $v_{i,\ell}$ denotes the ℓ-th row of $\boldsymbol{v}_i$, so that $\sigma_v^2 \in \{0, 1\}$ considers respectively the common and variable slopes setting. We set the slope population mean to $\beta = 1$.

In the simulations below, we denote CCE estimators as $\text{CCEP}_{\mathbf{A}}$ and $\text{CCEMG}_{\mathbf{A}}$, respectively, with the $\mathbf{A}$ subscript referring to the used specification of the CAs. We

[5]These numbers are based on the (simulated) distribution of the determinant of 2×3 matrices with elements drawn from $\mathcal{U}[0,2]$, which ranges roughly from 0 to 40 (with a long right tail) with $\mathbb{E}(d) \approx 9.2$.

include 3 specifications: 1) $\mathbf{A} = \mathbf{x}$: all CAs except for $\overline{\mathbf{y}}$, 2) $\mathbf{A} = \mathbf{x}_{inf}$: infeasible specifications with the optimal[6] sub-selection from $\overline{\mathbf{X}}$ such that $g = m_x$, 3) $\mathbf{A} = \widehat{\mathbf{x}}$: CAs selected with the IC from (3.1). Note, as such, that $m_x < g$ for $\mathbf{A} = \mathbf{x}$, $m_x = g$ for $\mathbf{A} = \mathbf{x}_{inf}$ and $\mathbf{A} = \widehat{\mathbf{x}}$ versions are estimated versions of the $\mathbf{A} = \mathbf{x}_{inf}$ specification. In the interest of space, we report the most relevant specifications for each experiment, but note that others are available upon request. All tests are performed at the norminal 5% significance level. Further, "boot$_{\mathbf{A}}$" denote bootstrap equivalents for the corresponding CCE specification, obtained from $B = 2000$ bootstrap samples generated with CS-resampling. Reported size for the bootstrap methods is from application of (2.9). As the main alternative to the CCE and bootstrap approaches, we include the 2SIV estimator recently proposed by Cui et al. (2022), where a two-stage PC method is used to arrive at an asymptotically unbiased estimator as $TN^{-1} \to \tau$, with $0 < \tau < \infty$. The approach also accommodates in its design potential distinct factors, and as such serves as a good benchmark for the CCE method. Clearly, as the 2SIV achieves the same goal as the CS-bootstrap, comparisons will be informative. We include the second stage IV estimator with the number of factors in both stages estimated using the eigenvalue ratio approach of Ahn and Horenstein (2013), as per the authors' suggestion.

4.1 Results: Homogeneous Slopes

We start with the results for homogeneous slopes. First, it is clear that standard asymptotic t-tests with CCEP cannot be trusted in case of distinct factors. In particular, Table 1 reveals the near-zero size for all asymptotic t-tests with CCEP. This occurs because the standard errors in (2.7) are inconsistent in this setting, and inference needs to be aided by means of the bootstrap. However, bootstrap inference performs well. We find that bias and size are adequate for boot$_{\mathbf{x}}$ when $m_x < g$. On the other hand, boot$_{\mathbf{x}_{inf}}$ ($m_x = g$) is slightly more accurate with an empirical size closer to the nominal one. As demonstrated in Theorem 2 (a), size distortions for boot$_{\mathbf{x}}$ are due to $m_x < g$, whereas the bootstrap was shown to be consistent if $m_x = g$, as is the case for boot$_{\mathbf{x}_{inf}}$. Results suggest, however, that the distortions for $m_x < g$ are not too large, and have a fairly minor effect on testing. The IC criterion in (3.1) can also clearly estimate the optimal set of averages for which $m_x = g$ well, at least given sufficiently large T.[7] Indeed, the boot$_{\widehat{\mathbf{x}}}$ estimator achieves practically the same bias and empirical size as its target, boot$_{\mathbf{x}_{inf}}$, when $T > 100$. This confirms the effectiveness of the combination of the IC selector and CS-bootstrap in the distinct factor case. Ultimately, we note that also the 2SIV estimator achieves a close-to-nominal size for sufficiently large T, but find that our bootstrap tests are

[6]The specified $g = m_x$ averages are optimal in the sense that $\|(\mathbf{\Gamma q}_{\bar{\mathbf{x}}})^{+}\|$ is minimized. For completeness, this optimal selection is $[\overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2]$.

[7]Selection frequencies in Table B-6 of Supplement B of De Vos and Stauskas (2024) confirm that $m_x = g$ is achieved with probability approaching 1, and shows that the same averages are selected as for the a priori unknown $\mathbf{x}_{inf}$ specification ($[\overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2]$).

generally more accurate, especially for smaller T. Comparison of the bias in Table 1 with that for the low-dependence factors (available upon request) also confirms the conclusion of Theorem 1 that asymptotic bias for CCEP is larger when correlation between $\mathbf{F_x}$ and $\mathbf{F}_y$ is stronger. As before, performance of the bootstrap is practically unaffected, whereas the 2SIV suffers some size distortions for $T < 100$.

Table 1: High dependence non-common factors

		$\sqrt{NT} \times bias$				*size*			
	$N \backslash T$	25	50	100	500	25	50	100	500
$\text{CCEP}_{\mathbf{x}}$	25	0.32	0.24	0.10	0.43	0.01	0.01	0.00	0.00
	50	0.29	0.26	0.16	0.19	0.02	0.01	0.01	0.00
	100	0.18	0.21	0.15	0.10	0.01	0.01	0.01	0.00
	500	0.09	0.02	0.03	0.15	0.01	0.01	0.00	0.01
$\text{CCEP}_{\mathbf{x}_{inf}}$	25	0.36	0.31	-0.07	0.54	0.02	0.02	0.00	0.00
	50	0.27	0.32	0.02	0.31	0.03	0.02	0.01	0.00
	100	0.16	0.22	-0.05	0.18	0.03	0.02	0.00	0.01
	500	0.02	0.05	0.05	0.20	0.02	0.03	0.00	0.01
$\text{CCEP}_{\widehat{\mathbf{x}}}$	25	0.41	0.32	-0.07	0.54	0.03	0.02	0.00	0.00
	50	0.26	0.32	0.02	0.31	0.03	0.02	0.01	0.00
	100	0.13	0.21	-0.05	0.18	0.04	0.02	0.00	0.01
	500	-0.10	0.03	0.05	0.20	0.05	0.03	0.00	0.01
$\text{boot}_{\mathbf{x}}$	25	0.13	0.02	-0.06	0.03	0.08	0.07	0.07	0.06
	50	0.13	0.08	0.03	-0.09	0.07	0.06	0.08	0.05
	100	0.05	0.07	0.08	-0.10	0.07	0.06	0.07	0.07
	500	0.04	-0.07	0.00	0.07	0.06	0.06	0.04	0.06
$\text{boot}_{\mathbf{x}_{inf}}$	25	0.18	0.04	-0.12	0.13	0.06	0.06	0.08	0.05
	50	0.11	0.09	0.03	-0.02	0.05	0.06	0.06	0.05
	100	0.04	0.03	-0.03	-0.08	0.06	0.04	0.06	0.06
	500	-0.04	-0.05	0.08	0.08	0.04	0.05	0.06	0.05
$\text{boot}_{\widehat{\mathbf{x}}}$	25	0.15	0.03	-0.12	0.13	0.07	0.06	0.08	0.05
	50	0.01	0.09	0.02	-0.02	0.06	0.06	0.06	0.05
	100	-0.11	0.02	-0.03	-0.08	0.07	0.05	0.06	0.06
	500	-0.35	-0.10	0.07	0.08	0.06	0.05	0.06	0.05
2SIV	25	0.28	0.16	0.06	0.03	0.15	0.11	0.08	0.08
	50	0.46	0.17	0.12	-0.12	0.09	0.07	0.07	0.06
	100	0.68	0.18	0.18	0.05	0.08	0.05	0.07	0.07
	500	1.67	0.36	0.06	0.03	0.21	0.06	0.05	0.05

Experiment parameters: $(d^u, \beta, \sigma^2, \sigma_\eta^2, \sigma_v^2, \rho_f) = (10, 1, 1, 1, 0, 0.7)$. This experiment features $m_y = 2$ y-specific factors $\mathbf{F}_y$ that are correlated ($\rho_f = 0.7$) with $m_x = 2$ x-specific factors $\mathbf{F_x}$. An $\widehat{\mathbf{x}}$ subscript denotes CCE specifications with CA selected with the IC criterion in (3.1), and $\mathbf{x}_{inf}$ is the infeasible CCEP specification with the optimal $g = 2$ averages from $\overline{\mathbf{X}}$ (optimal in terms of their information content on $\mathbf{F_x}$). These are $[\overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2]$. Size reported for $\text{boot}_{\mathbf{A}}$ estimators are for the bootstrap interval in (2.9) based on 2000 replications.

4.2 Results: Heterogeneous Slopes

For the heterogeneous case, we similarly begin with the CCEP estimator. Our immediate focus is on the plain $\text{CCEP}_{\mathbf{x}}$ estimator because the key message of Theorem 3 is its robustness to the distinct factors case. Table 2 corroborates this. We find that the estimator is virtually unbiased for all combinations of larger N and T, and that it displays minimal bias only if $N = 25$. However, any small sample bias is substantially smaller than in the homogeneous setting of Table 1. This observation carries over when we employ the infeasible selection of CAs ($\text{CCEP}_{\mathbf{x}_{inf}}$), where $g = m_x$. For both $\mathbf{A} \in \{\mathbf{x}, \mathbf{x}_{inf}\}$, the empirical size is similar and revolves closely around the nominal 0.05 level for all combinations of (N, T), with the exception of $N = 25$. This can be partially explained by the large N that CCE estimators require to approximate the factor space. Also, the slight distortions, especially those that occur in medium-sized samples, can be attributed to the fact that the heterogeneity $\boldsymbol{\nu}_i$ is simulated from a chi-squared distribution with $\sigma^2 = 1$, unlike in Pesaran and Tosetti (2011) or Stauskas (2022), where $\boldsymbol{\nu}_i$ is normal and $\sigma^2 = 0.02$. We also see that the bootstrap CCEP estimators behave similarly to the original sample ones both in terms of bias and size. Particularly, the infeasible $\text{boot}_{\mathbf{x}_{inf}}$ is almost identical to $\text{boot}_{\widehat{\mathbf{x}}}$, where the IC selector is employed in the first stage. The latter even performs slightly better for a small N and $T \geq 50$. Eventually, we see that both $\text{CCEP}_{\mathbf{A}}$ and $\text{boot}_{\mathbf{A}}$ for all versions of $\mathbf{A}$ perform very similarly to the 2SIV of Cui et al. (2022), which is specifically constructed to accommodate distinct factors. In fact, we note that plain CCEP showcases a better performance in terms of empirical size, especially in small and medium samples. Because 2SIV is a PC-based estimator, this can be explained by the fact that it needs not only a large N but also a large T to consistently estimate the factor space. Overall, the discussion implies that our theoretical predictions in Theorems 3, 5 and 6 are borne out well.

Table 2: High dependence non-common factors (CCEP)

	N \ T	$\sqrt{NT} \times bias$				size			
		25	50	100	500	25	50	100	500
$\text{CCEP}_{\mathbf{x}}$	25	-0.02	0.01	-0.01	-0.02	0.04	0.11	0.08	0.09
	50	0.01	0.01	0.00	0.00	0.06	0.05	0.06	0.06
	100	0.00	0.01	0.02	0.00	0.05	0.05	0.08	0.06
	500	0.01	0.00	0.00	0.00	0.06	0.06	0.05	0.06
$\text{CCEP}_{\mathbf{x}_{inf}}$	25	-0.01	0.01	-0.01	-0.02	0.08	0.11	0.07	0.09
	50	0.01	0.01	0.00	0.01	0.07	0.06	0.06	0.06
	100	0.00	0.00	0.02	0.00	0.06	0.06	0.08	0.06
	500	0.01	0.00	0.00	0.00	0.06	0.06	0.05	0.07
$\text{boot}_{\mathbf{x}_{inf}}$	25	-0.01	0.01	-0.02	-0.03	0.11	0.13	0.09	0.11
	50	0.01	0.01	0.00	0.00	0.10	0.06	0.08	0.08
	100	0.00	0.00	0.02	0.00	0.07	0.06	0.08	0.09
	500	0.01	0.00	0.00	0.00	0.06	0.06	0.05	0.06
$\text{boot}_{\widehat{\mathbf{x}}}$	25	-0.03	0.00	-0.02	-0.02	0.12	0.12	0.09	0.09
	50	0.00	0.01	0.00	0.00	0.10	0.07	0.07	0.08
	100	0.00	0.00	0.02	0.00	0.06	0.07	0.09	0.08
	500	0.01	0.00	0.00	0.00	0.07	0.06	0.05	0.06
2SIV	25	-0.03	0.00	-0.02	-0.03	0.11	0.12	0.11	0.10
	50	0.01	0.00	-0.01	0.00	0.10	0.07	0.06	0.07
	100	-0.01	0.00	0.02	0.00	0.06	0.05	0.08	0.07
	500	0.01	0.00	0.00	0.00	0.07	0.08	0.03	0.07

Experiment parameters: $(d_u, \beta, \sigma^2, \sigma_\eta^2, \sigma_v^2, \rho_f) = (5, 1, 1, 1, 1, 0.7)$. This experiment features $m_y = 2$ y-specific factors $\mathbf{F}_y$ that are correlated ($\rho_f = 0.7$) with $m_x = 2$ x-specific factors $\mathbf{F}_{\mathbf{x}}$. An $\mathbf{A} \in \{\widehat{\mathbf{x}}, \mathbf{x}_{inf}\}$ subscript denotes CCE specifications with CA selected from (3.1), and the infeasible CCEP specification with the optimal $g = 2$ averages from $\overline{\mathbf{X}}$ (optimal in terms of their information content on $\mathbf{F}_{\mathbf{x}}$), respectively. These are $[\overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2]$. Size reported for $\text{boot}_{\mathbf{A}}$ estimators are for the bootstrap interval in (2.9).

We further move on to Table 3, which contains results for the CCEMG estimator under heterogeneous slopes. The overall findings are fairly similar to the CCEP case, especially when it comes to bias. In line with our theory, we find that the plain CCEMG estimator is virtually unbiased even when $N \approx T$, and the empirical size hovers very closely to the nominal one. Again, some distortions can be attributed to the fact that a large N is needed to approximate the factor space, and $\boldsymbol{\nu}_i$ comes from a chi-squared distribution, which weakens normal approximations in finite samples. Plus, in comparison to the CCEP case, we can see smaller size distortions for $N = 25$ and $T \geq 100$ across the board for both the original and bootstrap estimator. Moreover, $\text{boot}_{\mathbf{A}}$ for both $\mathbf{A} \in \{\mathbf{x}_{inf}, \widehat{\mathbf{x}}\}$ performs slightly better than its CCEP counterpart for $(N, T) \leq 100$. Similarly to the CCEP case displayed in Table 2, all the considered estimators behave similarly to the 2SIV estimator. However, the plain CCEMG estimator no longer exhibits a clear size advantage, at least in small samples.

Table 3: High dependence non-common factors (CCEMG)

	$N \backslash T$	$\sqrt{NT} \times bias$ 25	50	100	500	*size* 25	50	100	500
$\text{CCEMG}_{\mathbf{x}}$	25	-0.03	0.01	0.00	-0.01	0.05	0.08	0.04	0.06
	50	0.02	0.00	-0.01	0.00	0.08	0.05	0.07	0.05
	100	0.00	0.00	0.01	0.00	0.05	0.04	0.08	0.07
	500	0.00	0.00	0.00	0.00	0.04	0.06	0.04	0.05
$\text{CCEMG}_{\mathbf{x}_{inf}}$	25	-0.03	0.01	0.00	-0.01	0.04	0.08	0.05	0.06
	50	0.02	0.00	-0.01	0.00	0.06	0.05	0.07	0.05
	100	-0.01	0.00	0.01	0.00	0.06	0.04	0.08	0.07
	500	0.00	0.00	0.00	0.00	0.06	0.06	0.04	0.05
$\text{boot}_{\mathbf{x}_{inf}}$	25	-0.03	0.00	0.00	-0.01	0.05	0.08	0.06	0.06
	50	0.01	0.00	-0.01	-0.01	0.06	0.06	0.07	0.05
	100	-0.01	0.00	0.01	0.00	0.06	0.04	0.09	0.07
	500	0.00	0.00	0.00	0.00	0.05	0.05	0.05	0.05
$\text{boot}_{\widehat{\mathbf{x}}}$	25	-0.03	0.00	0.00	-0.01	0.05	0.09	0.05	0.06
	50	0.02	0.00	-0.01	-0.01	0.06	0.05	0.07	0.06
	100	-0.01	0.00	0.01	0.00	0.05	0.04	0.08	0.07
	500	0.00	0.00	0.00	0.00	0.04	0.05	0.04	0.05
2SIV	25	-0.03	0.00	0.00	-0.02	0.06	0.07	0.04	0.06
	50	0.02	0.00	-0.01	-0.01	0.07	0.06	0.06	0.05
	100	-0.01	0.00	0.01	0.00	0.05	0.03	0.08	0.07
	500	0.00	0.00	0.00	0.00	0.07	0.07	0.05	0.06

Experiment parameters: $(d^u, \beta, \sigma^2, \sigma_\eta^2, \sigma_v^2, \rho_f) = (5, 1, 1, 1, 1, 0.7)$. This experiment features $m_y = 2$ y-specific factors $\mathbf{F}_y$ that are correlated ($\rho_f = 0.7$) with $m_x = 2$ x-specific factors $\mathbf{F}_\mathbf{x}$. An $\mathbf{A} \in \{\widehat{\mathbf{x}}, \mathbf{x}_{inf}\}$ subscript denotes CCE specifications with CA selected from (3.1), and the infeasible CCEMG specification with the optimal $g = 2$ averages from $\overline{\mathbf{X}}$ (optimal in terms of their information content on $\mathbf{F}_\mathbf{x}$), respectively. These are $[\overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2]$. Size reported for $\text{boot}_\mathbf{A}$ estimators are for the bootstrap interval in (2.9).

5 Application: climate shocks and economic growth

In this section, we illustrate our procedures to study the effect of climate shocks on economic growth as in Dell et al. (2012). Data are taken from the aforementioned paper, which constitutes an unbalanced panel dataset with $N = 127$ countries and an average timespan of observations of $\overline{T} = \frac{1}{N}\sum_{i=1}^N T_i = 39$ years, from 1961 to 2003. The authors regress the annual per capita economic growth rate $y_{i,t}$ on temperature and precipitation for both developed and developing nations, and find significant effects. The main model is

$$y_{i,t} = \theta_i + \alpha_{r_i,t} + \beta_1 Tmp_{i,t} + \beta_2 Prec_{i,t} + \varepsilon_{i,t}$$

where $Tmp_{i,t}$ denotes the temperature (in °C) for country i at time t, and $Prec_{i,t}$ is the precipitation level in 100mm units. θ_i are country fixed effects, and $\alpha_{r_i,t}$ represent time-and-region dummies to account for e.g. region-specific and time-varying labour productivity (technological progress). The latter can be seen as a restricted unobserved factor, with the assumption that regions have a common response/absorption speed. Following e.g. Cui et al. (2022), we generalize this assumption by replacing the set of dummies $\alpha_{r_i,t}$ with interactive effects:

$$y_{i,t} = \theta_i + \gamma_i' \mathbf{f}_{\mathbf{y},t} + \beta_1 Tmp_{i,t} + \beta_2 Prec_{i,t} + \varepsilon_{i,t} \quad (5.1)$$

where $\mathbf{f}_{\mathbf{y},t}$ now encompasses also time and region effects, but enables more flexible responses and captures potentially more growth-specific unobservables besides technological progress. These can entail e.g. business cycles, shifting trends and preferences, crises,... etc, with potentially heterogeneous responses. Naturally, the climate regressors can similarly be thought of as being affected by common factors, for instance the global temperature and precipitation climate, to which countries contribute or react depending on the characteristics of the land. Factors underlying the climate regressors, $\mathbf{f}_{\mathbf{x},t}$, are thus intuitively distinct from the factors directly affecting economic growth, but the two sets are nevertheless likely to be correlated. Indeed, historical changes in the global climate $\mathbf{f}_{\mathbf{x},t}$ (climate trends, deforestation, disasters, floods) have affected technological developments over the years. Technological progress to optimize agrucultural yield and production processes in difficult climates, for instance. Accordingly, investigating the impact of weather shocks on growth would seem to require controlling for this unobserved effect space for consistent estimates.

Given the above, we thus re-evaluate the model of Dell et al. (2012) with the CCE method to allow for interactive effects. As it can be argued that the setting is one with distinct correlated effects, we follow the theory above and focus on estimating the factor space in the regressors with $\bar{\mathbf{x}}_t = [\overline{Tmp}_t, \overline{Prec}_t]'$. Indeed, as economic growth (from theory) tends to feature more than 1 factor, $\mathbf{f}_{\mathbf{y},t}$ is in that case not estimable with a single cross-section average, and will be left in the error term of the model. As mentioned above, this requires the bootstrap for estimating standard errors and bias-correction of the pooled slope coefficients. In addition to pooled slopes, we also estimate an explicitly heterogeneous version of (5.1) with the CCEMG estimator. For both estimators, we employ the Information Criterion in (3.1) to select averages from $\bar{\mathbf{x}}_t$ such that $g = m_x$. In all of the regressions below, this resulted in the selection of a single average $\overline{Tmp}_t$, such that we use $\widehat{\mathbf{f}}_{\mathbf{x},t} = [1, \overline{Tmp}_t]'$ to estimate the factor space in all the reported regressions. Note that the one is added as an observed factor to directly capture the fixed effect θ_i.

Finally, to generalize further the approach with region-time dummies, we also more explicitly account for potential country-group specific factors by splitting the sample between developed and developing countries. The advantage in the CCE context is that the approximated factors are then also allowed to differ among the developed and developing nations, thus potentially better controlling for factors specific to each nation group. The initial downside for the pooled CCE estimator is that the N dimension is split in half, such that the relevant TN^{-1} ratio is increased from 0.31

in the full sample to roughly 0.6 in each sample, thus increasing the likelihood of distortive bias effects (cfr. Theorem 1). The proposed bootstrap toolbox, however, has been shown to remedy the implied increase in asymptotic bias. We accordingly report bootstrap-corrected slopes and confidence intervals as in (2.9) based on $B = 1999$ bootstrap samples obtained with CS resampling.

5.1 Results

Consider the results in Tables 4 and 5, which report respectively a linear and non-linear version of (5.1). An interesting general finding is that, after controlling for unobserved factors, we find no significant effects of temperature and precipitation shocks on economic growth in the developed countries in the dataset. This is the case for both the pooled and mean group estimates, and applies to both the linear and non-linear specification of the model. We thus find insufficient evidence to conclude that growth in developed economies is affected by temperature and precipitation shocks. This suggests economies that are less agriculture based. While bootstrap-corrected estimators show sizeable differences in the estimated slopes and associated confidence intervals, which are somewhat more narrow than the asymptotic ones, the conclusions align with that of the standard estimators. Conclusions are decidedly different, however, for the developing countries in the dataset, where the regressions demonstrate significant and non-linear effects on growth. In the linear model, the pooled CCE estimator finds that a 1°C rise in temperature is expected to decrease economic growth with 1.534 percentage points (p.p.), ceteris paribus, which is significant at the 1% level. The bootstrap results in a slight downward adjustment of this effect to a 1.53p.p. decrease, and similarly concludes a highly significant effect. Mean group estimates, however, are not significant. Regarding the effect of precipitation shocks, bootstrap inference with CCEP leads to different conclusions than the standard asymptotic test. Indeed, normal inference is potentially distorted by the distinct factor setting as per our theory, and CCEP finds no significant effect of precipitation on growth. The bootstrap confidence interval, however, appears more narrow than the asymptotic one, and leads to the conclusion that an additional 100mm of annual rainfall is expected to increase economic growth with 0.12p.p., ceteris paribus, which is significant at the 5% level. The mean group approach estimates this effect to be roughly 3 times larger, and both methods concur with significance at the 1% level. As expected from theory, the mean group estimator is unbiased and standard inference continues to apply under distinct factors. Accordingly, asymptotic and bootstrap results are very similar.

To explore potential non-linear effects, squares of both Temperature and Precipitation were added to the model. The effect of temperature was found to be linear, such that its square was removed in the estimation reported in Table 5. The resulting effect estimates are also similar in size to those in the linear specification. Regarding the effect of precipitation, the asymptotic and bootstrap method arrive at very different conclusions. Indeed, asymptotic inference with CCEP suggests no significant effect of precipitation, while bootstrap-corrected slopes and inference indicate that the effect of precipitation on growth may be non-linear. This also corresponds to common logic, as the effect of an additional 100mm of rainfall is likely to depend on how much rain has already fallen. Indeed, the bootstrap-corrected slope received a sizeable downward adjustment compared to the CCEP estimate, and the bootstrap confidence interval is also vastly different from the normal ones, indicating the potential distortive effects of the distinct factors. By consequence, we conclude that the bootstrap-corrected effect of precipitation is non-linear, and follows $0.414 - 2 \times 0.009 \times PRECIP$. Indeed, the marginal effect of additional rainfall turns negative when the precipitation level is already sufficiently high, while it is positive in relatively dry areas.

Table 4: Climate shocks and economic growth: linear model

		Developing			Developed		
variable	estimator	slope	LB	UB	slope	LB	UB
TEMP	CCEP	-1.534***	-2.292	-0.776	0.159	-0.413	0.731
	CCEPbt	-1.530***	-2.304	-0.708	0.138	-0.396	0.690
	CCEMG	-0.911	-2.527	0.705	0.211	-0.504	0.927
	CCEMGbt	-0.917	-2.641	0.582	0.194	-0.475	0.906
PRECIP	CCEP	0.121	-0.044	0.287	-0.057	-0.164	0.050
	CCEPbt	0.120**	0.018	0.214	-0.060	-0.152	0.016
	CCEMG	0.400***	0.123	0.676	0.036	-0.159	0.231
	CCEMGbt	0.399***	0.108	0.656	0.035	-0.166	0.222
	$(N, \overline{T}, \tau)$	65	39	0.60	62	39	0.63

Notes: *,** and *** denote respectively significance at the 10%, 5%, and 1% level. CCE estimators feature IC-selected averages ($\overline{Tmp}$ for all specifications). Bootstrap procedures (with a "bt" ending) employ 1999 replications.

Table 5: Climate shocks and economic growth: non-linear model

		Developing			Developed		
variable	estimator	slope	LB	UB	slope	LB	UB
TEMP	CCEP	-1.585***	-2.460	-0.711	0.091	-1.209	1.390
	CCEPbt	-1.634***	-2.485	-0.693	0.071	-0.502	0.638
	CCEMG	-0.734	-2.408	0.941	0.153	-0.615	0.921
	CCEMGbt	-0.789	-2.558	0.880	0.142	-0.582	0.929
PRECIP	CCEP	0.442	-0.829	1.714	-0.045	-4.458	4.367
	CCEPbt	0.414**	0.009	0.784	-0.049	-0.252	0.132
	CCEMG	3.118***	0.834	5.403	0.887	-0.467	2.241
	CCEMGbt	3.190**	0.825	5.722	0.846	-0.686	2.251
PRECIP^2	CCEP	-0.010	-0.055	0.036	0.000	-0.152	0.151
	CCEPbt	-0.009*	-0.017	0.001	0.000	-0.005	0.004
	CCEMG	-0.097	-0.398	0.204	-0.114	-0.326	0.098
	CCEMGbt	-0.096	-0.446	0.184	-0.106	-0.320	0.150
	$(N, \overline{T}, \tau)$	65	39	0.60	62	39	0.63

Notes: *,** and *** denote respectively significance at the 10%, 5%, and 1% level. CCE estimators feature IC-selected averages ($\overline{Tmp}$ for all specifications). Bootstrap procedures employ 1999 replications.

6 Conclusions

In this study we consider the practically relevant issue of CCE-based estimation when the dependent and explanatory variables are driven by distinct sets of factors, and their cross-section averages are not necessarily consistent for the space spanned by all of them. This generally distorts inference, unless in the specific case where the number of distinct factors underlying the dependent variable is equal to 1. To circumvent this problem, we develop a toolbox that can be seen as a CCE-counterpart of the Two-Stage Instrumental Variable (2SIV) approach of Cui et al. (2022). We employ a user-friendly cross-section bootstrap algorithm to approximate the asymptotic distribution that is affected by the unattended factors in the dependent variable. We derive conditions for bootstrap consistency and show that the algorithm and asymptoptic distributions remain the same in both homogeneous and heterogeneous panels, which means that asymptotically normal inference can ensue without a need to discriminate between the different cases. Our Monte Carlo simulations show that the theoretical predictions are born out well, and that our methodology performs well in comparison to alternative estimators.

References

Ahn, S. C. and Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3):1203–1227.

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

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Breitung, J. and Eickmeier, S. (2011). Testing for structural breaks in dynamic factor models. *Journal of Econometrics*, 163:71–84.

Chudik, A., Pesaran, M., and Tosetti, E. (2011). Weak and strong cross-section dependence and estimation of large panels. *The Econometrics Journal*, 14(1):C45–C90.

Cui, G., Norkutė, M., Sarafidis, V., and Yamagata, T. (2022). Two-stage instrumental variable estimation of linear panel data models with interactive effects. *The Econometrics Journal*, 25(2):340–361.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap methods and their application*. Number 1. Cambridge university press.

De Vos, I., Everaert, G., and Sarafidis, V. (2024). A method to evaluate the rank condition for CCE estimators. *Econometric Reviews*, 43(2-4):123–155.

De Vos, I. and Stauskas, O. (2024). Cross-section bootstrap for cce regressions. *Journal of Econometrics*, 240(1):105648.

De Vos, I. and Westerlund, J. (2019). On cce estimation of factor-augmented models when regressors are not linear in the factors. *Economics Letters*, 178:5–7.

Dell, M., Jones, B. F., and Olken, B. A. (2012). Temperature Shocks and Economic Growth: Evidence from the Last Half Century. *American Economic Journal: Macroeconomics*, 4(3):66–95.

Djogbenou, A., Gonçalves, S., and Perron, B. (2015). Bootstrap inference in regressions with estimated factors and serial correlation. *Journal of Time Series Analysis*, 36(3):481–502.

Eberhardt, M. and Teal, F. (2011). Econometrics for grumblers: a new look at the literature on cross-country growth empirics. *Journal of Economic Surveys*, 25(1):109–155.

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

Jiang, B., Yang, Y., Gao, J., and Hsiao, C. (2021). Recursive estimation in large panel data models: Theory and practice. *Journal of Econometrics*, 224(2):439–465.

Juodis, A. (2022). A regularization approach to common correlated effects estimation. *Journal of Applied Econometrics*, 37(4):788–810.

Kapetanios, G. (2008). A bootstrap procedure for panel data sets with many cross-sectional units. *Econometrics Journal*, 11(2):377–395.

Karabiyik, H., Reese, S., and Westerlund, J. (2017). On the role of the rank condition in CCE estimation of factor-augmented panel regressions. *Journal of Econometrics*, 197(1):60 – 64.

Karavias, Y., Narayan, P. K., and Westerlund, J. (2023). Structural breaks in interactive effects panels and the stock market reaction to covid-19. *Journal of Business & Economic Statistics*, 41(3):653–666.

Magdalinos, T. and Phillips, P. C. (2009). Limit theory for cointegrated systems with moderately integrated and moderately explosive regressors. *Econometric Theory*, 25(2):482–526.

Margaritella, L. and Westerlund, J. (2023). Using information criteria to select averages in cce. *The Econometrics Journal*, page utad009.

Moon, H. R. and Weidner, M. (2015). Linear regression for panel with unknown number of factors as interactive fixed effects. *Econometrica*, 83(4):1543–1579.

Norkutė, M. and Westerlund, J. (2021). The factor analytical approach in near unit root interactive effects panels. *Journal of Econometrics*, 221(2):569–590.

Pesaran, M. (2006). Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure. *Econometrica*, 74(4):967–1012.

Pesaran, M. and Tosetti, E. (2011). Large panels with common factors and spatial correlation. *Journal of Econometrics*, 161(2):182–202.

Stauskas, O. (2022). Complete theory for cce under heterogeneous slopes and general unknown factors. *Oxford Bulletin of Economics and Statistics*.

Westerlund, J. (2018). Cce in panels with general unknown factors. *The Econometrics Journal*, 21(3):264–276.

Westerlund, J. and Petrova, Y. (2018). Asymptotic collinearity in cce estimation of interactive effects models. *Economic Modelling*, 70:331–337.

Westerlund, J., Petrova, Y., and Norkute, M. (2019). Cce in fixed-t panels. *Journal of applied econometrics*, 34(5):746–761.

Westerlund, J. and Urbain, J. (2013). On the estimation and inference in factor-augmented panel regressions with correlated loadings. *Economics Letters*, 119(3):247–250.

Westerlund, J. and Urbain, J.-P. (2015). Cross-sectional averages versus principal components. *Journal of Econometrics*, 185(2):372–377.

Supplement to “Handling Distinct Correlated Effects with CCE”

Ovidijus Stauskas[1] and Ignace De Vos[2,3]

[1]BI Norwegian Business School, Department of Economics
[2]Vrije Universiteit Amsterdam, Department of Econometrics and Data Science
[3]Tinbergen Institute

May 17, 2024

Abstract

In this supplementary material we provide the proofs of Theorems 3 - 6 in the main text. Section 1 sets up assumptions, preliminary details and introduces to cross-section bootstrap. Section 2 states and explains the original and bootstrap sample results for homogeneous slopes derived in a separate study. In Section 3, Theorems 3 and 4 establish the asymptotic distribution of the CCEP and CCEMG estimators, respectively. Theorem 6 establishes bootstrap consistency for both CCEP and CCEMG bootstrap estimators. In Section 4, Theorem 5 demonstrates consistency of the asymptotic variance estimators, while Theorem 7 demonstrates the same for their bootstrap equivalents for completeness. The supplementary material is completed with the discussion on potentially non-stationary factors.

Contents

1 Preliminaries

1.1 Notation and Assumptions

In this supplement we use $\mathbf{A}^{+}$ to denote the Moore-Penrose pseudo-inverse of the matrix $\mathbf{A}$, $\text{rk}(\mathbf{A})$ for its rank, $\det(\mathbf{A})$ for the determinant and let $\|\mathbf{A}\| = \sqrt{\text{tr}\,(\mathbf{A}'\mathbf{A})}$ be the Euclidean (Frobenius) matrix norm. Let furthermore $\boldsymbol{\iota}_a$ be an a-rowed vector of ones and the $vec(.)$, $\otimes$ operators denote respectively the vectorization operation and the Kronecker products. Barred variables $\overline{\mathbf{A}}$ denote the cross-section average (CA) over the cross-section specific matrices $\mathbf{A}_i$ as in $\overline{\mathbf{A}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{A}_i$. For the analysis of the bootstrap, starred objects $\mathbf{A}^*$ denote *observed* variables (matrix or scalar) subject to bootstrap randomness (induced by the resampling weights). On the other hand, $\mathbf{A}_w$ denotes a weighted (by resampling weights) *unobserved* primitive of the model. On the other hand, $\mathbf{A}_w$ denotes a weighted (by resampling weights) *unobserved* primitive of the model. Bootstrap probability laws are formalized similarly to Galvao and Kato (2014). In particular, for any matrix bootstrap sequence $\mathbf{A}_n^*$, which depends on a generic index n, and a deterministic sequence $a_n \in \mathbb{R}_{++}$, we denote $\|\mathbf{A}_n^*\| = o_{p^*}(a_n)$ if for every $\epsilon > 0$ and $\delta > 0$ we have $\mathbb{P}(\mathbb{P}^*(a_n^{-1}\|\mathbf{A}_n^*\| > \epsilon) > \delta) \to 0$ as $n \to \infty$, where $\mathbb{P}^*(\cdot)$ is a bootstrap-induced measure. Accordingly, $\mathbf{A}_n^* = \mathbf{A}^* + o_{p^*}(1)$ implies $\|\mathbf{A}_n^* - \mathbf{A}^*\| = o_{p^*}(1)$ for a limiting bootstrap matrix $\mathbf{A}^*$. Similarly, we use $\|\mathbf{A}_n^*\| = O_{p^*}(a_n)$ if for every $\delta > 0$ and $\eta > 0$, there exists a constant $C > 0$, such that $\mathbb{P}(\mathbb{P}^*(a_n^{-1}\|\mathbf{A}_n^*\| > C) > \delta) < \eta$ for all $n \geq 1$. The symbols $\to_{p*}$ ($\to_p$) and $\to_{d^*}$ ($\to_d$) represent convergence in probability and distribution with respect to the induced (generic) probability measure.

We apply the following set of assumptions:

Assumption 1 *(Idiosyncratic errors) $\varepsilon_{i,t}$ and $\mathbf{v}_{i,t}$ are stationary variables, independent across i with $\mathbb{E}(\varepsilon_{i,t}) = 0$, $\mathbb{E}(\mathbf{v}_{i,t}) = \mathbf{0}_{k\times 1}$, $\sigma_i^2 = \mathbb{E}(\varepsilon_{i,t}^2)$, $\boldsymbol{\Sigma}_i = \mathbb{E}(\mathbf{v}_{i,t}\mathbf{v}_{i,t}')$, $\boldsymbol{\Omega}_i = \mathbb{E}(\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i')$, with $\boldsymbol{\Omega}_i, \boldsymbol{\Sigma}_i$ positive definite and $\mathbb{E}(\varepsilon_{i,t}^6) < \infty$, $\mathbb{E}(\|\mathbf{v}_{i,t}\|^6) < \infty$ for all i and t. Additionally, let $\tilde{\mathbf{u}}_{i,t} = (\varepsilon_{i,t}, \mathbf{v}_{i,t}')'$. Then*

$$\frac{1}{T^3}\sum_{t=1}^{T}\sum_{q=1}^{T}\sum_{r=1}^{T}\sum_{s=1}^{T}\|\mathbb{E}(\tilde{\mathbf{u}}_{i,t}\tilde{\mathbf{u}}_{i,q}'\tilde{\mathbf{u}}_{i,r}\tilde{\mathbf{u}}_{i,s}')\| = O(1), \quad \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\|\mathbb{E}(\tilde{\mathbf{u}}_{i,t}\tilde{\mathbf{u}}_{i,s}')\| = O(1)$$

as $T \to \infty$, whereas $\frac{1}{N}\sum_{i=1}^{N}\sigma_i^2 \to \sigma^2 < \infty$ and $\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i \to \boldsymbol{\Sigma} < \infty$ as $N \to \infty$.

Assumption 2 *(Distinct factors) Let $\mathbf{f}_t = (\mathbf{f}_{\mathbf{y}}', \mathbf{f}_{\mathbf{x}}')'$ be covariance stationary with $\mathbb{E}(\|\mathbf{f}_t\|^4) < \infty$, absolute summable autocovariances and $T^{-1}\mathbf{F}'\mathbf{F} \to^p \boldsymbol{\Sigma}_{\mathbf{F}}$ as $T \to \infty$, such that*

$$\boldsymbol{\Sigma}_{\mathbf{F}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{F}_y} & \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},\mathbf{y}}}' \\ \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},\mathbf{y}}} & \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}} \end{bmatrix}$$

with $\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},y}} = \text{plim}_{T\to\infty} T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}$ denoting the covariance between $\mathbf{F}_{\mathbf{x}}$ and $\mathbf{F}_{\mathbf{y}}$. Also $\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}$ and and $\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{y}}}$ are positive definite.

Assumption 3 *(Factor loadings, distinct factors) The factor loadings are given by*

$$\begin{aligned} \boldsymbol{\gamma}_i &= \boldsymbol{\gamma} + \boldsymbol{\eta}_{\gamma,i} \qquad \boldsymbol{\eta}_{\gamma,i} \sim IID(\mathbf{0}_{m_y\times 1}, \boldsymbol{\Omega}_\gamma) \\ \boldsymbol{\Gamma}_i &= \boldsymbol{\Gamma} + \boldsymbol{\eta}_{\Gamma,i} \qquad vec(\boldsymbol{\eta}_{\Gamma,i}) \sim IID(\mathbf{0}_{km_x\times 1}, \boldsymbol{\Omega}_\Gamma) \end{aligned}$$

where $\boldsymbol{\gamma}, \boldsymbol{\Gamma}$ are constant matrices, $\boldsymbol{\Sigma}_{\gamma\Gamma} = \mathbb{E}(\boldsymbol{\eta}_{\gamma,i} \otimes \boldsymbol{\eta}_{\Gamma,i})$ is a covariance matrix, $\boldsymbol{\eta}_{\gamma,i}, \boldsymbol{\eta}_{\Gamma,i}$ are independent across i and of the other model components, and $\|\boldsymbol{\gamma}\|$, $\|\boldsymbol{\Gamma}\|$, $\|\boldsymbol{\Sigma}_{\gamma\Gamma}\|$, $\|\boldsymbol{\Omega}_\gamma\|$, $\|\boldsymbol{\Omega}_\Gamma\|$ are finite.

Assumption 4 *(Rank condition) $\text{rk}(\overline{\boldsymbol{\Gamma}}\mathbf{q}_{\ddot{\mathbf{x}}}) = m$, with $\mathbf{q}_{\ddot{\mathbf{x}}}$ a $k \times g$ selector matrix.*

Assumption 5 *(Independence)* $\mathbf{f}_t, \varepsilon_{i,s}, \mathbf{v}_{j,l}, \widetilde{\boldsymbol{\eta}}_n$ *are mutually independent for all* i, j, n, t, s, l.

Assumption 6 *(Slope heterogeneity) The slopes* $\boldsymbol{\beta}_i$ *follow*

$$\boldsymbol{\beta}_i = \boldsymbol{\beta} + \boldsymbol{v}_i, \qquad \boldsymbol{v}_i \sim IID(\mathbf{0}_{k\times 1}, \boldsymbol{\Omega}_v)$$

with $\boldsymbol{\Omega}_v$ *a finite nonnegative definite* $k \times k$ *matrix and the* $\boldsymbol{v}_i$ *are independent of* $\mathbf{f}_t, \varepsilon_{i,s}, \mathbf{v}_{j,l}, \widetilde{\boldsymbol{\eta}}_n$ *for all* i, j, n, t, s, l.

Assumption 7 *(Identification)* $\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i} = T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i$, *with* $\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} = \overline{\mathbf{X}}\mathbf{q}_{\dot{\mathbf{x}}}$, *is non-singular for all* N, T, *and*

$$\mathbb{E}\left(\left\|(T^{-1}\mathbf{V}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{V}_i)^{-1}\right\|^2\right) < \infty$$

also when $\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} = \mathbf{F}_{\mathbf{x}}$.

1.2 Rotation Matrix: $m_x < g$ vs. $m_x = g$

Let $\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} = \overline{\mathbf{Z}}\mathbf{q}_{\dot{\mathbf{x}}} = \overline{\mathbf{X}}\mathbf{q}_{\dot{\mathbf{x}}}$, where $\overline{\mathbf{Z}} = [\overline{\mathbf{y}}, \overline{\mathbf{X}}]$ is the full set of available CAs and let $\mathbf{q}_{\dot{\mathbf{x}}} = [\mathbf{0}_{g\times 1}, \mathbf{q}_{\dot{\mathbf{x}}}']'$ be a $(1 + k) \times g$ selection matrix that picks g cross-section averages determined by $\mathbf{q}_{\dot{\mathbf{x}}}$ (a $k \times g$ matrix) exclusively from $\overline{\mathbf{X}}$, such that

$$\overline{\mathbf{X}}\mathbf{q}_{\dot{\mathbf{x}}} = (\mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}} + \overline{\mathbf{V}})\mathbf{q}_{\dot{\mathbf{x}}} = \mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}} + \overline{\mathbf{V}}_{\dot{\mathbf{x}}}. \tag{1.1}$$

Firstly, we consider $m_x < g$ case. To setup the key arguments in the proofs, we follow Karabiyik et al. (2017) and notice that because $\|\overline{\mathbf{V}}_{\dot{\mathbf{x}}}\| = O_p(N^{-1/2})$ for the fixed T, we have

$$\mathbb{P}\left(\text{rk}\left[T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\right] > \text{rk}\left[T^{-1}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}'\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}\right]\right) \to 1 \tag{1.2}$$

as $(N, T) \to \infty$, which means that the condition

$$\left|\text{rk}\left[T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\right] - \text{rk}\left[T^{-1}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}'\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}\right]\right| \to 0 \text{ almost surely,} \tag{1.3}$$

which ensures convergence in MP inverses (see Andrews, 1987), is violated. To take this into account, we introduce the following *rotation matrix*:

$$\overline{\mathbf{H}}_{\dot{\mathbf{x}}} = \begin{bmatrix} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}},m_x}^{-1} & -\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}},m_x}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}},-m_x} \\ \mathbf{0}_{(g-m_x)\times m_x} & \mathbf{I}_{g-m_x} \end{bmatrix} = [\overline{\mathbf{H}}_{\dot{\mathbf{x}},m_x}, \overline{\mathbf{H}}_{\dot{\mathbf{x}},-m_x}], \tag{1.4}$$

such that the average loading matrix is partitioned as $\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}} = [\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}},m_x}, \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}},-m_x}]$, where $\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}},m_x} \in \mathbb{R}^{m_x\times m_x}$ and $\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}},-m_x} \in \mathbb{R}^{m_x\times(g-m_x)}$ and $\mathbf{T}_{\dot{\mathbf{x}}}$ is the partitioning matrix. This leads to

$$\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}}} = \mathbf{F}_{\dot{\mathbf{x}}}^0 + \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}}}, \tag{1.5}$$

such that $\mathbf{F}_{\dot{\mathbf{x}}}^0 = [\mathbf{F}_{\mathbf{x}}, \mathbf{0}_{T\times(g-m_x)}]$ and $\overline{\mathbf{V}}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}}} = [\overline{\mathbf{V}}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}},m_x}, \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}},-m_x}]$. Because the upper-left block of $T^{-1}\overline{\mathbf{H}}_{\dot{\mathbf{x}}}'\mathbf{T}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}}}$ converges to $\boldsymbol{\Sigma}_{\mathbf{F}_x}$, but the lower-right block is $O_p(N^{-1})$, we still encounter a violation of (1.3). Eventually, we introduce

$$\mathbf{D}_N = \begin{bmatrix} \mathbf{I}_{m_x} & \mathbf{0}_{m_x\times(g-m_x)} \\ \mathbf{0}_{(g-m_x)\times m_x} & \sqrt{N}\mathbf{I}_{g-m_x} \end{bmatrix}. \tag{1.6}$$

Let $\mathbf{R}_{\dot{\mathbf{x}}} = \mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}}}\mathbf{D}_N$. This matrix ensures that

$$\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0 = \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\mathbf{R}_{\dot{\mathbf{x}}} = \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}}}\mathbf{D}_N = \mathbf{F}_{\dot{\mathbf{x}}}^0 + [\overline{\mathbf{V}}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}},m_x}, \sqrt{N}\overline{\mathbf{V}}_{\dot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}},-m_x}] = \mathbf{F}_{\dot{\mathbf{x}}}^0 + [\overline{\mathbf{V}}_{\dot{\mathbf{x}},m_x}^0, \overline{\mathbf{V}}_{\dot{\mathbf{x}},-m_x}^0] \tag{1.7}$$

does not have $g - m_x$ asymptotically degenerating columns since $\left\|\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}\right\| = O_p(1)$. This ensures that

$$\begin{aligned} T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} &= T^{-1}\mathbf{F}^{0\prime}_{\ddot{\mathbf{x}}}\mathbf{F}^0_{\ddot{\mathbf{x}}} + T^{-1}\mathbf{F}^{0\prime}_{\ddot{\mathbf{x}}}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}}} + T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}}}\mathbf{F}^0_{\ddot{\mathbf{x}}} + T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}}}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}}} \\ &= \mathbf{\Sigma}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}} + O_p(N^{-1/2}) + O_p(T^{-1/2}), \end{aligned} \tag{1.8}$$

where the limiting matrix is

$$\mathbf{\Sigma}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}} = \text{diag}\left[\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}, (T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})\right]. \tag{1.9}$$

This approximation holds because

$$\left\|T^{-1}\mathbf{F}^{0\prime}_{\ddot{\mathbf{x}}}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}}}\right\| = O_p(T^{-1/2}), \tag{1.10}$$

$$\left\|T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x}\overline{\mathbf{V}}^0_{m_x}\right\| = O_p(N^{-1}), \tag{1.11}$$

$$\left\|T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}\right\| = O_p(N^{-1/2}), \tag{1.12}$$

and so because $\left|\text{rk}\left[T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}\right] - \text{rk}\left[\mathbf{\Sigma}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right]\right| \to 0$ almost surely, we obtain

$$\left\|\left(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}\right)^+ - \mathbf{\Sigma}^+_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right\| = O_p(N^{-1/2}) + O_p(T^{-1/2}). \tag{1.13}$$

Because $\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} = \mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}}$ due to $\mathbf{R}_{\ddot{\mathbf{x}}} = \mathbf{T}_{\ddot{\mathbf{x}}}\overline{\mathbf{H}}_{\ddot{\mathbf{x}}}\mathbf{D}_N$ being a full rank matrix, by using the same steps as in S25 - S29 in Karabiyik et al. (2017), we then arrive at the following important expansion of projection matrices, which will play a key role in our proofs:

$$\begin{aligned} \mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} = \mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}} &= T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^+\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x} + T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}(T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{x}})^+\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \\ &+ T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}(T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{x}})^+\mathbf{F}'_{\mathbf{x}} + T^{-1}\mathbf{F}_{\mathbf{x}}(T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{x}})^+\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \\ &+ T^{-1}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}\left[(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}})^+ - \mathbf{\Sigma}^+_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right]\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}. \end{aligned} \tag{1.14}$$

However, if $m_x = g$, then (1.3) is not violated by constrution and by definition the rotation matrix becomes $\mathbf{R}_{\ddot{\mathbf{x}}} = \overline{\mathbf{\Gamma}}^{-1}_{\ddot{\mathbf{x}}}$ so that $\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} = \mathbf{M}_{\mathbf{F}_{\mathbf{x}}}$. Also, by the properties of the generalized inverse we have $\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} = \mathbf{M}_{\mathbf{F}_{\mathbf{x}}} = \mathbf{M}_{\mathbf{F}_{\mathbf{x}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}}$ and also $\mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}} = \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}$. Here, all the components are well behaved. Next, we simplify and analyze the decomposition in (1.14), given that now $m_x = g$ as

$$\begin{aligned} \mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}} = \mathbf{M}_{\mathbf{F}_{\mathbf{x}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} &= T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}(T^{-1}\widehat{\mathbf{F}}'_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^+\overline{\mathbf{V}}'_{\ddot{\mathbf{x}}} + T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}(T^{-1}\widehat{\mathbf{F}}'_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^+\overline{\mathbf{\Gamma}}'_{\ddot{\mathbf{x}}}\mathbf{F}'_{\mathbf{x}} \\ &+ T^{-1}\mathbf{F}_{\mathbf{x}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}(T^{-1}\widehat{\mathbf{F}}'_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^+\overline{\mathbf{V}}'_{\ddot{\mathbf{x}}} + T^{-1}\mathbf{F}_{\mathbf{x}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}[(T^{-1}\widehat{\mathbf{F}}'_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^+ - (\overline{\mathbf{\Gamma}}'_{\ddot{\mathbf{x}}}T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{x}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}})^+]\overline{\mathbf{\Gamma}}'_{\ddot{\mathbf{x}}}\mathbf{F}'_{\mathbf{x}}, \end{aligned} \tag{1.15}$$

where now because $\left\|T^{-1}\mathbf{F}'_{\mathbf{x}}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\right\| = O_p((NT)^{-1/2})$ and $\left\|T^{-1}\overline{\mathbf{V}}'_{\ddot{\mathbf{x}}}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\right\| = O_p(N^{-1})$ we have

$$\left\|T^{-1}\widehat{\mathbf{F}}'_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}} - \overline{\mathbf{\Gamma}}'_{\ddot{\mathbf{x}}}T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{x}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}\right\| = O_p(N^{-1}) + O_p((NT)^{-1/2}), \tag{1.16}$$

$$\left\|(T^{-1}\widehat{\mathbf{F}}'_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^+ - (\overline{\mathbf{\Gamma}}'_{\ddot{\mathbf{x}}}T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{x}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}})^+\right\| = O_p(N^{-1}) + O_p((NT)^{-1/2}). \tag{1.17}$$

1.3 Cross-Section Bootstrap

We begin this section by describing the sampling scheme as given in De Vos and Stauskas (2024) in terms of generic stack of b-rowed matrices $\mathbf{A} = (\mathbf{A}'_1, \mathbf{A}'_2, \ldots, \mathbf{A}'_N)'$. In what follows, $\to_{p*}$ and $\to_{d^*}$ represent convergence in probability and distribution with respect to the bootstrap induced probability measure, while $\mathbb{E}^*(.)$ stands for bootstrap expectation (conditionally on the sample). This is how the scheme works:

1. We model the pick of the matrix $\mathbf{A}_i$ from $\mathbf{A}$ through the $1 \times N$ selection vectors $\mathbf{w}_i = [w_{i,1}, \ldots, w_{i,N}]$, which are drawn from a multinomial distribution with 1 trial and N events with a probability of N^{-1}. Hence, each $\mathbf{w}_i$ is a unit-length vector with randomly realized 1 and zeros elsewhere. The index of the non-zero element in $\mathbf{w}_i$ denotes the unit (i^*) that is sampled from the stack $\mathbf{A}$ as unit i in the bootstrap sample.

2. The selection vectors are further collected in the $N \times N$ matrix $\mathbf{w} = [\mathbf{w}_1', \ldots, \mathbf{w}_N']'$, which outlines the allocation pattern in the bootstrap sample. In what follows,

$$\boldsymbol{\iota}_N'\mathbf{w} = \left[\sum_{i=1}^{N} w_{i,1}, \ldots, \sum_{i=1}^{N} w_{i,N}\right] = [s_1, \ldots, s_N] = \mathbf{s} \tag{1.18}$$

gives the total sampling frequency of each unit with the restriction $\sum_{i=1}^{N} s_i = N$. The random vector $\mathbf{s}$ is a multinomial vector, where the coordinate s_i for every i has expectation 1, variance of $1 - N^{-1}$, covariance between s_i and s_j of $-N^{-1}$ and a probability mass of N^{-1}.

3. We ultimately define the *cross-section bootstrap operator* $\mathbf{W}_b = (\mathbf{w} \otimes \mathbf{I}_b) \in \mathbb{R}^{bN \times bN}$ which, given a stack $\mathbf{A}$ of b-rowed matrices, produces a random draw with replacement of size N: $\mathbf{W}_b\mathbf{A} = \mathbf{A}^*$. An example with $N = 2$ and $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{b \times c}$ would be

$$\mathbf{W}_b \begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix} = \left(\begin{bmatrix}1,0\\ 1,0\end{bmatrix} \otimes \mathbf{I}_b\right) \begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix} = \begin{bmatrix}\mathbf{A}\\ \mathbf{A}\end{bmatrix} \quad \text{or} \quad \mathbf{W}_b \begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix} = \left(\begin{bmatrix}1,0\\ 0,1\end{bmatrix} \otimes \mathbf{I}_b\right) \begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix} = \begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix}.$$

The operator has the property $\mathbf{W}_b'\mathbf{W}_b = \mathbf{w}'\mathbf{w} \otimes \mathbf{I}_b = \text{diag}(\mathbf{s} \otimes \boldsymbol{\iota}_b')$, because $\mathbf{w}'\mathbf{w} = \text{diag}(\mathbf{s})$. Let also $\mathbf{A}_b = N^{-1}(\boldsymbol{\iota}_N' \otimes \mathbf{I}_b)$ be the *cross-section average operator* for stacked $b-$rowed matrices. Then, by using the Kronecker properties, the CA of the bootstrap sample is obtained by

$$\mathbf{A}_b\mathbf{A}^* = \mathbf{A}_b\mathbf{W}_b\mathbf{A} = N^{-1}(\boldsymbol{\iota}_N' \otimes \mathbf{I}_b)(\mathbf{w} \otimes \mathbf{I}_b)\mathbf{A} = N^{-1}(\mathbf{s} \otimes \mathbf{I}_b)\mathbf{A} = \frac{1}{N}\sum_{i=1}^{N} s_i\mathbf{A}_i, \tag{1.19}$$

which means that every summand is assigned a multinomial weight, such that $\mathbb{E}^*(\mathbf{A}_b\mathbf{A}^*) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{A}_i$.

We implement the steps 1 - 3 above in the CCE context. We stack the T-rowed matrices over the individuals:

$$\mathbf{X} = \underline{\mathbf{F}}_{\mathbf{x}}\,\underline{\boldsymbol{\Gamma}} + \mathbf{V} \in \mathbb{R}^{NT \times k} \tag{1.20}$$

where $\mathbf{X} = [\mathbf{X}_1', \ldots, \mathbf{X}_N']'$, $\underline{\mathbf{F}}_{\mathbf{x}} = (\mathbf{I}_N \otimes \mathbf{F}_{\mathbf{x}})$, $\underline{\boldsymbol{\Gamma}} = [\boldsymbol{\Gamma}_1', \ldots, \boldsymbol{\Gamma}_N']'$ and $\mathbf{V} = [\mathbf{V}_1', \ldots, \mathbf{V}_N']'$. Then, the draw is given by

$$\mathbf{X}^* = \mathbf{W}_T\mathbf{X} = (\mathbf{w} \otimes \mathbf{I}_T)(\mathbf{I}_N \otimes \mathbf{F}_{\mathbf{x}})\underline{\boldsymbol{\Gamma}} + \mathbf{W}_T\mathbf{V} = (\mathbf{I}_N \otimes \mathbf{F}_{\mathbf{x}})(\mathbf{w} \otimes \mathbf{I}_{m_x})\underline{\boldsymbol{\Gamma}} + \mathbf{W}_T\mathbf{V} = \underline{\mathbf{F}}_{\mathbf{x}}\mathbf{W}_{m_x}\underline{\boldsymbol{\Gamma}} + \mathbf{W}_T\mathbf{V}. \tag{1.21}$$

Simultaneously, the same is performed on $\mathbf{y} = [\mathbf{y}_1', \ldots, \mathbf{y}_N']' \in \mathbb{R}^{NT \times 1}$, such that

$$\begin{aligned}\mathbf{y}^* &= \mathbf{W}_T\mathbf{y} = \mathbf{W}_T\mathbf{X}\boldsymbol{\beta} + (\mathbf{w} \otimes \mathbf{I}_T)(\mathbf{I}_N \otimes \mathbf{F}_{\mathbf{y}})\underline{\boldsymbol{\gamma}} + \mathbf{W}_T\boldsymbol{\varepsilon} = (\mathbf{I}_N \otimes \mathbf{F}_{\mathbf{y}})(\mathbf{w} \otimes \mathbf{I}_{m_y})\underline{\boldsymbol{\gamma}} + \mathbf{W}_T\boldsymbol{\varepsilon}\\ &= \mathbf{X}^*\boldsymbol{\beta} + \underline{\mathbf{F}}_{\mathbf{y}}\mathbf{W}_{m_y}\underline{\boldsymbol{\gamma}} + \mathbf{W}_T\boldsymbol{\varepsilon}.\end{aligned} \tag{1.22}$$

By using the same Kronecker product properties as in (1.21), we can show that the cross-section average of the bootstrap sample has the following expression:

$$\widehat{\mathbf{F}}_{\mathbf{x}}^* = \overline{\mathbf{X}}^* = \mathbf{A}_T\mathbf{X}^* = \mathbf{A}_T\mathbf{W}_T\mathbf{X} = \mathbf{A}_T\mathbf{W}_T(\underline{\mathbf{F}}_{\mathbf{x}}\,\underline{\boldsymbol{\Gamma}} + \mathbf{V}) = \mathbf{F}_{\mathbf{x}}\mathbf{A}_{m_x}\mathbf{W}_{m_x}\underline{\boldsymbol{\Gamma}} + \mathbf{A}_T\mathbf{W}_T\mathbf{V} = \mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}}_w + \overline{\mathbf{V}}_w \tag{1.23}$$

where $\overline{\mathbf{\Gamma}}_w = \frac{1}{N}\sum_{i=1}^N s_i\mathbf{\Gamma}_i$ and $\overline{\mathbf{V}}_w = \frac{1}{N}\sum_{i=1}^N s_i\mathbf{V}_i$. By implementing the selection of the averages, we get

$$\widehat{\mathbf{F}}^*_{\ddot{\mathbf{x}}} = \overline{\mathbf{X}}^*\mathbf{q}_{\ddot{\mathbf{x}}} = (\mathbf{F}_{\mathbf{x}}\overline{\mathbf{\Gamma}}_w + \overline{\mathbf{V}}_w)\mathbf{q}_{\ddot{\mathbf{x}}} = \mathbf{F}_{\mathbf{x}}\overline{\mathbf{\Gamma}}_{w,\ddot{\mathbf{x}}} + \overline{\mathbf{V}}_{w,\ddot{\mathbf{x}}}. \tag{1.24}$$

This representation ensures that $\overline{\mathbf{\Gamma}}_{w,,\ddot{\mathbf{x}}} \to_{p^*} \mathbf{\Gamma}_{\ddot{\mathbf{x}}}$ as $N \to \infty$, and in turn $\overline{\mathbf{\Gamma}}^+_{w,\ddot{\mathbf{x}}} \to_{p^*} \mathbf{\Gamma}^+_{\ddot{\mathbf{x}}}$. This confirms that the asymptotic information content in the cross-section averages is replicated in the bootstrap samples. Therefore, Assumption 3 holds in the original sample and in the bootstrap environment. Recall that asymptotic singularity of $T^{-1}\widehat{\mathbf{F}}'_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}$ under $m_x < g$ is the fundamental observation in the asymptotic analysis, which requires introduction of the steps in (1.4) - (1.13). Hence, this information is also mapped to its bootstrap equivalent $T^{-1}\widehat{\mathbf{F}}^{*\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^*_{\ddot{\mathbf{x}}}$.

2 Homogeneous Slopes

2.1 Pooled Estimator: Original Sample

Theorem 1. *Under Assumptions 1 - 5 as $(N,T) \to \infty$ such that $TN^{-1} \to \tau < \infty$ the following asymptotic representations:*

(a) If $m_x < g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\ddot{\mathbf{x}}} - \boldsymbol{\beta}) \to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \mathbf{\Sigma}^{-1}(\mathbf{\Psi} + \mathbf{\Psi}_f)\mathbf{\Sigma}^{-1}\right) + \mathbf{\Sigma}^{-1}(\sqrt{\tau}\mathbf{h}_1 + \mathbf{h}_2)$$

with $\mathbf{\Psi} = \lim_{N,T\to\infty}\frac{1}{N}\sum_{i=1}^N \mathbb{E}\left(T^{-1}\mathbf{V}_i'\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{V}_i\right)$, $\mathbf{h}_1 = \mathbf{h}_{1,1} + \mathbf{h}_{1,2} - \mathbf{h}_{1,3}$, where

$$\begin{aligned}
\mathbf{h}_{1,1} &= \mathbf{\Sigma}'_{\gamma\Gamma} vec\left((\mathbf{\Gamma}^+_{\ddot{\mathbf{x}}})'\mathbf{q}'_{\ddot{\mathbf{x}}}\mathbf{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\mathbf{H}_{\dot{\mathbf{x}},m_x}\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x},y}}\right),\\
\mathbf{h}_{1,2} &= \widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}\mathbf{\Gamma}'(\mathbf{\Gamma}^+_{\ddot{\mathbf{x}}})'\mathbf{q}'_{\ddot{\mathbf{x}}}\mathbf{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\mathbf{H}_{\dot{\mathbf{x}},m_x}\mathbf{\Sigma}^+_{\mathbf{F}_{\mathbf{x}}}\mathbf{\Sigma}_{\mathbf{F}_{x,y}}\gamma,\\
\mathbf{h}_{1,3} &= \widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}\mathbf{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\mathbf{H}_{\dot{\mathbf{x}},m_x}\mathbf{\Sigma}^+_{\mathbf{F}_{\mathbf{x}}}\mathbf{\Sigma}_{\mathbf{F}_{x,y}}\gamma,
\end{aligned} \tag{2.1}$$

with $\mathbf{\Gamma}_{\ddot{\mathbf{x}}} = \mathbf{\Gamma}\mathbf{q}_{\ddot{\mathbf{x}}}$, and $\mathbf{T}_{\dot{\mathbf{x}}}$ is a $g\times g$ partitioning matrix such that $\mathbf{\Gamma}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}} = [\mathbf{\Gamma}_{\dot{\mathbf{x}},m_x}, \mathbf{\Gamma}_{\dot{\mathbf{x}},-m_x}]$, where $\mathbf{\Gamma}_{\dot{\mathbf{x}},m_x}$ is an $m_x \times m_x$ full rank matrix, $\mathbf{\Gamma}_{\dot{\mathbf{x}},-m_x}$ is $m_x \times (g - m_x)$, and $\mathbf{H}_{\dot{\mathbf{x}},m_x} = [\mathbf{\Gamma}^{-1}_{\dot{\mathbf{x}},m_x}, \mathbf{0}_{m_x\times(g-m_x)}]'$. Lastly,

$$\begin{aligned}
\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}} &= \text{diag}\left([\mathbf{1}_{(\overline{\mathbf{X}}_1\notin\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})}, \mathbf{1}_{(\overline{\mathbf{X}}_2\notin\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})}, \ldots, \mathbf{1}_{(\overline{\mathbf{X}}_k\notin\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})}]\right),\\
\mathbf{\Psi}_f &= \lim_{N,T\to\infty}\frac{1}{N}\sum_{i=1}^N \mathbb{E}\left[\mathbf{\Xi}_{\dot{\mathbf{x}},y,i}\left(T^{-1}vec\left(\mathbf{V}_i'\mathbf{F}\right)vec\left(\mathbf{V}_i'\mathbf{F}\right)'\right)\mathbf{\Xi}'_{\dot{\mathbf{x}},y,i}\right] \textit{ with}\\
\mathbf{h}_2 &= \mathbf{\Sigma}'_{\gamma\Gamma}\left(\mathbf{\Sigma}_{\mathbf{F}^0_{\ddot{\mathbf{x}},y}} \otimes \mathbf{D}_{\dot{\mathbf{x}},g-m_x}\mathbf{H}'_{\dot{\mathbf{x}}}\mathbf{T}'_{\dot{\mathbf{x}}}\mathbf{q}'_{\ddot{\mathbf{x}}}\mathbf{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}\mathbf{\Gamma}^+_{\ddot{\mathbf{x}}}\right)' vec\left(\sqrt{T}\left[(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\mathbf{x}}\widehat{\mathbf{F}}^0_{\mathbf{x}})^+ - \mathbf{\Sigma}^+_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right]\right) + \mathbf{h}_2(\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}),
\end{aligned}$$

where $\mathbf{h}_2(\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}})$ involves the terms depending on $(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\mathbf{x}}\widehat{\mathbf{F}}^0_{\mathbf{x}})^+ - \mathbf{\Sigma}^+_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}$, which disappear if $\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}} = \mathbf{0}_{k\times k}$. Next, for $\mathbf{F}_{\mathbf{x}} = \mathbf{F}\mathbf{p}_{\mathbf{x}}$ and $\mathbf{F}_{\mathbf{y}} = \mathbf{F}\mathbf{p}_{\mathbf{y}}$ we have

$$\begin{aligned}
\mathbf{\Xi}_{\dot{\mathbf{x}},y,i} &= \boldsymbol{\eta}'_{\gamma,i}\left(\mathbf{p}_{\mathbf{y}} - \mathbf{p}_{\mathbf{x}}\mathbf{\Sigma}^+_{\mathbf{F}_{\mathbf{x}}}\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x},y}}\right)' \otimes \mathbf{I}_k + \mathbf{\Sigma}'_{\gamma\Gamma}\left[\left(\mathbf{p}_{\mathbf{x}}\mathbf{\Sigma}^+_{\mathbf{F}_{\mathbf{x}}}\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x},y}} \otimes \mathbf{q}_{\ddot{\mathbf{x}}}\mathbf{\Gamma}^+_{\ddot{\mathbf{x}}}\right)' - \left(\mathbf{p}_{\mathbf{y}} \otimes (\mathbf{I}_k - \mathbf{D}_{\dot{\mathbf{x}},-m_x}\mathbf{\Sigma})\mathbf{q}_{\ddot{\mathbf{x}}}\mathbf{\Gamma}^+_{\ddot{\mathbf{x}}}\right)'\right]\\
&\quad + \mathbf{\Xi}_{\dot{\mathbf{x}},y,i}(\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}),\\
\mathbf{D}_{\dot{\mathbf{x}},g-m_x} &= \text{diag}(\mathbf{0}_{m_x}, \mathbf{I}_{g-m_x}),\\
\mathbf{D}_{\dot{\mathbf{x}},-m_x} &= \operatorname*{plim}_{N,T\to\infty} \mathbf{q}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}},-m_x}\left(T^{-1}\overline{\mathbf{V}}^{0\prime}_{-m_x}\overline{\mathbf{V}}^0_{-m_x}\right)^+\overline{\mathbf{H}}'_{\dot{\mathbf{x}},-m_x}\mathbf{T}'_{\dot{\mathbf{x}}}\mathbf{q}'_{\ddot{\mathbf{x}}},
\end{aligned}$$

where $\mathbf{\Xi}_{\dot{\mathbf{x}},y,i}(\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}})$ summarizes the terms that disappear if $\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}} = \mathbf{0}_{k\times k}$.

(b) If $m_x = g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Psi} + \widetilde{\boldsymbol{\Psi}}_f)\boldsymbol{\Sigma}^{-1}\right) + \sqrt{\tau}\boldsymbol{\Sigma}^{-1}\widetilde{\mathbf{h}}_1,$$

with $\boldsymbol{\Gamma}_{\ddot{\mathbf{x}}} = \boldsymbol{\Gamma}\mathbf{q}_{\ddot{\mathbf{x}}}$, $\widetilde{\mathbf{h}}_1 = \widetilde{\mathbf{h}}_{1,1} + \widetilde{\mathbf{h}}_{1,2} - \widetilde{\mathbf{h}}_{1,3}$, where

$$\begin{aligned}
\widetilde{\mathbf{h}}_{1,1} &= \boldsymbol{\Sigma}'_{\gamma\Gamma} vec\left((\boldsymbol{\Gamma}^{+}_{\ddot{\mathbf{x}}})'\mathbf{q}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}(\boldsymbol{\Gamma}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}\boldsymbol{\Gamma}_{\ddot{\mathbf{x}}})^{+}\boldsymbol{\Gamma}_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},y}}\right),\\
\widetilde{\mathbf{h}}_{1,2} &= \widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}'(\boldsymbol{\Gamma}^{+}_{\ddot{\mathbf{x}}})'\mathbf{q}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}(\boldsymbol{\Gamma}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}\boldsymbol{\Gamma}_{\ddot{\mathbf{x}}})^{+}\boldsymbol{\Gamma}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{x,y}}\gamma,\\
\widetilde{\mathbf{h}}_{1,3} &= \widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}(\boldsymbol{\Gamma}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}\boldsymbol{\Gamma}_{\ddot{\mathbf{x}}})^{+}\boldsymbol{\Gamma}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{x,y}}\gamma.
\end{aligned} \tag{2.2}$$

Also,

$$\begin{aligned}
\widetilde{\boldsymbol{\Psi}}_f &= \lim_{N,T\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left[\boldsymbol{\Theta}_{\dot{\mathbf{x}},y,i}\left(T^{-1}vec\left(\mathbf{V}'_i\mathbf{F}\right)vec\left(\mathbf{V}'_i\mathbf{F}\right)'\right)\boldsymbol{\Theta}'_{\dot{\mathbf{x}},y,i}\right],\\
\boldsymbol{\Theta}_{\dot{\mathbf{x}},y,i} &= \boldsymbol{\eta}'_{\gamma,i}\left(\mathbf{p}_{\mathbf{y}} - \mathbf{p}_{\mathbf{x}}\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},y}}\right)'\otimes\mathbf{I}_k + \boldsymbol{\Sigma}'_{\gamma\Gamma}\left[\left(\mathbf{p}_{\mathbf{x}}\boldsymbol{\Sigma}^{+}_{\mathbf{F}_{\mathbf{x}}}\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},y}} - \mathbf{p}_{\mathbf{y}}\right)\otimes\mathbf{q}_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}^{+}_{\ddot{\mathbf{x}}}\right]' + \boldsymbol{\Theta}_{\dot{\mathbf{x}},y,i}(\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}}),
\end{aligned}$$

where $\boldsymbol{\Xi}_{\dot{\mathbf{x}},y,i}(\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}})$ summarizes terms that disappear if $\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}} = \mathbf{0}_{k\times k}$.

Proof. See the proof of parts (a) and (b) of Proposition 1 in De Vos and Stauskas (2024).

2.2 Pooled Estimator: Bootstrap Distribution

Theorem 2. *Under Assumptions 1 - 5 we have as $(N,T)\to\infty$ such that $TN^{-1}\to\tau<\infty$ the following asymptotic representations:*

(a) If $m_x < g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}^{*}_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \to_{d^*} \mathcal{N}(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Psi} + \boldsymbol{\Psi}_f)\boldsymbol{\Sigma}^{-1}) + \boldsymbol{\Sigma}^{-1}(\sqrt{\tau}\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}^{+})$$

where $\mathbf{h}^{+} = 2(\mathbf{h}^{*}_2 - \mathbf{h}_2)$ and

$$\mathbf{h}^{*}_2 = \boldsymbol{\Sigma}'_{\gamma\Gamma}\left(\boldsymbol{\Sigma}_{\mathbf{F}^0_{\ddot{\mathbf{x}},y}}\otimes\mathbf{D}_{\dot{\mathbf{x}},g-m_x}\mathbf{H}'_{\ddot{\mathbf{x}}}\mathbf{T}'_{\dot{\mathbf{x}}}\mathbf{q}'_{\ddot{\mathbf{x}}}\boldsymbol{\Sigma}\mathbf{q}_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}^{+}_{\ddot{\mathbf{x}}}\right)' vec\left(\sqrt{T}\left[(T^{-1}\widehat{\mathbf{F}}^{0*\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^{*0}_{\ddot{\mathbf{x}}})^{+} - \boldsymbol{\Sigma}^{+}_{w,\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right]\right) + \mathbf{h}^{*}_2(\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}})$$

with $\boldsymbol{\Sigma}_{\mathbf{F}^0_{w,\ddot{\mathbf{x}},v}} = \operatorname{diag}\left[\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}, (T^{-1}\overline{\mathbf{V}}^{0\prime}_{w,\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^{0}_{w,\ddot{\mathbf{x}},-m_x})\right]$ .The remaining quantities are as defined in Theorem 1.

(b) If $m_x = g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}^{*}_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \to_{d^*} \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Psi} + \widetilde{\boldsymbol{\Psi}}_f)\boldsymbol{\Sigma}^{-1}\right) + \sqrt{\tau}\boldsymbol{\Sigma}^{-1}\widetilde{\mathbf{h}}_1,$$

where the quantities are the same as in Theorem 1 (b), and we have under the same conditions:

$$\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^{*}_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \le x] - \mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \le x]\right| \to_p 0,$$

where the inequalities should be interpreted coordinate-wise.

Proof. See the proof of part (a) and (b) of Proposition 2 in De Vos and Stauskas (2024).

3 Heterogeneous Slopes

3.1 Pooled Estimator

Theorem 3. *Under Assumptions 1 - 7, for either $m_x < g$ or $m_x = g$ as $(N,T) \to \infty$*

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_\nu\boldsymbol{\Sigma}^{-1}\right),$$

where $\boldsymbol{\Sigma} = \text{plim}_{N\to\infty} \frac{1}{NT}\sum_{i=1}^N \mathbf{V}_i'\mathbf{V}_i$ and $\boldsymbol{\Psi}_\nu = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^N \boldsymbol{\Sigma}_i\boldsymbol{\Omega}_\nu\boldsymbol{\Sigma}_i$.

Proof. To begin with, let $m_x < g$. We use the model

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{F}_\mathbf{y}\gamma_i + \boldsymbol{\varepsilon}_i, \tag{3.1}$$
$$\mathbf{X}_i = \mathbf{F}_\mathbf{x}\boldsymbol{\Gamma}_i + \mathbf{V}_i, \tag{3.2}$$

which leads to the expansion of the CCEP estimator in the following way:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} &= \left(\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&= \left(\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&= \left(\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\beta}_i + \frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_\mathbf{y}\gamma_i + \frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right) \\
&= \boldsymbol{\beta} + \left(\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i + \frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_\mathbf{y}(\gamma + \boldsymbol{\eta}_{\gamma,i}) + \frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right).
\end{aligned} \tag{3.3}$$

This leads to

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) &= \left(\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i \\
&+ \left(\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_\mathbf{y}\boldsymbol{\eta}_{\gamma,i} \\
&+ \left(\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \\
&+ \left(\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_\mathbf{y}\gamma \\
&= \mathbf{I} + \mathbf{II} + \mathbf{III} + \mathbf{IV}
\end{aligned} \tag{3.4}$$

By using the fact that $\mathbf{F}_\mathbf{x} = (\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} - \overline{\mathbf{V}}_{\dot{\mathbf{x}}})\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^+$, $\mathbf{X}_i = (\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} - \overline{\mathbf{V}}_{\dot{\mathbf{x}}})\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^+\boldsymbol{\Gamma}_i + \mathbf{V}_i$ and hence $\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} = \mathbf{0}_{T\times k}$, we obtain

$$\begin{aligned}
\frac{1}{NT}\sum_{i=1}^N \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i &= \frac{1}{NT}\sum_{i=1}^N (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^+\boldsymbol{\Gamma}_i)'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}(\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^+\boldsymbol{\Gamma}_i) \\
&= \frac{1}{N}\sum_{i=1}^N T^{-1}\mathbf{V}_i'\mathbf{V}_i + O_p(T^{-1/2}) \\
&= \boldsymbol{\Sigma} + O_p(T^{-1/2}),
\end{aligned} \tag{3.5}$$

which comes directly from Lemma B-7 leading up to Theorem 4 in De Vos and Stauskas (2024), in addition to $T^{-1}\mathbf{V}_i'\mathbf{V}_i = \mathbf{\Sigma}_i + O_p(T^{-1/2})$. There it is assumed that $\mathbf{F_x} = \mathbf{F_y} = \mathbf{F}$ and $\widehat{\mathbf{F}} = [\overline{\mathbf{y}}, \overline{\mathbf{X}}]$, which means that (3.5) is a special case and the same rate of convergence applies. By using the same Lemma B-7 and Theorem 4 in De Vos and Stauskas (2024) in connection to (3.5) we have that

$$\mathbf{III} = \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i = o_p(1) \tag{3.6}$$

and

$$\mathbf{I} = \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i = \mathbf{\Sigma}^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (T^{-1} \mathbf{V}_i' \mathbf{V}_i) \boldsymbol{\nu}_i + o_p(1), \tag{3.7}$$

which means that the slope heterogeneity dominates $\boldsymbol{\varepsilon}_i$ in the asymptotic distribution. Again, these results follow, because in the heterogeneous slope analysis in De Vos and Stauskas (2024) we have $\mathbf{F_x} = \mathbf{F_y} = \mathbf{F}$ and $\widehat{\mathbf{F}} = [\overline{\mathbf{y}}, \overline{\mathbf{X}}]$, thus the rates of convergence here are preserved or faster when only $\overline{\mathbf{X}}$ is employed. As such,

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) &= \mathbf{\Sigma}^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (T^{-1} \mathbf{V}_i' \mathbf{V}_i) \boldsymbol{\nu}_i \\
&\quad + \underbrace{\left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1}}_{O_p(1)} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F_y} \boldsymbol{\eta}_{\gamma,i} \\
&\quad + \underbrace{\left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1}}_{O_p(1)} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F_y} \boldsymbol{\gamma} + o_p(1).
\end{aligned} \tag{3.8}$$

Note that $\mathbf{IV}$ is algebraically equal to $\mathbf{0}_k$ if $\mathbf{q}_{\dot{\mathbf{x}}} = \mathbf{I}_k$. Otherwise, it has nearly identical structure to $\mathbf{II}$. Therefore, we will now examine $\mathbf{II}$, and we will focus on its numerator. Because $\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} = \mathbf{M}_{\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}}$ since $\mathbf{R}_{\dot{\mathbf{x}}} = \mathbf{T}_{\dot{\mathbf{x}}} \overline{\mathbf{H}}_{\dot{\mathbf{x}}} \mathbf{D}_N$ is full-rank, we now decompose the numerator of $\mathbf{II}$ as

$$\begin{aligned}
\frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F_y} \boldsymbol{\eta}_{\gamma,i} &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F_y} \boldsymbol{\eta}_{\gamma,i} \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' \mathbf{M}_{\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}} \mathbf{F_y} \boldsymbol{\eta}_{\gamma,i} \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' \mathbf{M}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} \mathbf{F_y} \boldsymbol{\eta}_{\gamma,i} \\
&\quad - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' (\mathbf{M}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}}) \mathbf{F_y} \boldsymbol{\eta}_{\gamma,i} \\
&= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' \mathbf{F_y} \boldsymbol{\eta}_{\gamma,i} - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' \mathbf{P}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} \mathbf{F_y} \boldsymbol{\eta}_{\gamma,i} \\
&\quad - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' (\mathbf{M}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}}) \mathbf{F_y} \boldsymbol{\eta}_{\gamma,i} \\
&= \mathbf{A} - \mathbf{B} - \mathbf{C}.
\end{aligned} \tag{3.9}$$

We start from $\mathbf{A}$, which leads to

$$\mathbf{A} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} - \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}$$
$$= O_p(T^{-1/2}), \tag{3.10}$$

because

$$\begin{aligned}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\| &= \left\|\frac{1}{N}\sum_{i=1}^{N} T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\sqrt{N}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\| \\ &\leq \frac{1}{N}\sum_{i=1}^{N}\left\|\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i\right\|\left\|\boldsymbol{\eta}_{\gamma,i}\right\|\left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\right\| \\ &= O_p(T^{-1/2})\end{aligned} \tag{3.11}$$

and by cross-section independence of the error terms

$$\begin{aligned}\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\|^2\right) &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left(T^{-2}\mathbf{tr}\left[\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\boldsymbol{\eta}_{\gamma,j}'\mathbf{F}_{\mathbf{y}}'\mathbf{V}_j\right]\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(T^{-2}\mathbf{tr}\left[\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\boldsymbol{\eta}_{\gamma,i}'\mathbf{F}_{\mathbf{y}}'\mathbf{V}_i\right]\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(T^{-2}\mathbf{tr}\left[\boldsymbol{\eta}_{\gamma,i}'\mathbf{F}_{\mathbf{y}}'\mathbf{V}_i\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right]\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{tr}\left[\mathbb{E}(\boldsymbol{\eta}_{\gamma,i}\boldsymbol{\eta}_{\gamma,i}')\mathbb{E}(T^{-2}\mathbf{F}_{\mathbf{y}}'\mathbf{V}_i\mathbf{V}_i'\mathbf{F}_{\mathbf{y}})\right]\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{tr}\left[\mathbb{E}(\boldsymbol{\eta}_{\gamma,i}\boldsymbol{\eta}_{\gamma,i}')\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}(\mathbf{f}_{\mathbf{y},t}\mathbf{v}_{i,t}'\mathbf{v}_{i,s}\mathbf{f}_{\mathbf{y},s}')\right]\right) \\ &= O(T^{-1})\end{aligned} \tag{3.12}$$

due to summable covariances. Further, we look into $\mathbf{B}$, and in particular we get

$$\begin{aligned}\mathbf{B} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} - \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\ &= O_p(T^{-1/2}),\end{aligned}$$

because

$$\begin{aligned}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\| &= \left\|\frac{1}{N}\sum_{i=1}^{N} T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\sqrt{N}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\| \\ &\leq \frac{1}{N}\sum_{i=1}^{N}\left\|\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i\right\|\left\|\boldsymbol{\eta}_{\gamma,i}\right\|\left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\ddot{\mathbf{x}}}^0\right\|\left\|(T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\ddot{\mathbf{x}}}^0)^{+}\right\|\left\|T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\right\| \\ &= O_p(T^{-1/2})\end{aligned} \tag{3.13}$$

as $\left\| \sqrt{N} T^{-1} \overline{\mathbf{V}}'_{\dot{\mathbf{x}}} \mathbf{F}^0_{\dot{\mathbf{x}}} \right\| = O_p(T^{-1/2})$ and

$$\begin{aligned}
\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{V}'_i \mathbf{P}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| &= \left\| vec \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{V}'_i \mathbf{P}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right) \right\| \\
&= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\boldsymbol{\eta}'_{\gamma,i} \otimes T^{-1} \mathbf{V}'_i \mathbf{F}^0_{\dot{\mathbf{x}}} \right) vec \left((T^{-1} \mathbf{F}^{0\prime}_{\dot{\mathbf{x}}} \mathbf{F}^0_{\dot{\mathbf{x}}})^{+} T^{-1} \mathbf{F}^{0\prime}_{\dot{\mathbf{x}}} \mathbf{F}_{\mathbf{y}} \right) \right\| \\
&\leq \underbrace{\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\boldsymbol{\eta}'_{\gamma,i} \otimes T^{-1} \mathbf{V}'_i \mathbf{F}^0_{\dot{\mathbf{x}}} \right) \right\|}_{O_p(T^{-1/2})} \left\| vec \left((T^{-1} \mathbf{F}^{0\prime}_{\dot{\mathbf{x}}} \mathbf{F}^0_{\dot{\mathbf{x}}})^{+} T^{-1} \mathbf{F}^{0\prime}_{\dot{\mathbf{x}}} \mathbf{F}_{\mathbf{y}} \right) \right\| \\
&= O_p(T^{-1/2})
\end{aligned} \tag{3.14}$$

by the exact same argument as in (3.12). Particularly, by using the Kronecker properties, cross-section independence of the error terms and $\text{tr}(\mathbf{A}'\mathbf{A}) = \text{tr}(\mathbf{A}\mathbf{A}')$, we obtain

$$\begin{aligned}
\mathbb{E}\left(\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\boldsymbol{\eta}'_{\gamma,i} \otimes T^{-1} \mathbf{V}'_i \mathbf{F}^0_{\dot{\mathbf{x}}} \right) \right\|^2 \right) &= \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{E}\left(\text{tr}\left[\boldsymbol{\eta}'_{\gamma,i} \boldsymbol{\eta}_{\gamma,j} \otimes T^{-2} \mathbf{V}'_i \mathbf{F}^0_{\dot{\mathbf{x}}} \mathbf{F}^{0\prime}_{\dot{\mathbf{x}}} \mathbf{V}_j \right] \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\left(\text{tr}\left[\boldsymbol{\eta}'_{\gamma,i} \boldsymbol{\eta}_{\gamma,i} \otimes T^{-2} \mathbf{V}'_i \mathbf{F}^0_{\dot{\mathbf{x}}} \mathbf{F}^{0\prime}_{\dot{\mathbf{x}}} \mathbf{V}_i \right] \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\left(\boldsymbol{\eta}'_{\gamma,i} \boldsymbol{\eta}_{\gamma,i} \right) \text{tr}\left[\mathbb{E}\left(T^{-2} \mathbf{V}'_i \mathbf{F}^0_{\dot{\mathbf{x}}} \mathbf{F}^{0\prime}_{\dot{\mathbf{x}}} \mathbf{V}_i \right) \right] \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\left(\boldsymbol{\eta}'_{\gamma,i} \boldsymbol{\eta}_{\gamma,i} \right) \frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \text{tr}\left[\mathbb{E}\left(\mathbf{v}_{i,t} \mathbf{f}^{0\prime}_{\dot{\mathbf{x}},t} \mathbf{f}^0_{\dot{\mathbf{x}},s} \mathbf{v}'_{i,s} \right) \right] \\
&= O(T^{-1}).
\end{aligned} \tag{3.15}$$

Lastly, we show that $\mathbf{C}$ is negligible as well. To demonstrate this, we re-state the fact that

$$\begin{aligned}
\mathbf{M}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}} = {} & T^{-1} \overline{\mathbf{V}}^0_{-m_x} (T^{-1} \overline{\mathbf{V}}^{0\prime}_{-m_x} \overline{\mathbf{V}}^0_{-m_x})^{+} \overline{\mathbf{V}}^{0\prime}_{-m_x} + T^{-1} \overline{\mathbf{V}}^0_{m_x} (T^{-1} \mathbf{F}'_{\mathbf{x}} \mathbf{F}_{\mathbf{x}})^{+} \overline{\mathbf{V}}^{0\prime}_{m_x} \\
& + T^{-1} \overline{\mathbf{V}}^0_{m_x} (T^{-1} \mathbf{F}'_{\mathbf{x}} \mathbf{F}_{\mathbf{x}})^{+} \mathbf{F}'_{\mathbf{x}} + T^{-1} \mathbf{F}_{\mathbf{x}} (T^{-1} \mathbf{F}'_{\mathbf{x}} \mathbf{F}_{\mathbf{x}})^{+} \overline{\mathbf{V}}^{0\prime}_{m_x} \\
& + T^{-1} \widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}} \left[(T^{-1} \widehat{\mathbf{F}}^{0\prime}_{\dot{\mathbf{x}}} \widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}})^{+} - \boldsymbol{\Sigma}^{+}_{\mathbf{F}^0_{\dot{\mathbf{x}},v}} \right] \widehat{\mathbf{F}}^{0\prime}_{\dot{\mathbf{x}}},
\end{aligned} \tag{3.16}$$

which comes from performing the same manipulations as in S25 - S29 from the supplementary material of Karabiyik et al. (2017). Therefore, we obtain

$$\begin{aligned}
\mathbf{C} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0}(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0}(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'T^{-1}\mathbf{F}_{\mathbf{x}}(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0}\left[(T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^{+}\right]\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&= \mathbf{C1} + \mathbf{C2} + \mathbf{C3} + \mathbf{C4} + \mathbf{C5}, \qquad (3.17)
\end{aligned}$$

where each of the terms is negligible. We will start with **C1** and **C5**, which require the most work. In particular,

$$\begin{aligned}
\mathbf{C1} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i'T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&- \frac{1}{N}\sum_{i=1}^{N} T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\sqrt{N}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i'T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} + O_p(T^{-1/2}), \qquad (3.18)
\end{aligned}$$

since

$$\begin{aligned}
&\left\|\frac{1}{N}\sum_{i=1}^{N} T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\sqrt{N}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\| \\
&\leq \left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\right\|\left\|(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\right\|\left\|\|T^{-1}\sqrt{N}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|\overline{\boldsymbol{\Gamma}}^{+}\boldsymbol{\Gamma}_i\right\|\left\|\boldsymbol{\eta}_{\gamma,i}\right\| \\
&= O_p(T^{-1/2}). \qquad (3.19)
\end{aligned}$$

By defining $\widehat{\mathbf{D}}_{\dot{\mathbf{x}},-m_x} = \mathbf{q}_{\ddot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}},-m_x}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{H}}_{\dot{\mathbf{x}},-m_x}\mathbf{q}_{\ddot{\mathbf{x}}}'$, the first term can be simplified in the following way:

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^{+}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i}\\
&=\frac{1}{\sqrt{N}T^2}\sum_{i=1}^{N}N\mathbf{V}_i'\overline{\mathbf{V}}\mathbf{q}_{\ddot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}},-m_x}(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^{+}\overline{\mathbf{H}}_{\dot{\mathbf{x}},-m_x}\mathbf{q}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}'\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i}\\
&=\frac{1}{N\sqrt{N}T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}\mathbf{V}_i'\mathbf{V}_j\widehat{\mathbf{D}}_{\dot{\mathbf{x}},-m_x}\mathbf{V}_l'\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i}\\
&=\sum_{u=1}^{k}\sum_{v=1}^{k}\widehat{d}_{\dot{\mathbf{x}},-m_x,u,v}\frac{1}{N\sqrt{N}T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}\mathbf{V}_i'\mathbf{V}_j^{(u)}\mathbf{V}_l^{(v)\prime}\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i},
\end{aligned}
\tag{3.20}
$$

where $\widehat{d}_{\dot{\mathbf{x}},-m_x,u,v}$ is an element in row u and column v in $\mathbf{D}_{\dot{\mathbf{x}},-m_x}$. Therefore,

$$
\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^{+}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i}\right\|\\
&=\left\|\sum_{u=1}^{k}\sum_{v=1}^{k}\widehat{d}_{\dot{\mathbf{x}},-m_x,u,v}\frac{1}{N\sqrt{N}T^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}\mathbf{V}_i'\mathbf{V}_j^{(u)}\mathbf{V}_l^{(v)\prime}\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i}\right\|\\
&\leq\sum_{u=1}^{k}\sum_{v=1}^{k}\left|\widehat{d}_{\dot{\mathbf{x}},-m_x,u,v}\right|\frac{1}{\sqrt{N}}\underbrace{\left\|\frac{1}{NT^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{l=1}^{N}\mathbf{V}_i'\mathbf{V}_j^{(u)}\mathbf{V}_l^{(v)\prime}\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i}\right\|}_{O_p(T^{-1/2})}\\
&=O_p((NT)^{-1/2}),
\end{aligned}
\tag{3.21}
$$

where the $O_p(T^{-1/2})$ component is established in (2.80) of the supplementary material of De Vos and Stauskas (2024), where they demonstrate the the normalized triple sum of with the triples of the same variable multiplied by the fourth independent variable follows this order under our assumptions. Indeed, $\{\mathbf{f}_{\mathbf{y}}'\boldsymbol{\eta}_{\gamma,i}\}_{t=1}^{T}$ is a zero-mean process independent from the model errors. Alternatively, this can be demonstrated with

$$
\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^{+}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i}\right\|\\
&=\left\|vec\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^{+}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i}\right)\right\|\\
&=\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\eta}_{\gamma,i}'\otimes T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}\right)vec\left((T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^{+}T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\mathbf{F_y}\right)\right\|\\
&\leq\frac{1}{\sqrt{T}}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\eta}_{\gamma,i}'\otimes T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}\right)\right\|\left\|(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^{+}\right\|\left\|T^{-1/2}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\mathbf{F_y}\right\|\\
&=O_p(T^{-1})+O_p((NT)^{-1/2}),
\end{aligned}
\tag{3.22}
$$

although at a slightly different rate. Nevertheless, this rate is sufficient show that in summary

$$
\|\mathbf{C1}\|=\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^{+}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i}\right\|=O_p(T^{-1/2}).
\tag{3.23}
$$

We next move on to **C5**:

$$\begin{aligned}
\mathbf{C5} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)' T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+}\right]\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i' T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+}\right]\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&- \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\dot{\mathbf{x}}}' T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+}\right]\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i' T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+}\right]\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} + O_p(T^{-1/2}) + O_p(N^{-1/2}) \qquad (3.24)
\end{aligned}$$

since

$$\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\dot{\mathbf{x}}}' T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+}\right]\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\| \\
&\leq \left\|(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+}\right\|\left\|T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\right\|\left\|T^{-1}\sqrt{N}\overline{\mathbf{V}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i\right\|\left\|\boldsymbol{\eta}_{\gamma,i}\right\| \\
&= O_p(T^{-1/2}) + O_p(N^{-1/2}), \qquad (3.25)
\end{aligned}$$

because $\left\|T^{-1}\sqrt{N}\overline{\mathbf{V}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\right\| \leq \left\|T^{-1}\sqrt{N}\overline{\mathbf{V}}_{\dot{\mathbf{x}}}'\overline{\mathbf{V}}_{\dot{\mathbf{x}}}^{0}\right\| + \left\|T^{-1}\sqrt{N}\overline{\mathbf{V}}_{\dot{\mathbf{x}}}'\mathbf{F}_{\dot{\mathbf{x}}}^{0}\right\| = \left\|T^{-1}\sqrt{N}\overline{\mathbf{V}}_{\dot{\mathbf{x}}}'\overline{\mathbf{V}}_{\dot{\mathbf{x}}}^{0}\right\| + O_p(T^{-1/2}) = O_p(1)$ and $\left\|T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\right\| = O_p(1)$. Next up, we re-write the first term in vectorized form to obtain

$$\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i' T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+}\right]\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\| \\
&= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\eta}_{\gamma,i}' \otimes T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\right) vec\left(\underbrace{\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+}\right]}_{O_p(N^{-1/2}) + O_p(T^{-1/2})} T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\right)\right\| \\
&\leq \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\eta}_{\gamma,i}' \otimes T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\right)\right\|\left\|vec\left(\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+}\right] T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\right)\right\| \\
&\leq \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\eta}_{\gamma,i}' \otimes T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0}\right)\right\|\left\|(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+}\right\|\left\|T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\right\| \\
&= O_p(N^{-1}) + O_p(T^{-1}), \qquad (3.26)
\end{aligned}$$

because the first component is asymptotically negligible, as well. Particularly, by using cross-section independence of the loadings, multiplication properties of the Kronecker product and the fact that $\mathrm{tr}(\mathbf{A}'\mathbf{A}) =$

$\text{tr}(\mathbf{A}\mathbf{A}')$, we obtain

$$\begin{aligned}
&\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\eta}'_{\gamma,i}\otimes T^{-1}\mathbf{V}'_i\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}\right)\right\|^2\right)\\
&=\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left(\text{tr}\left[\boldsymbol{\eta}_{\gamma,i}\boldsymbol{\eta}_{\gamma,j}\otimes T^{-1}\mathbf{V}'_i\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\dot{\mathbf{x}}}\mathbf{V}_j)\right]\right)\\
&=\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\text{tr}\left[\boldsymbol{\eta}'_{\gamma,i}\boldsymbol{\eta}_{\gamma,i}\otimes T^{-1}\mathbf{V}'_i\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\dot{\mathbf{x}}}\mathbf{V}_i)\right]\right)\\
&=\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\boldsymbol{\eta}'_{\gamma,i}\boldsymbol{\eta}_{\gamma,i}\text{tr}\left[T^{-1}\mathbf{V}'_i\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\dot{\mathbf{x}}}\mathbf{V}_i)\right]\right)\\
&=\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\boldsymbol{\eta}'_{\gamma,i}\boldsymbol{\eta}_{\gamma,i}\right)\mathbb{E}\left(\text{tr}\left[T^{-1}\mathbf{V}'_i\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\dot{\mathbf{x}}}\mathbf{V}_i)\right]\right)\\
&=O(N^{-1})+O(T^{-1}),
\end{aligned}\tag{3.27}$$

because $\left\|T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\dot{\mathbf{x}}}\mathbf{V}_i\right\|\leq\left\|T^{-1}\mathbf{V}'_i\overline{\mathbf{V}}^0_{\dot{\mathbf{x}}}\right\|+\|T^{-1}\mathbf{V}'_i\mathbf{F}^0_{\dot{\mathbf{x}}}\|=(O_p(N^{-1/2})+O_p(T^{-1/2}))+O_p(T^{-1/2})=O_p(N^{-1/2})+O_p(T^{-1/2})$. This means that overall

$$\begin{aligned}
\mathbf{C5}&=\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}^+_{\dot{\mathbf{x}}}\boldsymbol{\Gamma}_i)'T^{-1}\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}\left[(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\dot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}})^+-\boldsymbol{\Sigma}^+_{\mathbf{F}^0_{\dot{\mathbf{x}},v}}\right]\widehat{\mathbf{F}}^{0\prime}_{\dot{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\\
&=O_p(N^{-1/2})+O_p(T^{-1/2}).
\end{aligned}\tag{3.28}$$

We will finish by analysing **C2**, **C3** and **C4**, which all have a similar structure. For instance,

$$\begin{aligned}
\|\mathbf{C2}\|&=\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i)'T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}(T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{x}})^+\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\|\\
&\leq\left\|(T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{x}})^+\right\|\left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x}\mathbf{F}_{\mathbf{y}}\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i)'\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x})\right\|\left\|\boldsymbol{\eta}_{\gamma,i}\right\|\\
&=O_p(T^{-1/2})\left(O_p(N^{-1})+O_p((NT)^{-1/2})\right)\\
&=O_p(N^{-1}T^{-1/2})+O_p(N^{-1/2}T^{-1})
\end{aligned}\tag{3.29}$$

and

$$\begin{aligned}
\|\mathbf{C3}\|&=\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i)'T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}(T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{x}})^+\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\|\\
&\leq\left\|T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{y}}\right\|\left\|(T^{-1}\mathbf{F}'_{\mathbf{x}}\mathbf{F}_{\mathbf{x}})^+\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i)'\sqrt{N}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}\right\|\left\|\boldsymbol{\eta}_{\gamma,i}\right\|\\
&=O_p(N^{-1/2})+O_p(T^{-1/2}),
\end{aligned}\tag{3.30}$$

since $\left\| \sqrt{N} T^{-1} \mathbf{V}_i' \overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x} \right\| = O_p(N^{-1/2}) + O_p(T^{-1/2})$ and $\left\| T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x} \right\| = O_p(N^{-1/2})$. Finally,

$$
\begin{aligned}
\|\mathbf{C4}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}} \mathbf{\Gamma}_i)' T^{-1} \mathbf{F}_{\mathbf{x}} (T^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{x}})^+ \overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&\leq \left\| \sqrt{N} T^{-1} \overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \mathbf{F}_{\mathbf{y}} \right\| \left\| (T^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{x}})^+ \right\| \frac{1}{N} \sum_{i=1}^{N} \left\| T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}} \mathbf{\Gamma}_i)' \mathbf{F}_{\mathbf{x}} \right\| \left\| \boldsymbol{\eta}_{\gamma,i} \right\| \\
&= O_p(T^{-1/2}) \left(O_p(T^{-1/2}) + O_p((NT)^{-1/2}) \right) \\
&= O_p(T^{-1}).
\end{aligned} \tag{3.31}
$$

Hence, by combining the rates of **C1** - **C5**, we have that

$$
\begin{aligned}
\|\mathbf{C}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}} \mathbf{\Gamma}_i)' (\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}}) \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}),
\end{aligned} \tag{3.32}
$$

and in connection to the rates of **A** and **B**, we obtain

$$
\begin{aligned}
\|\mathbf{II}\| &= \left\| \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&\leq \underbrace{\left\| \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \right\|}_{O_p(1)} \underbrace{\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\|}_{O_p(N^{-1/2}) + O_p(T^{-1/2})} \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}).
\end{aligned} \tag{3.33}
$$

We are left to deal with **IV**. Note that it follows exactly the same analysis as **II** and will retain the same order results if we replace $\boldsymbol{\eta}_{\gamma,i}$ with $\boldsymbol{\gamma}$ in any of the equations above, because the steps do not depend on the statistical properties of the loadings. For example, (3.12) and (3.15) are solely driven by the covariance summability and not the loading properties. By using $\text{tr}(\mathbf{A}'\mathbf{A}) = \text{tr}(\mathbf{A}\mathbf{A}')\text{T}$, this gives

$$
\begin{aligned}
\mathbb{E} \left(\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{V}_i' \mathbf{F}_{\mathbf{y}} \gamma \right\|^2 \right) &= \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{E} \left(T^{-2} \mathbf{tr} \left[\mathbf{V}_i' \mathbf{F}_{\mathbf{y}} \gamma \gamma' \mathbf{F}_{\mathbf{y}}' \mathbf{V}_j \right] \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left(T^{-2} \mathbf{tr} \left[\mathbf{V}_i' \mathbf{F}_{\mathbf{y}} \gamma \gamma' \mathbf{F}_{\mathbf{y}}' \mathbf{V}_i \right] \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left(T^{-2} \mathbf{tr} \left[\gamma' \mathbf{F}_{\mathbf{y}}' \mathbf{V}_i \mathbf{V}_i' \mathbf{F}_{\mathbf{y}} \gamma \right] \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \left(\text{tr} \left[\gamma \gamma' \mathbb{E} (T^{-2} \mathbf{F}_{\mathbf{y}}' \mathbf{V}_i \mathbf{V}_i' \mathbf{F}_{\mathbf{y}}) \right] \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \left(\text{tr} \left[\gamma \gamma' \frac{1}{T^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \mathbb{E} (\mathbf{f}_{\mathbf{y},t} \mathbf{v}_{i,t}' \mathbf{v}_{i,s} \mathbf{f}_{\mathbf{y},s}') \right] \right) \\
&= O(T^{-1}),
\end{aligned} \tag{3.34}
$$

and similarly by cross-section independence

$$\begin{aligned}
\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\gamma}'\otimes T^{-1}\mathbf{V}_i'\mathbf{F}_{\dot{\mathbf{x}}}^0\right)\right\|^2\right) &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left(\mathrm{tr}\left[\boldsymbol{\gamma}'\boldsymbol{\gamma}\otimes T^{-2}\mathbf{V}_i'\mathbf{F}_{\dot{\mathbf{x}}}^0\mathbf{F}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{V}_j\right]\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\mathrm{tr}\left[\boldsymbol{\gamma}'\boldsymbol{\gamma}\otimes T^{-2}\mathbf{V}_i'\mathbf{F}_{\dot{\mathbf{x}}}^0\mathbf{F}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{V}_i\right]\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\gamma}'\boldsymbol{\gamma}\mathrm{tr}\left[\mathbb{E}\left(T^{-2}\mathbf{V}_i'\mathbf{F}_{\dot{\mathbf{x}}}^0\mathbf{F}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{V}_i\right)\right]\\
&= \boldsymbol{\gamma}'\boldsymbol{\gamma}\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathrm{tr}\left[\mathbb{E}\left(\mathbf{v}_{i,t}\mathbf{f}_{\dot{\mathbf{x}},t}^{0\prime}\mathbf{f}_{\dot{\mathbf{x}},s}^{0}\mathbf{v}_{i,s}'\right)\right]\\
&= O(T^{-1}).
\end{aligned} \tag{3.35}$$

The two exceptions are (3.26) and (3.21), which slightly change. In particular,

$$\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0)^+-\boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^0}^+\right]\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}\right\|\\
&=\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\gamma}'T^{-1}\mathbf{F}_{\mathbf{y}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0\otimes T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0\right)\right\|\left\|vec\left(\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0)^+-\boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^0}^+\right]\right)\right\|\\
&\leq\left\|\boldsymbol{\gamma}'T^{-1}\mathbf{F}_{\mathbf{y}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0\otimes\sqrt{N}T^{-1}\overline{\mathbf{V}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0\right\|\left\|vec\left(\underbrace{\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0)^+-\boldsymbol{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^0}^+\right]}_{O_p(N^{-1/2})+O_p(T^{-1/2})}\right)\right\|\\
&=O_p(N^{-1/2})+O_p(T^{-1/2}),
\end{aligned} \tag{3.36}$$

because $\sqrt{N}T^{-1}\overline{\mathbf{V}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0$ is bounded. Also,

$$\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^+\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}\right\|\\
&=\left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\widehat{\boldsymbol{\Sigma}}_{\mathbf{v}_{\ddot{\mathbf{x}},-m_x}^0}^+T^{-1}(\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)'\mathbf{F}_y\boldsymbol{\gamma}\right\|\\
&\leq\left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\right\|\left\|\widehat{\boldsymbol{\Sigma}}_{\mathbf{v}_{\ddot{\mathbf{x}},-m_x}^0}^+\right\|\left\|T^{-1}(\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)'\mathbf{F}_y\boldsymbol{\gamma}\right\|=O_p(T^{-1/2}).
\end{aligned} \tag{3.37}$$

This means that

$$\begin{aligned}
\|\mathbf{IV}\|&\leq\underbrace{\left\|\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\|}_{O_p(1)}\underbrace{\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}\right\|}_{O_p(N^{-1/2})+O_p(T^{-1/2})}\\
&=O_p(N^{-1/2})+O_p(T^{-1/2}).
\end{aligned} \tag{3.38}$$

By putting the results together, we simplify (3.8) and obtain the asymptotic distribution by standard

Lindeberg-Lévy Central Limit Theorem:

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) &= \boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(T^{-1}\mathbf{V}_i'\mathbf{V}_i)\boldsymbol{\nu}_i + o_p(1) \\
&= \boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i\boldsymbol{\nu}_i + o_p(1) \\
&\to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1}\right)
\end{aligned} \tag{3.39}$$

as $(N, T) \to \infty$, where $\boldsymbol{\Psi}_{\nu} = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i\boldsymbol{\Omega}_{\nu}\boldsymbol{\Sigma}_i$. The simplification comes from

$$\begin{aligned}
\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[(T^{-1}\mathbf{V}_i'\mathbf{V}_i) - \boldsymbol{\Sigma}_i\right]\boldsymbol{\nu}_i\right\|^2\right) &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{tr}\left[\mathbb{E}\left(\left[(T^{-1}\mathbf{V}_i'\mathbf{V}_i) - \boldsymbol{\Sigma}_i\right]\boldsymbol{\nu}_i\boldsymbol{\nu}_j\left[T^{-1}\mathbf{V}_j'\mathbf{V}_j) - \boldsymbol{\Sigma}_j\right]'\right)\right] \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbf{tr}\left[\mathbb{E}\left(\left[(T^{-1}\mathbf{V}_i'\mathbf{V}_i) - \boldsymbol{\Sigma}_i\right]\boldsymbol{\Omega}_{\nu}\left[T^{-1}\mathbf{V}_i'\mathbf{V}_i) - \boldsymbol{\Sigma}_i\right]'\right)\right] \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbf{tr}\left[\boldsymbol{\Omega}_{\nu}\mathbb{E}\left(\left[(T^{-1}\mathbf{V}_i'\mathbf{V}_i) - \boldsymbol{\Sigma}_i\right]'\left[(T^{-1}\mathbf{V}_i'\mathbf{V}_i) - \boldsymbol{\Sigma}_i\right]\right)\right] \\
&= O(T^{-1}).
\end{aligned} \tag{3.40}$$

Now, we let $m_x = g$, which means that we will use the expansion

$$\begin{aligned}
\mathbf{M}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}} = \mathbf{M}_{\mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} &= T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' + T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{x}}' \\
&+ T^{-1}\mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' + T^{-1}\mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+} - (\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}'T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}})^{+}]\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{x}}'.
\end{aligned} \tag{3.41}$$

Under $m_x = g$ case the results of De Vos and Stauskas (2024) hold, and so we arrive at the approximation in (3.8), where the remainder is of even lower order. In order to verify that the results hold, we only look at the most complex term **C** in (3.9) as the analysis of **A** and **B** would stay exactly the same and they will be negligible. This is so, because

$$(\mathbf{F}^{0\prime}_{\dot{\mathbf{x}}}\mathbf{F}^0_{\dot{\mathbf{x}}})^{+} = \begin{bmatrix} \mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}} & \mathbf{0}_{m_x\times(g-m_x)} \\ \mathbf{0}_{(g-m_x)\times m_x} & \mathbf{0}_{(g-m_x)} \end{bmatrix}^{+} = \begin{bmatrix} (\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+} & \mathbf{0}_{m_x\times(g-m_x)} \\ \mathbf{0}_{(g-m_x)\times m_x} & \mathbf{0}_{(g-m_x)} \end{bmatrix},$$

leading to

$$\begin{aligned}
\mathbf{P}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} = \mathbf{F}^0_{\dot{\mathbf{x}}}(\mathbf{F}^{0\prime}_{\dot{\mathbf{x}}}\mathbf{F}^0_{\dot{\mathbf{x}}})^{+}\mathbf{F}^{0\prime}_{\dot{\mathbf{x}}} &= \begin{bmatrix} \mathbf{F}_{\mathbf{x}}, \mathbf{0}_{T\times(g-m_x)} \end{bmatrix}\begin{bmatrix} (\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+} & \mathbf{0}_{m_x\times(g-m_x)} \\ \mathbf{0}_{(g-m_x)\times m_x} & \mathbf{0}_{(g-m_x)} \end{bmatrix}\begin{bmatrix} \mathbf{F}_{\mathbf{x}}' \\ \mathbf{0}_{(g-m_x)\times T} \end{bmatrix} \\
&= \mathbf{F}_{\mathbf{x}}'(\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+}\mathbf{F}_{\mathbf{x}} = \mathbf{P}_{\mathbf{F}_{\mathbf{x}}}.
\end{aligned}$$

Then, particularly for $\mathbf{C}$, we have

$$
\begin{aligned}
\|\mathbf{C}\| &= \left\| \frac{1}{N} \sum_{i=1}^{N} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' (\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0}) \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&\leq \left\| \frac{1}{N} \sum_{i=1}^{N} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&+ \left\| \frac{1}{N} \sum_{i=1}^{N} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&+ \left\| \frac{1}{N} \sum_{i=1}^{N} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \mathbf{F}_{\mathbf{x}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&+ \left\| \frac{1}{N} \sum_{i=1}^{N} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \mathbf{F}_{\mathbf{x}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}} [(T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} - (\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}' T^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{x}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}})^{+}] \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}),
\end{aligned} \tag{3.42}
$$

which is driven by the highest order component

$$
\begin{aligned}
&\left\| \frac{1}{N} \sum_{i=1}^{N} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&\leq \left\| (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \right\| \left\| T^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \right\| \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\eta}_{\gamma,i} \right\| \|\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}\| \left\| \sqrt{N} T^{-1} \mathbf{V}_i' \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \right\| \\
&+ \left\| (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \right\| \left\| T^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \right\| \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\eta}_{\gamma,i} \right\| \|\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}\| \left\| \sqrt{N} T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i \right\| \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}).
\end{aligned} \tag{3.43}
$$

The same order result will hold in the expansion equivalent to (3.9) in case of $\mathbf{IV}$, when we replace $\boldsymbol{\eta}_{\gamma,i}$ with γ. By looking at the equivalent leading term, we obtain

$$
\begin{aligned}
&\left\| \frac{1}{N} \sum_{i=1}^{N} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \gamma \right\| \\
&\leq \left\| (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \right\| \left\| T^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \right\| \frac{1}{N} \sum_{i=1}^{N} \|\gamma\| \|\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}\| \left\| \sqrt{N} T^{-1} \mathbf{V}_i' \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \right\| \\
&+ \left\| (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \right\| \left\| T^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \right\| \frac{1}{N} \sum_{i=1}^{N} \|\gamma\| \|\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}\| \left\| \sqrt{N} T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i \right\| \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}).
\end{aligned} \tag{3.44}
$$

3.2 Mean Group Estimator

Theorem 4. *Under Assumptions 1 - 7, for either $m_x < g$ or $m_x = g$ as $(N, T) \to \infty$ with $TN^{-1} \to \tau > 0$*

$$
\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\ddot{\mathbf{x}}} - \boldsymbol{\beta}) \to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Omega}_{\boldsymbol{\nu}}\right),
$$

where $\boldsymbol{\Omega}_{\boldsymbol{\nu}} = \mathbb{E}(\boldsymbol{\nu}_i \boldsymbol{\nu}_i')$.

Proof. Firstly, we assume $m_x < g$. We expand the CCEMG estimator in the following way:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} &= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}(\mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i) \\
&= \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\beta}_i + \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \\
&= \boldsymbol{\beta} + \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\nu}_i + \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i,
\end{aligned} \tag{3.45}$$

which implies that

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\nu}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i \\
&\quad + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \\
&= \mathbf{I} + \mathbf{II} + \mathbf{III}.
\end{aligned} \tag{3.46}$$

Clearly, $\mathbf{I}$ is asymptotically normal by the standard arguments:

$$\mathbf{I} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\nu}_i \to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Omega}_{\nu}\right), \tag{3.47}$$

as $(N, T) \to \infty$. We further move to $\mathbf{III}$, which is much simpler than its analog in Theorem 6 of De Vos and Stauskas (2024). In particular, in the later study, $\bar{\boldsymbol{\varepsilon}}$ is used to approximate the factor space via $\widehat{\mathbf{F}} = [\overline{\mathbf{y}}, \overline{\mathbf{X}}]$, which makes the numerator and the denominator dependent for each i. In the current case, we only use $\overline{\mathbf{X}}$ and hence (any subset of) $\overline{\mathbf{V}}$, which is independent from $\boldsymbol{\varepsilon}_i$ for all i. This implies that $\mathbf{III}$ is mean-zero and by our assumptions on existence of moments, we obtain

$$\begin{aligned}
&\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\|^2\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left(\mathrm{tr}\left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-2}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_j'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_j\left(T^{-1}\mathbf{X}_j'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_j\right)^{-1}\right]\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\mathrm{tr}\left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-2}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right]\right) = O(T^{-1}),
\end{aligned} \tag{3.48}$$

which comes from the fact that $\left\|T^{-1/2}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\| = O_p(1)$. This can easily be seen from the expansion similar to (3.9)

$$\begin{aligned}
T^{-1/2}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i &= T^{-1/2}(\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \\
&= T^{-1/2}(\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\boldsymbol{\varepsilon}_i - T^{-1/2}(\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\mathbf{P}_{\mathbf{F}_{\dot{\mathbf{x}}}^0}\boldsymbol{\varepsilon}_i \\
&\quad - T^{-1/2}(\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}_{\dot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0})\boldsymbol{\varepsilon}_i,
\end{aligned} \tag{3.49}$$

where the leading terms are the ones with $\mathbf{V}_i$ from the left, because $\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}$ will either preserve the same order or bring it down. Clearly,

$$\left\|T^{-1/2}\mathbf{V}_i'\boldsymbol{\varepsilon}_i\right\| = O_p(1), \tag{3.50}$$

$$\left\|T^{-1/2}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\boldsymbol{\varepsilon}_i\right\| = O_p(N^{-1/2}) \tag{3.51}$$

under our assumptions. Next,

$$\left\|T^{-1/2}\mathbf{V}_i'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\boldsymbol{\varepsilon}_i\right\| \leq \left\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_{\ddot{\mathbf{x}}}^0\right\| \left\|(T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\ddot{\mathbf{x}}}^0)^+\right\| \left\|T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\boldsymbol{\varepsilon}_i\right\| = O_p(T^{-1/2}), \tag{3.52}$$

$$\left\|T^{-1/2}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}^{+\prime}\overline{\mathbf{V}}'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\boldsymbol{\varepsilon}_i\right\| \leq \left\|T^{-1/2}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\ddot{\mathbf{x}}}^0\right\| \left\|(T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\ddot{\mathbf{x}}}^0)^+\right\| \left\|T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\boldsymbol{\varepsilon}_i\right\| = O_p((NT)^{-1/2}). \tag{3.53}$$

Eventually, by using the expansion in (3.16), we obtain

$$\begin{aligned}\left\|T^{-1/2}\mathbf{V}_i'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\boldsymbol{\varepsilon}_i\right\| &\leq \left\|T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\right\| \left\|(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^+\right\| \left\|T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\boldsymbol{\varepsilon}_i\right\| \\ &+ \left\|T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0\right\| \left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+\right\| \left\|T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\boldsymbol{\varepsilon}_i\right\| \\ &+ \left\|T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0\right\| \left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+\right\| \left\|T^{-1/2}\mathbf{F}_{\mathbf{x}}'\boldsymbol{\varepsilon}_i\right\| \\ &+ \left\|T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{x}}\right\| \left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+\right\| \left\|T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\boldsymbol{\varepsilon}_i\right\| \\ &+ \left\|T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0\right\| \left\|\left[(T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0)^+ - \boldsymbol{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^+\right]\right\| \left\|T^{-1/2}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\boldsymbol{\varepsilon}_i\right\| \\ &= O_p(N^{-1/2}) + O_p(T^{-1/2}),\end{aligned} \tag{3.54}$$

and

$$\begin{aligned}\left\|T^{-1/2}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}^{+\prime}\overline{\mathbf{V}}'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\boldsymbol{\varepsilon}_i\right\| &\leq \left\|T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\right\| \left\|(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^+\right\| \left\|T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\boldsymbol{\varepsilon}_i\right\| \\ &+ \left\|T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}^{+\prime}\overline{\mathbf{V}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0\right\| \left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+\right\| \left\|T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\boldsymbol{\varepsilon}_i\right\| \\ &+ \left\|T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0\right\| \left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+\right\| \left\|T^{-1/2}\mathbf{F}_{\mathbf{x}}'\boldsymbol{\varepsilon}_i\right\| \\ &+ \left\|T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{x}}\right\| \left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+\right\| \left\|T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\boldsymbol{\varepsilon}_i\right\| \\ &+ \left\|T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0\right\| \left\|\left[(T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0)^+ - \boldsymbol{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^+\right]\right\| \left\|T^{-1/2}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\boldsymbol{\varepsilon}_i\right\| \\ &= O_p(N^{-1/2}),\end{aligned} \tag{3.55}$$

since $\left\|T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\right\| = O_p(N^{-1/2}) + O_p(T^{-1/2})$, $\left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\right\| = O_p(N^{-1/2})$, $\left\|T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\boldsymbol{\varepsilon}_i\right\| = O_p(1)$, $\left\|T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{x}}\right\| = O_p(T^{-1/2})$, $\left\|T^{-1/2}\mathbf{F}_{\mathbf{x}}'\boldsymbol{\varepsilon}_i\right\| = O_p(1)$, $\left\|T^{-1/2}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\boldsymbol{\varepsilon}_i\right\| = O_p(1)$ and the rest of the terms are of a lower order. Therefore,

$$\left\|T^{-1/2}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\| = O_p(1), \tag{3.56}$$

$$\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\| = O_p(T^{-1/2}) \tag{3.57}$$

and hence

$$\|\mathbf{III}\| = \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\| = O_p(N^{-1/2}) + O_p(T^{-1/2}). \tag{3.58}$$

We will proceed with **II**. In particular, we can re-write it as

$$
\begin{aligned}
\mathbf{II} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \boldsymbol{\Sigma}_i^{-1}\right]T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \mathbf{A} + \mathbf{B},
\end{aligned} \tag{3.59}
$$

which is not the "sharpest" split of this term, but as we will see, the restriction on N, T expansion will be needed anyway. Here we will focus on **A**, first. We have

$$
\begin{aligned}
\mathbf{A} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\mathbf{F}_{\mathbf{y}}\gamma_i - \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\mathbf{P}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\
&- \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}})\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \mathbf{A_1} - \mathbf{A_2} - \mathbf{A_3},
\end{aligned} \tag{3.60}
$$

where $\|\mathbf{A}_1\| = O_p(T^{-1/2})$, because

$$
\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \leq \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Sigma}_i^{-1}\right\|\left\|\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i\right\|\|\gamma_i\| = O_p(T^{-1/2}), \tag{3.61}
$$

and by the cross-section independence of $\mathbf{V}_i$

$$
\begin{aligned}
\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\gamma_i\right\|^2\right) &= \frac{1}{N}\sum_{i=1}^{N}\mathrm{tr}\left[\mathbb{E}\left(\gamma_i\gamma_i'\right)\mathbb{E}\left(\boldsymbol{\Sigma}_i^{-1}T^{-2}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\mathbf{F}_{\mathbf{y}}'\mathbf{V}_i\boldsymbol{\Sigma}_i^{-1}\right)\right] \\
&= O(T^{-1})
\end{aligned} \tag{3.62}
$$

since $\left\|T^{-1}\mathbf{F}_{\mathbf{y}}'\mathbf{V}_i\right\| = O_p(T^{-1/2})$. The term **A2** follows a similar structure, because

$$
\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{P}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\leq \left\|(T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\ddot{\mathbf{x}}}^{0})^{+}\right\|\left\|T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\right\|\left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}'\mathbf{F}^0_{\ddot{\mathbf{x}}}\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Sigma}_i^{-1}\right\|\left\|\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i\right\|\|\gamma_i\| = O_p(T^{-1/2})
\end{aligned} \tag{3.63}
$$

and

$$
\begin{aligned}
\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\mathbf{V}_i'\mathbf{P}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| &= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\gamma_i'\otimes\boldsymbol{\Sigma}_i^{-1}T^{-1}\mathbf{V}_i'\mathbf{F}^0_{\ddot{\mathbf{x}}}\right)vec\left[(T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\ddot{\mathbf{x}}}^{0})^{+}T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\right]\right\| \\
&\leq \underbrace{\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\gamma_i'\otimes\boldsymbol{\Sigma}_i^{-1}T^{-1}\mathbf{V}_i'\mathbf{F}^0_{\ddot{\mathbf{x}}}\right)\right\|}_{O_p(T^{-1/2})}\left\|(T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\ddot{\mathbf{x}}}^{0})^{+}T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\right\| \\
&= O_p(T^{-1/2}),
\end{aligned} \tag{3.64}
$$

where the order comes by exactly the same argument as in (3.62) by using the Kronecker properties:

$$\begin{aligned}\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\gamma_i'\otimes\mathbf{\Sigma}_i^{-1}T^{-1}\mathbf{V}_i'\mathbf{F}_{\ddot{\mathbf{x}}}^0\right)\right\|^2\right) &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left(\mathrm{tr}\left[\gamma_i'\gamma_j\otimes\mathbf{\Sigma}_i^{-1}T^{-2}\mathbf{V}_i'\mathbf{F}_{\ddot{\mathbf{x}}}^0\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{V}_j\mathbf{\Sigma}_j^{-1}\right]\right)\\ &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\mathrm{tr}\left[\gamma_i'\gamma_i\otimes\mathbf{\Sigma}_i^{-1}T^{-2}\mathbf{V}_i'\mathbf{F}_{\ddot{\mathbf{x}}}^0\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{V}_i\mathbf{\Sigma}_i^{-1}\right]\right)\\ &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\gamma_i'\gamma_i\right)\mathrm{tr}\left[\mathbb{E}\left(\mathbf{\Sigma}_i^{-1}T^{-2}\mathbf{V}_i'\mathbf{F}_{\ddot{\mathbf{x}}}^0\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{V}_i\mathbf{\Sigma}_i^{-1}\right)\right]\\ &= O(T^{-1}).\end{aligned} \tag{3.65}$$

We now move to **A3**, where we again use (3.16):

$$\begin{aligned}\mathbf{A3} &= \frac{1}{N}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}\sqrt{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}-\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i\\ &= \frac{1}{N}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}\sqrt{N}T^{-1}\mathbf{V}_i'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}-\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i-\frac{1}{N}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}\sqrt{N}T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}-\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i\\ &= \mathbf{A3.1}-\mathbf{A3.2},\end{aligned} \tag{3.66}$$

such that

$$\begin{aligned}\|\mathbf{A3.1}\| \le{}& \frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Sigma}_i^{-1}\right\|\left\|\sqrt{N}T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\right\|\left\|(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^{+}\right\|\left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|\\ &+\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Sigma}_i^{-1}\right\|\left\|\sqrt{N}T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0\right\|\left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+}\right\|\left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|\\ &+\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Sigma}_i^{-1}\right\|\left\|\sqrt{N}T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0\right\|\left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+}\right\|\left\|T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\gamma_i\right\|\\ &+\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Sigma}_i^{-1}\right\|\left\|T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{x}}\right\|\left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+}\right\|\left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|\\ &+\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Sigma}_i^{-1}\right\|\left\|T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0\right\|\sqrt{N}\left\|\left[(T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0)^{+}-\mathbf{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^{+}\right]\right\|\left\|T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|\\ ={}& O_p(N^{-1/2})+O_p(T^{-1/2})\end{aligned} \tag{3.67}$$

if we assume that $TN^{-1}=O(1)$. Under this restriction, the first term, which is the dominant one, also becomes negligible, because $\left\|\sqrt{N}T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\right\| = \sqrt{N}(O_p(N^{-1/2})+O_p(T^{-1/2})) = O_p(1)$ then. A similar logic applies to the last term, because $\sqrt{N}\left\|\left[(T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0)^{+}-\mathbf{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^{+}\right]\right\| = O_p(1)$, $\left\|T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| = O_p(1)$ and the total order is driven by the terms of the form $\left\|\sqrt{N}T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}\right\| = O_p(N^{-1/2}) + O_p(T^{-1/2})$. Further,

$$\begin{aligned}
\|\mathbf{A3.2}\| \leq & \frac{1}{N}\sum_{i=1}^N \left\|\boldsymbol{\Sigma}_i^{-1}\left\|\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i\right\|\right\| \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}\right\| \left\|(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^+\right\| \left\|T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
& + \frac{1}{N}\sum_{i=1}^N \left\|\boldsymbol{\Sigma}_i^{-1}\right\| \left\|\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i\right\| \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}\right\| \left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+\right\| \left\|T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
& + \frac{1}{N}\sum_{i=1}^N \left\|\boldsymbol{\Sigma}_i^{-1}\right\| \left\|\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i\right\| \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}\right\| \left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+\right\| \left\|T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
& + \frac{1}{N}\sum_{i=1}^N \left\|\boldsymbol{\Sigma}_i^{-1}\right\| \left\|\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i\right\| \left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{x}}\right\| \left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+\right\| \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
& + \frac{1}{N}\sum_{i=1}^N \left\|\boldsymbol{\Sigma}_i^{-1}\right\| \left\|\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i\right\| \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}\right\| \left\|\left[(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}})^+ - \boldsymbol{\Sigma}^+_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right]\right\| \left\|T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
& = O_p(N^{-1/2}) + O_p(T^{-1/2}) \qquad (3.68)
\end{aligned}$$

by similar arguments, but we do not need $TN^{-1} = O(1)$. This means that overall

$$\|\mathbf{A}\| = \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N \boldsymbol{\Sigma}_i^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| = O_p(N^{-1/2}) + O_p(T^{-1/2}). \qquad (3.69)$$

Eventually, we move to term $\mathbf{B}$, which gives

$$\begin{aligned}
\|\mathbf{B}\| &= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N \left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \boldsymbol{\Sigma}_i^{-1}\right] T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\leq \underbrace{\sqrt{N}\sup_i \left\|\left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \boldsymbol{\Sigma}_i^{-1}\right]\right\|}_{O_p(1) \text{ if } TN^{-1} = O(1)} \frac{1}{N}\sum_{i=1}^N \left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}), \qquad (3.70)
\end{aligned}$$

where the order is dictated by $\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|$, because $\left\|\left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \boldsymbol{\Sigma}_i^{-1}\right]\right\| = O_p(T^{-1/2})$ uniformly as discussed below (3.5). Therefore, we have

$$\begin{aligned}
\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| &\leq \left\|T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i)'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + \left\|T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i)'\mathbf{P}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\quad + \left\|T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}})\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\leq \left\|T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}^+_{\ddot{\mathbf{x}}}\boldsymbol{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}})\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + O_p(T^{-1/2}) \\
&= \|\mathbf{a}\| + \|\mathbf{b}\| + O_p(T^{-1/2}), \qquad (3.71)
\end{aligned}$$

where the dominating order of the remainder is given by the first two terms since $\left\|T^{-1}\mathbf{V}_i'\mathbf{F}^0_{\ddot{\mathbf{x}}}\right\| = O_p(T^{-1/2})$ and $\left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}^0_{\ddot{\mathbf{x}}}\right\| = O_p((NT)^{-1/2})$, and also $\left\|T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\right\| = O_p(T^{-1/2})$, $\left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\right\| = O_p((NT)^{-1/2})$. By using the expansion in (3.16) and recognizing the fact that the terms involving $\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}$ from the left will either preserve the same order or bring it down similarly to (3.49), we obtain the following from the remaining

a and **b** terms:

$$\begin{aligned}
\|\mathbf{a}\| = \left\| T^{-1}\mathbf{V}_i'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \leq & \underbrace{\left\| T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0 \right\|}_{O_p(N^{-1/2})+O_p(T^{-1/2})} \left\| (T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^+ \right\| \underbrace{\left\| T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\|}_{O_p(T^{-1/2})} \\
& + \left\| T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& + \underbrace{\left\| T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\|}_{O_p(N^{-1})+O_p((NT)^{-1/2})} \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& + \left\| T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{x}} \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& + \left\| T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0 \right\| \left\| \left[(T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0)^+ - \mathbf{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^+ \right] \right\| \left\| T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& = O_p(T^{-1}) + O_p(N^{-1}) + O_p((NT)^{-1/2}),
\end{aligned} \tag{3.72}$$

and

$$\begin{aligned}
\|\mathbf{b}\| = \left\| T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \leq & \underbrace{\left\| T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0 \right\|}_{O_p(N^{-1/2})} \left\| (T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^+ \right\| \underbrace{\left\| T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\|}_{O_p(T^{-1/2})} \\
& + \left\| T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& + \underbrace{\left\| T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\|}_{O_p(N^{-1})} \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& + \left\| T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{x}} \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& + \left\| T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0 \right\| \left\| \left[(T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0)^+ - \mathbf{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^+ \right] \right\| \left\| T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& = O_p(T^{-1}) + O_p(N^{-1}) + O_p((NT)^{-1/2}),
\end{aligned} \tag{3.73}$$

with the drivers of the order indicated. In summary,

$$\begin{aligned}
\|\mathbf{B}\| &\leq \sqrt{N}\sup_i \left\| \left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i \right)^{-1} - \mathbf{\Sigma}_i^{-1} \right] \right\| \frac{1}{N}\sum_{i=1}^N \left\| T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
&= O_p(N^{-1}) + O_p(T^{-1/2}),
\end{aligned} \tag{3.74}$$

under $TN^{-1} = O(1)$ and so

$$\|\mathbf{II}\| = \left\| \frac{1}{\sqrt{N}}\sum_{i=1}^N \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i \right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| = O_p(N^{-1/2}) + O_p(T^{-1/2}), \tag{3.75}$$

which ultimately leads to

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) &= \frac{1}{\sqrt{N}}\sum_{i=1}^N \boldsymbol{\nu}_i + O_p(N^{-1/2}) + O_p(T^{-1/2}) \\
&\to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \mathbf{\Omega}_{\boldsymbol{\nu}}\right)
\end{aligned} \tag{3.76}$$

as $(N,T) \to \infty$ under $TN^{-1} = O(1)$.

We now let $m_x = g$, which means that we will again use the expansion in (3.41). Because now the convergence rate will be quicker, (3.58) will hold as well, therefore it is sufficient to check **II** in the expansion (3.45) and in particular we start with $\mathbf{A}_3$ as the analysis of $\mathbf{A}_1$ and $\mathbf{A}_2$ will be the same and these terms will be negligible. Hence,

$$\begin{aligned}
\|\mathbf{A}_3\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbf{\Sigma}_i^{-1} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' (\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}}) \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&\le \left\| \frac{1}{N} \sum_{i=1}^{N} \mathbf{\Sigma}_i^{-1} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&+ \left\| \frac{1}{N} \sum_{i=1}^{N} \mathbf{\Sigma}_i^{-1} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&+ \left\| \frac{1}{N} \sum_{i=1}^{N} \mathbf{\Sigma}_i^{-1} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' T^{-1} \mathbf{F}_{\mathbf{x}} \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&+ \left\| \frac{1}{N} \sum_{i=1}^{N} \mathbf{\Sigma}_i^{-1} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' T^{-1} \mathbf{F}_{\mathbf{x}} \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}} [(T^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}})^{+} - (\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}' T^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{x}} \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}})^{+}] \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}),
\end{aligned} \tag{3.77}$$

which is driven the highest order term which is almost identical to (3.43). Note that we still use the restriction $TN^{-1} = O(1)$ just as in case of $m_x < g$. We need this to show that **B** term is negligible as well, because **II** is generally not mean-zero. Particularly, the split

$$\begin{aligned}
\mathbf{II} &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (T^{-1} \mathbf{V}_i' \mathbf{M}_{\mathbf{F}_{\mathbf{x}}} \mathbf{V}_i)^{-1} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i \\
&+ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[\left(T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - (T^{-1} \mathbf{V}_i' \mathbf{M}_{\mathbf{F}_{\mathbf{x}}} \mathbf{V}_i)^{-1} \right] T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i
\end{aligned} \tag{3.78}$$

is not viable, because even if $\left\| \left(T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - (T^{-1} \mathbf{V}_i' \mathbf{M}_{\mathbf{F}_{\mathbf{x}}} \mathbf{V}_i)^{-1} \right\| = o_p(N^{-1/2})$ under $g = m_x$, the first component is not mean-zero, and its second moment is not negligible. Hence, we implement

$$\begin{aligned}
\mathbf{II} &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbf{\Sigma}_i^{-1} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i \\
&+ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[\left(T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - \mathbf{\Sigma}_i^{-1} \right] T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i = \mathbf{A} + \mathbf{B}
\end{aligned} \tag{3.79}$$

where $\|\mathbf{A}\| = o_p(1)$ still as $\left\| \mathbf{\Sigma}_i^{-1} \right\| = O(1)$ and deterministic. Under $m_x = g$ we have that

$$\sqrt{N} \left\| \left(T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - \mathbf{\Sigma}_i^{-1} \right\| = O_p(1) \tag{3.80}$$

under $TN^{-1} = O(1)$ uniformly, and hence

$$\|\mathbf{B}\| \le \sqrt{N} \sup_i \left\| \left(T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - \mathbf{\Sigma}_i^{-1} \right\| \frac{1}{N} \sum_{i=1}^{N} \left\| T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| = o_p(1). \tag{3.81}$$

3.3 Bootstrap Distributions

Theorem 6. *Under Assumptions 1 - 7, for either $m_x < g$ or $m_x = g$ as $(N, T) \to \infty$,*

(a) $\sup_{x \in \mathbb{R}^{k\times 1}} \left| \mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \leq x] - \mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \leq x] \right| \to_p 0,$

(b) $\sup_{x \in \mathbb{R}^{k\times 1}} \left| \mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}) \leq x] - \mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \leq x] \right| \to_p 0,$

where inequalities are to be interpreted coordinate wise.

Proof. (a) We assume $m_x < g$. Let $\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} = \mathbf{I}_T - \widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}(\widehat{\mathbf{F}}^{*\prime}_{\dot{\mathbf{x}}}\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}})^+\widehat{\mathbf{F}}^{*\prime}_{\dot{\mathbf{x}}}$ and $\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} = (\mathbf{I}_N \otimes \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}})$. We derive the CCEP estimator from the bootstrap sample:

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} &= \left(\mathbf{X}^{*\prime}\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}^*\right)^{-1}\mathbf{X}^{*\prime}\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{y}^* \\
&= \left(\mathbf{X}'\mathbf{W}_T'\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{W}_T\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}_T'\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{W}_T\mathbf{y} \\
&= \left(\mathbf{X}'\mathbf{W}_T'\mathbf{W}_T\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}_T'\mathbf{W}_T\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{y} \\
&= \left(\mathbf{X}'\text{diag}(\mathbf{s}\otimes\boldsymbol{\iota}_T')\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}\right)^{-1}\mathbf{X}'\text{diag}(\mathbf{s}\otimes\boldsymbol{\iota}_T')\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{y} \\
&= \left(\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*}\mathbf{y}_i \\
&= \boldsymbol{\beta} + \left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i + \frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right),
\end{aligned}
\tag{3.82}
$$

which implies that

$$
\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) = &\left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} \\
&\times\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right).
\end{aligned}
\tag{3.83}
$$

Next, we can write

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) &= \sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) - \sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \\
&= \left(\frac{1}{NT} \sum_{i=1}^N s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i - \frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i \right) \\
&+ \left(\frac{1}{NT} \sum_{i=1}^N s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i - \frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i \right) \\
&+ \left(\frac{1}{NT} \sum_{i=1}^N s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i - \frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i \right) \\
&+ \left[\left(\frac{1}{NT} \sum_{i=1}^N s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - \left(\frac{1}{NT} \sum_{i=1}^N \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \right] \\
&\times \left(\frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i + \frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i + \frac{1}{\sqrt{N}} \sum_{i=1}^N T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i \right).
\end{aligned} \tag{3.84}$$

In what follows, we will use the crucial lemma from Cheng and Huang (2010), which connects the rates of convergence in bootstrap and original (unconditional) probability measures. Particularly, given a vector valued statistic $\triangle_n$ which depends on $\mathbf{Z}_1, \dots \mathbf{Z}_n$ and multinomial weights $s_1, \dots, s_n$ (independent from model primitives), then for a deterministic sequence a_n we have

$$\triangle_n = O_{p^*}(a_n) \textit{ in probability} \Leftrightarrow \triangle_n = O_p(a_n) \textit{ unconditionally.}$$

Due to this result, we have

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) &= \left(\frac{1}{NT} \sum_{i=1}^{N} s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i \right) \\
&+ \left(\frac{1}{NT} \sum_{i=1}^{N} s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i \right) \\
&+ \left(\frac{1}{NT} \sum_{i=1}^{N} s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i \right) \\
&+ \underbrace{\left[\left(\frac{1}{NT} \sum_{i=1}^{N} s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \right]}_{o_{p^*}(1)} \\
&\times \underbrace{\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i + \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} + \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i \right)}_{O_{p^*}(1)} \\
&= \left(\frac{1}{NT} \sum_{i=1}^{N} s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i \right) \\
&+ \left(\frac{1}{NT} \sum_{i=1}^{N} s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i \right) \\
&+ \left(\frac{1}{NT} \sum_{i=1}^{N} s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i \right) \\
&+ o_{p^*}(1) \\
&= \mathbf{I} + \mathbf{II} + \mathbf{III} + o_{p^*}(1) \qquad (3.85)
\end{aligned}$$

in probability, where $\left\| \left(\frac{1}{NT} \sum_{i=1}^{N} s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \right\| = o_{p^*}(1)$ by Theorem 2 in De Vos and Stauskas (2024). By using the bootstrap consistency results from the same study,

$$\begin{aligned}
\|\mathbf{III}\| &\leq \left\| \left(\frac{1}{NT} \sum_{i=1}^{N} s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \right\| \left(\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i \right\| + \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \boldsymbol{\varepsilon}_i \right\| \right) \\
&= o_{p^*}(1) \qquad (3.86)
\end{aligned}$$

in probability and

$$\begin{aligned}
\mathbf{I} &= \left(\frac{1}{NT} \sum_{i=1}^{N} s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i - \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \boldsymbol{\nu}_i \right) \\
&= \boldsymbol{\Sigma}^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (s_i - 1) \boldsymbol{\Sigma}_i \boldsymbol{\nu}_i + O_{p^*}(T^{-1/2}) \\
&\to_{d^*} \mathcal{N}\left(\mathbf{0}_{k \times 1}, \boldsymbol{\Sigma}^{-1} \boldsymbol{\Psi}_{\nu} \boldsymbol{\Sigma}^{-1} \right) \qquad (3.87)
\end{aligned}$$

in probability. We are left with evaluating **II**. For this, we introduce the bootstrap rotation matrix

$$\overline{\mathbf{H}}_{w,\dot{\mathbf{x}}} = [\overline{\mathbf{H}}_{w,\dot{\mathbf{x}},m_x}, \overline{\mathbf{H}}_{w,\dot{\mathbf{x}},-m_x}] = \begin{bmatrix} \overline{\mathbf{\Gamma}}^{-1}_{w,\ddot{\mathbf{x}},m_x} & -\overline{\mathbf{\Gamma}}^{-1}_{w,\ddot{\mathbf{x}},m_x}\overline{\mathbf{\Gamma}}_{w,\ddot{\mathbf{x}}-m_x} \\ \mathbf{0}_{(g-m_x)\times m_x} & \mathbf{I}_{g-m_x} \end{bmatrix}, \qquad \mathbf{D}_N = \begin{bmatrix} \mathbf{I}_{m_x} & \mathbf{0}_{m_x\times(g-m_x)} \\ \mathbf{0}_{(g-m_x)\times m_x} & \sqrt{N}\mathbf{I}_{g-m_x} \end{bmatrix} \tag{3.88}$$

with its limiting matrix $\mathbf{H}_{\dot{\mathbf{x}}} = [\mathbf{H}_{\dot{\mathbf{x}},m_x}, \mathbf{H}_{\dot{\mathbf{x}},-m_x}] = \begin{bmatrix} \mathbf{\Gamma}^{-1}_{\ddot{\mathbf{x}},m_x} & -\mathbf{\Gamma}^{-1}_{\ddot{\mathbf{x}},m_x}\mathbf{\Gamma}_{\ddot{\mathbf{x}},-m_x} \\ \mathbf{0}_{(g-m_x)\times m_x} & \mathbf{I}_{g-m_x} \end{bmatrix}$ such that

$$\widehat{\mathbf{F}}^{0*}_{\dot{\mathbf{x}}} = \widehat{\mathbf{F}}^{*}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{w,\dot{\mathbf{x}}}\mathbf{D}_N = \mathbf{F}^0_{\dot{\mathbf{x}}} + [\overline{\mathbf{V}}_{w,\dot{\mathbf{x}}}\overline{\mathbf{H}}_{w,\dot{\mathbf{x}},m_x}, \sqrt{N}\overline{\mathbf{V}}_{w,\dot{\mathbf{x}}}\overline{\mathbf{H}}_{w,\dot{\mathbf{x}},-m_x}] = \mathbf{F}^0_{\dot{\mathbf{x}}} + [\overline{\mathbf{V}}^0_{w,\dot{\mathbf{x}},m_x}, \overline{\mathbf{V}}^0_{w,\ddot{\mathbf{x}},-m_x}]. \tag{3.89}$$

From now on, we can repeat exactly the same steps as in the analysis of **II** (and **IV**, which is now merged together) in the original sample by using independence of bootstrap weights from the model primitives, the rate conversion lemma of Cheng and Huang (2010) and a few key results, such as

$$(1)\quad \left\|\overline{\mathbf{V}}_{w,\ddot{\mathbf{x}}}\right\| = O_{p^*}(N^{-1/2}), \tag{3.90}$$

$$(2)\quad \left\|T^{-1}\overline{\mathbf{V}}'_{w,\ddot{\mathbf{x}}}\mathbf{V}_i\right\| = O_{p^*}(N^{-1}) + O_{p^*}((NT)^{-1/2}), \tag{3.91}$$

$$(3)\quad \left\|\left(T^{-1}\widehat{\mathbf{F}}^{0*\prime}_{\dot{\mathbf{x}}}\widehat{\mathbf{F}}^{0*}_{\dot{\mathbf{x}}}\right)^{+} - \mathbf{\Sigma}^{+}_{w,\mathbf{F}^0_{\dot{\mathbf{x}},v}}\right\| = O_{p^*}(N^{-1/2}) + O_{p^*}(T^{-1/2}), \tag{3.92}$$

$$(4)\quad \mathbb{E}(s_i) = 1, \tag{3.93}$$

$$(5)\quad \mathrm{V}ar(s_i) = \mathbb{E}[(s_i-1)^2] = 1 - N^{-1} \quad \textit{(multinomial variance)} \tag{3.94}$$

where

$$\mathbf{\Sigma}_{w,\mathbf{F}^0_{\dot{\mathbf{x}}},v} = \operatorname{diag}\left[\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}, (T^{-1}\overline{\mathbf{V}}^{0\prime}_{w,\dot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{w,\dot{\mathbf{x}},-m_x})\right]. \tag{3.95}$$

Therefore,

$$\begin{aligned}
\|\mathbf{II}\| &= \left\|\left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i - \frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right)\right\| \\
&\le \left\|\left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\|\left\|\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i - \frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right)\right\| \\
&\le \underbrace{\left\|\left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\|}_{O_{p^*}(1)}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|\right) \\
&= O_{p^*}(N^{-1/2}) + O_{p^*}(T^{-1/2}).
\end{aligned} \tag{3.96}$$

Note how (3.94) ensures that whenever we analyze mean-square convergence, we will obtain the expectation of the square of the main object of analysis, plus a lower order term, hence the limits will stay the same. Hence,

$$\begin{aligned}
\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| &\le \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N (s_i-1)T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&= O_{p^*}(N^{-1/2}) + O_{p^*}(T^{-1/2}).
\end{aligned} \tag{3.97}$$

In summary, we obtain

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) = \boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(s_i - 1)\boldsymbol{\Sigma}_i\boldsymbol{\nu}_i + o_{p^*}(1)$$
$$\to_{d^*} \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1}\right) \tag{3.98}$$

as $(N, T) \to \infty$ in probability. The consistency holds uniformly by multivariate Polya's Theorem, similarly to the argument in Gonçalves and Perron (2014). The latter states that when $\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \to_d \mathcal{N}(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1})$ (proven in Theorem 1), then

$$\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}(\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \le x) - \boldsymbol{\Phi}(x; \mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1})\right| \to 0,$$

where $\boldsymbol{\Phi}(x; \boldsymbol{\mu}, \boldsymbol{\Omega})$ is the Gaussian CDF with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\Omega}$. Hence, uniformity follows if also

$$\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}^*(\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \le x) - \boldsymbol{\Phi}(x; \mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1})\right| \to_p 0$$

which is in turn guaranteed by Polya's Theorem because (3.98) holds in probability. Hence, uniform consistency follows:

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \le x] - \mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \le x]\right| \\
&= \sup_{x\in\mathbb{R}^{k\times 1}}\left| \left(\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \le x] - \boldsymbol{\Phi}(x; \mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1}\right)\right. \\
&- \left.\left(\mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \le x] - \boldsymbol{\Phi}(x; \mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1})\right) \right| \\
&\le \sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \le x] - \boldsymbol{\Phi}(x; \mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1}\right| \\
&+ \sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \le x] - \boldsymbol{\Phi}(x; \mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1})\right| \\
&= o_p(1),
\end{aligned} \tag{3.99}$$

which completes the proof.

The argument for $m_x = g$ is exact the same as in the discussion of Theorem 3.

(b) The bootstrap CCEMG estimator is given by

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} &= \frac{1}{N}\sum_{i=1}^{N} s_i\left(\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&= \frac{1}{N}\sum_{i=1}^{N} s_i\boldsymbol{\beta}_i + \frac{1}{N}\sum_{i=1}^{N} s_i\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + \frac{1}{N}\sum_{i=1}^{N} s_i\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \\
&= \boldsymbol{\beta}\frac{1}{N}\underbrace{\sum_{i=1}^{N} s_i}_{N} + \frac{1}{N}\sum_{i=1}^{N} s_i\boldsymbol{\nu}_i + \frac{1}{N}\sum_{i=1}^{N} s_i\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\
&+ \frac{1}{N}\sum_{i=1}^{N} s_i\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i,
\end{aligned} \tag{3.100}$$

hence

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \boldsymbol{\nu}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i$$
$$+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i, \tag{3.101}$$

and so

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}) &= \sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) - \sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(s_i - 1)\boldsymbol{\nu}_i \\
&+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right) \\
&+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right) \\
&+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right] \\
&\times \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right) \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}(s_i - 1)\boldsymbol{\nu}_i \\
&+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right) \\
&+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right) + o_{p^*}(1) \\
&= \mathbf{I} + \mathbf{II} + \mathbf{III} + o_{p^*}(1)
\end{aligned} \tag{3.102}$$

in probability, because

$$\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right]\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right)\right\| \\
&\leq \sqrt{N}\sup_i \left\|\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\| \frac{1}{N}\sum_{i=1}^{N}|s_i|\left(\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\| + \left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|\right) \\
&= \sqrt{N}\sup_i \left\|\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\| \frac{1}{N}\sum_{i=1}^{N}|s_i|\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\| \\
&+ \sqrt{N}\sup_i \left\|\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\| \left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&= o_{p^*}(1)
\end{aligned} \tag{3.103}$$

as $TN^{-1} = O(1)$ in analogy to (3.70). Then

$$\begin{aligned}
\|\mathbf{III}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{X}_i \right)^{-1} \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\boldsymbol{\varepsilon}_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \right) \right\| \\
&\leq \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{X}_i \right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\boldsymbol{\varepsilon}_i \right\| + \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{X}_i \right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \right\| \\
&= o_{p^*}(1) \qquad (3.104)
\end{aligned}$$

in analogy to (3.48) by using the fact that bootstrap weights are independent from the model primitives and the results in (3.90) - (3.92). Further,

$$\begin{aligned}
\mathbf{II} &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{X}_i \right)^{-1} \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{F}_{\mathbf{y}}\gamma_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right) \\
&= \underbrace{\frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \boldsymbol{\Sigma}_i^{-1} \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{F}_{\mathbf{y}}\gamma_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right)}_{o_{p^*}(1) \text{ in analogy to (3.69)}} \\
&\quad + \underbrace{\frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{X}_i \right)^{-1} - \boldsymbol{\Sigma}_i^{-1} \right] \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{F}_{\mathbf{y}}\gamma_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right)}_{o_{p^*}(1) \text{ in analogy to (3.70)}} \\
&= o_{p^*}(1) \qquad (3.105)
\end{aligned}$$

under $TN^{-1} = O(1)$ by using the independence of the bootstrap weights from the model primitives. Eventually,

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}) &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (s_i - 1)\boldsymbol{\nu}_i + o_{p^*}(1) \\
&\rightarrow_{d^*} \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Omega}_{\boldsymbol{\nu}}\right) \qquad (3.106)
\end{aligned}$$

as $(N, T) \to \infty$ in probability. Similarly to part (a), consistency holds uniformly by multivariate Polya's Theorem. We have

$$\sup_{x \in \mathbb{R}^{k\times 1}} \left| \mathbb{P}(\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \leq x) - \boldsymbol{\Phi}(x; \mathbf{0}_{k\times 1}, \boldsymbol{\Omega}_{\boldsymbol{\nu}}) \right| \to 0.$$

Hence, uniformity follows if also

$$\sup_{x \in \mathbb{R}^{k\times 1}} \left| \mathbb{P}^*(\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}) \leq x) - \boldsymbol{\Phi}(x; \mathbf{0}_{k\times 1}, \boldsymbol{\Omega}_{\boldsymbol{\nu}}) \right| \to_p 0$$

which is in turn guaranteed by Polya's Theorem because (3.106) holds in probability. Hence, uniform consistency follows:

$$
\begin{aligned}
&\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}}-\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})\leq x]-\mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}-\boldsymbol{\beta})\leq x]\right| \\
&=\sup_{x\in\mathbb{R}^{k\times 1}}\Big|\left(\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}}-\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})\leq x]-\boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1},\boldsymbol{\Omega}_{\boldsymbol{\nu}}\right) \\
&-\left(\mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}-\boldsymbol{\beta})\leq x]-\boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1},\boldsymbol{\Omega}_{\boldsymbol{\nu}})\right)\Big| \\
&\leq\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}}-\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})\leq x]-\boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1},\boldsymbol{\Omega}_{\boldsymbol{\nu}})\right| \\
&+\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}-\boldsymbol{\beta})\leq x]-\boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1},\boldsymbol{\Omega}_{\nu})\right| \\
&=o_p(1), \qquad (3.107)
\end{aligned}
$$

which completes the proof.

The argument for $m_x=g$ is exact the same as in the discussion of Theorem 4.

4 Variance Estimators

Theorem 5. *Under Assumptions 1 - 7, for either $m_x<g$ or $m_x=g$ as $(N,T)\to\infty$*

(a) $N\widehat{\boldsymbol{\Theta}}_{CCEP,\dot{\mathbf{x}}}\to_p\boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1}$

(b) $N\widehat{\boldsymbol{\Theta}}_{CCEMG,\dot{\mathbf{x}}}\to_p\boldsymbol{\Omega}_{\nu}$.

Proof. (a) The proofs for either $m_x<g$ or $m_x=g$ are identical as in the latter case the remainder will be of even lower order. Let $\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}=T^{-1}\mathbf{X}_i\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i$. We firstly find the workable expression of $\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i}-\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})$. Notice how

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i}-\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}&=\widehat{\mathbf{Q}}^{-1}_{\dot{\mathbf{x}},i}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i-\frac{1}{N}\sum_{i=1}^N\widehat{\mathbf{Q}}^{-1}_{\dot{\mathbf{x}},i}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&=\boldsymbol{\nu}_i-\frac{1}{N}\sum_{i=1}^N\boldsymbol{\nu}_i+\widehat{\mathbf{Q}}^{-1}_{\dot{\mathbf{x}},i}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i+T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right) \\
&-\frac{1}{N}\sum_{i=1}^N\widehat{\mathbf{Q}}^{-1}_{\dot{\mathbf{x}},i}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i+T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right) \\
&=\boldsymbol{\nu}_i+o_p(1), \qquad (4.1)
\end{aligned}
$$

because $\frac{1}{N}\sum_{i=1}^N\boldsymbol{\nu}_i=O_p(N^{-1/2})$, $\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\|=o_p(1)$ and $\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|=o_p(1)$, which come directly from (3.49) and (3.71), respectively. Also,

$$
\begin{aligned}
&\left\|\frac{1}{N}\sum_{i=1}^N\widehat{\mathbf{Q}}^{-1}_{\dot{\mathbf{x}},i}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i+T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right)\right\| \\
&\leq\sup_i\left\|\widehat{\mathbf{Q}}^{-1}_{\dot{\mathbf{x}},i}\right\|\frac{1}{N}\sum_{i=1}^N\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|+\sup_i\left\|\widehat{\mathbf{Q}}^{-1}_{\dot{\mathbf{x}},i}\right\|\frac{1}{N}\sum_{i=1}^N\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\| \\
&=o_p(1). \qquad (4.2)
\end{aligned}
$$

Therefore, because $\left\|\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}\right\| = O_p(1)$, we have that

$$\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}) = \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}\boldsymbol{\nu}_i + o_p(1). \tag{4.3}$$

By using this, we obtain

$$\begin{aligned}
N\widehat{\boldsymbol{\Theta}}_{CCEP,\dot{\mathbf{x}}} &= N\left[\left(\frac{1}{N}\sum_{i=1}^{N}\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}\right)^{-1}\frac{1}{N(N-1)}\sum_{i=1}^{N}\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})'\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}\left(\frac{1}{N}\sum_{i=1}^{N}\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}\right)^{-1}\right] \\
&= \left(\frac{1}{N}\sum_{i=1}^{N}\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}\right)^{-1}\frac{1}{N-1}\sum_{i=1}^{N}\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}\boldsymbol{\nu}_i\boldsymbol{\nu}_i'\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}\left(\frac{1}{N}\sum_{i=1}^{N}\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}\right)^{-1} + o_p(1) \\
&= \left(\frac{1}{N}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'\mathbf{V}_i\right)^{-1}\frac{1}{N-1}\sum_{i=1}^{N}(T^{-1}\mathbf{V}_i'\mathbf{V}_i)\boldsymbol{\nu}_i\boldsymbol{\nu}_i'(T^{-1}\mathbf{V}_i'\mathbf{V}_i)\left(\frac{1}{N}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'\mathbf{V}_i\right)^{-1} + o_p(1) \\
&\to_p \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1}
\end{aligned} \tag{4.4}$$

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

(b) The result comes immediately from (4.1):

$$\begin{aligned}
N\widehat{\boldsymbol{\Theta}}_{CCEMG,\dot{\mathbf{x}}} &= \frac{1}{N-1}\sum_{i=1}^{N}(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})' \\
&= \frac{1}{N-1}\sum_{i=1}^{N}\boldsymbol{\nu}_i\boldsymbol{\nu}_i' + o_p(1) \\
&\to_p \boldsymbol{\Omega}_{\nu}
\end{aligned} \tag{4.5}$$

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

Theorem 7. *Under Assumptions 1 - 7, for either $m_x < g$ or $m_x = g$ as $(N,T)\to\infty$*

(a) $N\widehat{\boldsymbol{\Theta}}^{*}_{CCEP,\dot{\mathbf{x}}} \to_{p^*} \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1}$

(b) $N\widehat{\boldsymbol{\Theta}}^{*}_{CCEMG,\dot{\mathbf{x}}} \to_{p^*} \boldsymbol{\Omega}_{\nu}$.

Proof. (a) The proofs for either $m_x < g$ or $m_x = g$ are again identical since in the latter case the remainder

will be of even lower order in bootstrap probability measure. Generally, the proof follows Theorem 5 closely. Let $\widehat{\mathbf{Q}}^*_{\dot{\mathbf{x}},i} = T^{-1}\mathbf{X}_i\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i$. The first part of the workable expression of $\widehat{\mathbf{Q}}^*_{\dot{\mathbf{x}},i}(\widehat{\boldsymbol{\beta}}^*_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}})$ is given by

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}^*_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} &= \widehat{\mathbf{Q}}^{*-1}_{\dot{\mathbf{x}},i}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{y}_i - \frac{1}{N}\sum_{i=1}^{N}s_i\widehat{\mathbf{Q}}^{*-1}_{\dot{\mathbf{x}},i}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&= \boldsymbol{\nu}_i - \frac{1}{N}\sum_{i=1}^{N}s_i\boldsymbol{\nu}_i + \widehat{\mathbf{Q}}^{*-1}_{\dot{\mathbf{x}},i}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right) \\
&\quad - \frac{1}{N}\sum_{i=1}^{N}s_i\widehat{\mathbf{Q}}^{*-1}_{\dot{\mathbf{x}},i}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right) \\
&= \boldsymbol{\nu}_i + o_{p^*}(1),
\end{aligned} \tag{4.6}$$

since $\frac{1}{N}\sum_{i=1}^{N} s_i \boldsymbol{\nu}_i = O_{p^*}(N^{-1/2})$, $\left\| T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\boldsymbol{\varepsilon}_i \right\| = o_{p^*}(1)$ and $\left\| T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| = o_{p^*}(1)$, which come from the proof of Theorem 6. Also,

$$\begin{aligned}
&\left\| \frac{1}{N}\sum_{i=1}^{N} s_i \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*-1}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{F}_{\mathbf{y}}\gamma_i + T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\boldsymbol{\varepsilon}_i \right) \right\| \\
&\leq \sup_i \left\| \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*-1} \right\| \frac{1}{N}\sum_{i=1}^{N} |s_i| \left\| T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| + \sup_i \left\| \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*-1} \right\| \frac{1}{N}\sum_{i=1}^{N} |s_i| \left\| T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\boldsymbol{\varepsilon}_i \right\| \\
&= o_{p^*}(1). \qquad (4.7)
\end{aligned}$$

Therefore, because $\left\| \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*} \right\| = O_{p^*}(1)$, we have that

$$\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*}(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i}^{*} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}^{*}) = \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*}\boldsymbol{\nu}_i + o_p(1). \qquad (4.8)$$

Based on these arguments, we again obtain

$$\begin{aligned}
&N\widehat{\boldsymbol{\Theta}}_{CCEP,\dot{\mathbf{x}}}^{*} \\
&= N\left[\left(\frac{1}{N}\sum_{i=1}^{N} s_i \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*} \right)^{-1} \frac{1}{N(N-1)}\sum_{i=1}^{N} s_i \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*}(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i}^{*} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}^{*})(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i}^{*} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}^{*})'\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*} \left(\frac{1}{N}\sum_{i=1}^{N} s_i \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*} \right)^{-1} \right] \\
&= \left(\frac{1}{N}\sum_{i=1}^{N} s_i \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*} \right)^{-1} \frac{1}{N-1}\sum_{i=1}^{N} s_i \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*}\boldsymbol{\nu}_i\boldsymbol{\nu}_i'\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*} \left(\frac{1}{N}\sum_{i=1}^{N} s_i \widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{*} \right)^{-1} + o_{p^*}(1) \\
&= \left(\frac{1}{N}\sum_{i=1}^{N} T^{-1}s_i \mathbf{V}_i'\mathbf{V}_i \right)^{-1} \frac{1}{N-1}\sum_{i=1}^{N} s_i (T^{-1}\mathbf{V}_i'\mathbf{V}_i)\boldsymbol{\nu}_i\boldsymbol{\nu}_i'(T^{-1}\mathbf{V}_i'\mathbf{V}_i) \left(\frac{1}{N}\sum_{i=1}^{N} s_i T^{-1}\mathbf{V}_i'\mathbf{V}_i \right)^{-1} + o_{p^*}(1) \\
&\rightarrow_{p^*} \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1} \qquad (4.9)
\end{aligned}$$

as $(N,T) \rightarrow \infty$.

(b) Similarly to Theorem 5, the result comes immediately from (4.6):

$$\begin{aligned}
N\widehat{\boldsymbol{\Theta}}_{CCEMG,\dot{\mathbf{x}}}^{*} &= \frac{1}{N-1}\sum_{i=1}^{N} s_i (\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i}^{*} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}^{*})(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i}^{*} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}^{*})' \\
&= \frac{1}{N-1}\sum_{i=1}^{N} s_i \boldsymbol{\nu}_i\boldsymbol{\nu}_i' + o_{p^*}(1) \\
&\rightarrow_{p^*} \boldsymbol{\Omega}_{\nu} \qquad (4.10)
\end{aligned}$$

as $(N,T) \rightarrow \infty$.

5 Discussion on General Unknown Factors

Proposition 1. *Under Assumptions 1 - 7 for $m_x < g$ as $(N,T) \rightarrow \infty$ with $TN^{-1} \rightarrow \tau > 0$, plus a covariance stationary $\mathbf{F}_{\mathbf{y}}$, we have the following asymptotic representations:*

(a) (*heterogeneous case*) $\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) = \boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i\boldsymbol{\nu}_i + o_p(1),$

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\nu}_i + o_p(1).$$

If, in addition, $\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},\mathbf{y}}}$ is deterministic, then

(b) (*homogeneous case*) $\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta})$

$$= \boldsymbol{\Sigma}^{-1}\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[T^{-1/2}\mathbf{V}_i'\boldsymbol{\varepsilon}_i + \boldsymbol{\Theta}_i\mathbf{D}_{T,F,m}^{-1}\text{vec}(\mathbf{V}_i'\mathbf{F})\right] + \sqrt{\tau}\mathbf{h}_1(\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{xy}}}) + \mathbf{h}_2\right) + o_p(1),$$

where $\boldsymbol{\Theta}_i$ is a random matrix that is a function of loadings. Also, $\mathbf{h}_1$ and $\mathbf{h}_2$ are equivalents of the respective terms in Theorem 1.

Proof. (a) We begin with the CCEP estimator under heterogeneous slopes, where we use the expansion

$$\begin{aligned}\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) &= \left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i \\ &+ \left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \\ &+ \left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\ &= \mathbf{I} + \mathbf{II} + \mathbf{III},\end{aligned} \tag{5.1}$$

which can be simplified by using results from Stauskas (2022), where general factors under heterogeneous slopes were explored under common $\mathbf{F}$. Particularly, by using Theorem 2 in the latter study, we obtain

$$\mathbf{I} = \boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i\boldsymbol{\nu}_i + o_p(1) \tag{5.2}$$

and

$$\|\mathbf{II}\| = o_p(1), \tag{5.3}$$

where the negligible terms are of the same or lower order because $\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} \subset \widehat{\mathbf{F}}$ is independent form $\boldsymbol{\varepsilon}_i$ for all i. By the same Theorem 2,

$$\left\|\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\| = O_p(1), \tag{5.4}$$

therefore

$$\begin{aligned}\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) &= \boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i\boldsymbol{\nu}_i \\ &+ \underbrace{\left(\frac{1}{NT}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}}_{O_p(1)}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + o_p(1).\end{aligned} \tag{5.5}$$

In what remains, we will focus on the numerator of **III**, which can be further decomposed into

$$\begin{aligned}
\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\
&-\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}-\mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}})\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{F}_{\mathbf{y}}\gamma_i-\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{P}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\
&-\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}-\mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}})\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \mathbf{A}-\mathbf{B}-\mathbf{C}, \qquad (5.6)
\end{aligned}$$

where clearly,

$$\|\mathbf{A}\| \le \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| = O_p(T^{-1/2})$$

as it follows (3.10) since $\mathbf{F}_{\mathbf{y}}$ is assumed to be covariance stationary. Moving on to **B**, we obtain

$$\begin{aligned}
\|\mathbf{B}\| &= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{P}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{P}_{\mathbf{F}_{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\le \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'\mathbf{P}_{\mathbf{F}_{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{P}_{\mathbf{F}_{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\le T^{-1/2}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\gamma_i'\otimes\mathbf{V}_i'\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1}\right)\right\|\left\|(\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+}\right\|\left\|\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\mathbf{D}_{T,\mathbf{y}}^{-1}\right\| \\
&+ T^{-1/2}\left\|\frac{1}{N}\sum_{i=1}^{N}(\gamma_i\otimes\mathbf{\Gamma}_i)\right\|\left\|\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\sqrt{N}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1}\right\|\left\|(\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+}\right\|\left\|\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\mathbf{D}_{T,\mathbf{y}}^{-1}\right\| \\
&= O_p(T^{-1/2}), \qquad (5.7)
\end{aligned}$$

because $T^{-1/2}\mathbf{F}_{\mathbf{y}} = \mathbf{F}_{\mathbf{y}}\mathbf{I}_{m_y}T^{-1/2} = \mathbf{F}_{\mathbf{y}}\mathbf{D}_{T,\mathbf{y}}^{-1}$. We finally move to **C**, where we use the expansion adapted from Westerlund (2018):

$$\begin{aligned}
\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}-\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}} = \mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}-\mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}} &= T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^{+}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}+\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+}\mathbf{D}_{T,\mathbf{x}}^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \\
&+\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x}\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+}\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'+\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+}\mathbf{D}_{T,\mathbf{x}}^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \\
&+\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}\mathbf{D}_{T,\widehat{x}}^{-1}\left[(\mathbf{D}_{T,\widehat{x}}^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}\mathbf{D}_{T,\widehat{x}}^{-1})^{+}-\mathbf{\Sigma}^{+}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right]\mathbf{D}_{T,\widehat{x}}^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \\
&+\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1}(\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+}-(\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}'\mathbf{F}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+})\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}', \qquad (5.8)
\end{aligned}$$

where

$$\mathbf{D}_{T,\hat{\mathbf{x}}} = \text{diag}\left[\mathbf{D}_{T,\mathbf{x}}, \sqrt{T}\mathbf{I}_{k-m_x}\right], \tag{5.9}$$

which is needed to handle the $k > m_x$ case, and

$$\left\|(\mathbf{D}_{T,\hat{\mathbf{x}}}^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0}\mathbf{D}_{T,\hat{\mathbf{x}}}^{-1})^{+} - \mathbf{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^{0}}^{+}\right\| = O_p(N^{-1/2}) + O_p(T^{-\kappa/2}), \tag{5.10}$$

which now takes into account the general factors. We split $\mathbf{C}$ into

$$\begin{aligned}
\mathbf{C} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^{0}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0}})\mathbf{F}_{\mathbf{y}}\gamma_i &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^{0}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0}})\mathbf{F}_{\mathbf{y}}\gamma_i \\
&\quad - \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^{0}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0}})\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \mathbf{C}_1 - \mathbf{C}_2,
\end{aligned} \tag{5.11}$$

which we split further according to the components of (5.8). Hence,

$$\begin{aligned}
\|\mathbf{C}_{11}\| &= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i' T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i' T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma\right\| \\
&+ \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i' T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\| \\
&\leq \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}\right\| \left\|(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\right\| \left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma\right\| \\
&+ \left\|vec\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i' T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right)\right\| \\
&= \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}\right\| \left\|(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\right\| \left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma\right\| \\
&+ \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\eta}_{\gamma,i}' \otimes T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}\right) vec\left((T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\right)\right\| \\
&\leq \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}\right\| \left\|(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\right\| \left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma\right\| \\
&+ \frac{1}{\sqrt{T}}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\eta}_{\gamma,i}' \otimes T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}\right)\right\| \left\|(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+}\right\| \left\|T^{-1/2}\mathbf{F}_{\mathbf{y}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}\right\| \\
&= O_p(T^{-1/2}),
\end{aligned} \tag{5.12}$$

which is driven by the first additive term, because

$$\frac{1}{\sqrt{T}}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\boldsymbol{\eta}_{\gamma,i}' \otimes T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0}\right)\right\| = O_p(T^{-1}) + O_p((NT)^{-1/2}). \tag{5.13}$$

Next up

$$\begin{aligned}
\|\mathbf{C}_{12}\| &= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0}\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+}\mathbf{D}_{T,\mathbf{x}}^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\leq \frac{1}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0}\right\| \left\|\sqrt{T}\mathbf{D}_{T,\mathbf{x}}^{-1}\right\|^2 \left\|\sqrt{N}T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \|\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}\| \\
&= O_p(N^{-1}T^{-1/2}) + O_p(N^{-1/2}T^{-1})
\end{aligned} \tag{5.14}$$

and

$$
\begin{aligned}
\|\mathbf{C}_{13}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{V}_i' \overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x} \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{\Sigma}_{\mathbf{F}_\mathbf{x}}^{+} \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_\mathbf{x}' \mathbf{F}_\mathbf{y} \gamma_i \right\| \\
&\leq \frac{1}{N} \sum_{i=1}^{N} \left\| \sqrt{N} T^{-1} \mathbf{V}_i' \overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},m_x} \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \|\mathbf{\Sigma}_{\mathbf{F}_\mathbf{x}}\| \left\| \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_\mathbf{x}' \mathbf{F}_\mathbf{y} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \|\gamma_i\| \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}), \qquad (5.15)
\end{aligned}
$$

and

$$
\begin{aligned}
\|\mathbf{C}_{14}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{V}_i' \mathbf{F}_\mathbf{x} \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{\Sigma}_{\mathbf{F}_\mathbf{x}}^{+} \mathbf{D}_{T,\mathbf{x}}^{-1} \overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \mathbf{F}_\mathbf{y} \gamma_i \right\| \\
&\leq \frac{1}{T} \frac{1}{N} \sum_{i=1}^{N} \left\| \mathbf{V}_i' \mathbf{F}_\mathbf{x} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \sqrt{N} T^{-1/2} \overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \mathbf{F}_\mathbf{y} \gamma_i \right\| \left\| \mathbf{\Sigma}_{\mathbf{F}_\mathbf{x}}^{+} \right\| \\
&= O_p(T^{-1}). \qquad (5.16)
\end{aligned}
$$

Eventually, we obtain the following:

$$
\begin{aligned}
\|\mathbf{C}_{15}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \left[(\mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1})^{+} - \mathbf{\Sigma}^{+}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}} \right] \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \mathbf{F}_\mathbf{y} \gamma_i \right\| \\
&\leq \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \left[(\mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1})^{+} - \mathbf{\Sigma}^{+}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}} \right] \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \mathbf{F}_\mathbf{y} \gamma \right\| \\
&+ \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \left[(\mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1})^{+} - \mathbf{\Sigma}^{+}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}} \right] \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \mathbf{F}_\mathbf{y} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&\leq \left\| T^{-1/2} \sqrt{N} \overline{\mathbf{V}}' \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right\| \left\| (\mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1})^{+} - \mathbf{\Sigma}^{+}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}} \right\| \left\| \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \mathbf{F}_\mathbf{y} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \|\gamma\| \\
&+ \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\boldsymbol{\eta}_{\gamma,i}' \otimes T^{-1/2} \mathbf{V}_i' \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right) \right\| \left\| (\mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1})^{+} - \mathbf{\Sigma}^{+}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}} \right\| \left\| \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \mathbf{F}_\mathbf{y} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \\
&= O_p(N^{-1/2}) + O_p(T^{-\kappa/2}), \qquad (5.17)
\end{aligned}
$$

which is driven by the first additive term, because

$$
\begin{aligned}
\left\| T^{-1/2} \sqrt{N} \overline{\mathbf{V}}' \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right\| &\leq \left\| \sqrt{N} T^{-1} \overline{\mathbf{V}}' \overline{\mathbf{V}}^0_{\ddot{\mathbf{x}}} \right\| \left\| \sqrt{T} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right\| + T^{-1/2} \left\| \sqrt{N} \overline{\mathbf{V}}' \mathbf{F}^0_{\mathbf{x}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right\| \\
&= O_p(1) \qquad (5.18)
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E} \left(\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\boldsymbol{\eta}_{\gamma,i}' \otimes T^{-1/2} \mathbf{V}_i' \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right) \right\|^2 \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \operatorname{tr} \left[\mathbb{E} \left(\boldsymbol{\eta}_{\gamma,i}' \boldsymbol{\eta}_{\gamma,j} \right) \otimes \mathbb{E} \left(T^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \mathbf{V}_j \right) \right] \\
&= \frac{1}{N} \sum_{i=1}^{N} \operatorname{tr} \left[\mathbb{E} \left(\boldsymbol{\eta}_{\gamma,i}' \boldsymbol{\eta}_{\gamma,i} \right) \otimes \mathbb{E} \left(T^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \mathbf{V}_i \right) \right] \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left(\boldsymbol{\eta}_{\gamma,i}' \boldsymbol{\eta}_{\gamma,i} \right) \operatorname{tr} \left[\mathbb{E} \left(T^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}} \mathbf{V}_i \right) \right] = O(1). \qquad (5.19)
\end{aligned}
$$

Finally,

$$\begin{aligned}
\|\mathbf{C}_{16}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{V}_i' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} (\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} - (\mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}' \mathbf{F} \mathbf{D}_{T,\mathbf{x}}^{-1})^{+}) \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&\leq \frac{1}{\sqrt{T}} \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\gamma_i' \otimes \mathbf{V}_i' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \right) \right\| \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} - (\mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}' \mathbf{F} \mathbf{D}_{T,\mathbf{x}}^{-1})^{+} \right\| \left\| \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \\
&= O_p(T^{-(\kappa+1)/2}), \qquad (5.20)
\end{aligned}$$

which follows from

$$\begin{aligned}
\mathbb{E}\left(\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\gamma_i' \otimes \mathbf{V}_i' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \right) \right\|^2 \right) &= \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{E}\left(\operatorname{tr}\left[\gamma_i' \gamma_j \otimes \mathbf{V}_i' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{V}_j \right] \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\left(\operatorname{tr}\left[\gamma_i' \gamma_i \otimes \mathbf{V}_i' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{V}_i \right] \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\left(\gamma_i' \gamma_i \right) \operatorname{tr}\left[\mathbb{E}\left(\mathbf{V}_i' \mathbf{F}_{\mathbf{x}}' \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{V}_i \right) \right] \\
&= O(1). \qquad (5.21)
\end{aligned}$$

We next, move on to $\mathbf{C}_2$ and split it further according to (5.8), as well. Therefore,

$$\begin{aligned}
\|\mathbf{C}_{21}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0} (T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&\leq \frac{1}{\sqrt{T}} \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| \sqrt{N} T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0} \right\| \left\| (T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0})^{+} \right\| \left\| T^{-1/2} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&= O_p(T^{-1/2}). \qquad (5.22)
\end{aligned}$$

Next,

$$\begin{aligned}
\|\mathbf{C}_{22}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0} \mathbf{D}_{T,\mathbf{x}}^{-1} \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} \mathbf{D}_{T,\mathbf{x}}^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&\leq \frac{1}{N\sqrt{T}} \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| N T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0} \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\|^2 \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} \right\| \left\| T^{-1/2} \sqrt{N} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&= O_p(N^{-1} T^{-1/2}), \qquad (5.23)
\end{aligned}$$

and

$$\begin{aligned}
\|\mathbf{C}_{23}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0} \mathbf{D}_{T,\mathbf{x}}^{-1} \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&\leq \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| T^{-1} \sqrt{N} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0} \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} \right\| \left\| \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \|\gamma_i\| \\
&= O_p(N^{-1/2}) \qquad (5.24)
\end{aligned}$$

with

$$\|\mathbf{C}_{24}\| = \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{\Gamma}_i' \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} \mathbf{D}_{T,\mathbf{x}}^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\|$$

$$\leq \frac{1}{T} \frac{1}{N} \sum_{i=1}^{N} \left\| \mathbf{\Gamma}_i' \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| \sqrt{N} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| T^{-1/2} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\|$$

$$= O_p(N^{-1/2} T^{-1}) \tag{5.25}$$

and

$$\|\mathbf{C}_{25}\| = \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{\Gamma}_i' \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \left[(\mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1})^{+} - \mathbf{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^{0}}^{+} \right] \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\|$$

$$\leq \frac{1}{N} \sum_{i=1}^{N} \left\| \mathbf{\Gamma}_i' \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| \sqrt{N} T^{-1/2} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right\| \left\| (\mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1})^{+} - \mathbf{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^{0}}^{+} \right\|$$

$$\times \left\| \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime} \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \|\gamma_i\|$$

$$= O_p(N^{-1/2}) + O_p(T^{-\kappa/2}). \tag{5.26}$$

Finally,

$$\|\mathbf{C}_{26}\| = \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{\Gamma}_i' \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} (\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} - (\mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}' \mathbf{F} \mathbf{D}_{T,\mathbf{x}}^{-1})^{+}) \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \gamma_i \right\|$$

$$\leq \frac{1}{\sqrt{T}} \frac{1}{N} \sum_{i=1}^{N} \left\| \mathbf{\Gamma}_i' \overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| \sqrt{N} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} - (\mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}' \mathbf{F} \mathbf{D}_{T,\mathbf{x}}^{-1})^{+} \right\|$$

$$\times \left\| \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \|\gamma_i\|$$

$$= O_p(T^{-(\kappa+1)/2}). \tag{5.27}$$

Overall, we have

$$\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| = O_p(N^{-1/2}) + O_p(T^{-\kappa/2}), \tag{5.28}$$

and so

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) = \mathbf{\Sigma}^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbf{\Sigma}_i \boldsymbol{\nu}_i$$

$$+ \underbrace{\left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1}}_{O_p(1)} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i + o_p(1)$$

$$= \mathbf{\Sigma}^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \mathbf{\Sigma}_i \boldsymbol{\nu}_i + o_p(1), \tag{5.29}$$

which completes the proof.

Clearly, the same result holds under $m_x = g$. The subtle and important difference is using a different expansion:

$$\begin{aligned}\mathbf{M}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}} &= \mathbf{M}_{\mathbf{F}_{\mathbf{x}}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \\ &= \mathbf{M}_{\mathbf{F}_{\mathbf{x}}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1}} \\ &= \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1}\mathbf{D}_{T,\mathbf{x}}^{-1}(\mathbf{D}_{T,\mathbf{x}}^{-1}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+}\mathbf{D}_{T,\mathbf{x}}^{-1}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1\prime}\overline{\mathbf{V}}_{\dot{\mathbf{x}}}' \\ &+ \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1}\mathbf{D}_{T,\mathbf{x}}^{-1}(\mathbf{D}_{T,\mathbf{x}}^{-1}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+}\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}' \\ &+ \mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1}(\mathbf{D}_{T,\mathbf{x}}^{-1}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+}\mathbf{D}_{T,\mathbf{x}}^{-1}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1\prime}\overline{\mathbf{V}}_{\dot{\mathbf{x}}}' \\ &+ \mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1}\left[(\mathbf{D}_{T,\mathbf{x}}^{-1}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{-1}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+} - \mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+}\right]\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}' \\ &+ \mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1}\left[\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^{+} - (\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+}\right]\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}', \end{aligned} \tag{5.30}$$

which is just a different rotation of (3.41), which uses the fact that $\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}$ is square and invertible. This rotation is necessary in the non-stationary case in order to ensure that we work with $\left\|\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{V}_i\right\|$ and similar terms that are bounded in probability, because $\left\|\mathbf{D}_{T,\mathbf{x}}^{-1}\overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}'\mathbf{F}_{\mathbf{x}}'\mathbf{V}_i\right\|$ may not be, unless $\mathbf{F}_{\mathbf{x}}$ is of single integration order.

The proof of CCEMG case closely follows Theorem 4:

$$\begin{aligned}\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\nu}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\ &+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\nu}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\ &+ o_p(1), \end{aligned} \tag{5.31}$$

where the second term is negligible under general unknown factors as shown in Theorem 1 of Stauskas (2022), but in the current case the remainder terms are of even lower order, because only $\mathbf{F}_{\mathbf{x}}$ is being approximated. To show that the first term is negligible as well, we use the interim results leading to the same Theorem 1 of Stauskas (2022), which tell that

$$\left\|\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1} - \mathbf{\Sigma}_i^{-1}\right\| = O_p(N^{-1}) + O_p(T^{-1/2}) + O_p(N^{-1/2}T^{-\kappa/2}). \tag{5.32}$$

Therefore,

$$\begin{aligned}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| &\leq \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\ &+ \sqrt{N}\sup_i\left\|\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \mathbf{\Sigma}_i^{-1}\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\ &= o_p(1) \end{aligned} \tag{5.33}$$

under $TN^{-1} = O(1)$, because by the same steps as in the CCEP part, we have that $\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| = o_p(1)$. The first term is negligible as well, because it is almost identical (5.6), as scaling by $\mathbf{\Sigma}_i^{-1}$ will not change the orders. For example, by implementing the same decomposition as in (5.6), we get

$$\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \leq \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| = O_p(T^{-1/2}), \tag{5.34}$$

or

$$\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'\mathbf{P}_{\mathbf{F}_{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\leq \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}T^{-1}\mathbf{V}_i'\mathbf{P}_{\mathbf{F}_{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{P}_{\mathbf{F}_{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\leq T^{-1/2}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\gamma_i'\otimes\mathbf{\Sigma}_i^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1}\right)\right\| \left\|(\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+}\right\| \left\|\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\mathbf{D}_{T,\mathbf{y}}^{-1}\right\| \\
&+ T^{-1/2}\left\|\frac{1}{N}\sum_{i=1}^{N}\left(\gamma_i\otimes\mathbf{\Gamma}_i\mathbf{\Sigma}_i^{-1}\right)\right\| \left\|\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\sqrt{N}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1}\right\| \left\|(\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+}\right\| \left\|\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\mathbf{D}_{T,\mathbf{y}}^{-1}\right\| \\
&= O_p(T^{-1/2}).
\end{aligned} \tag{5.35}$$

Lastly,

$$\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}-\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\leq \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}T^{-1}\mathbf{V}_i'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}-\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&+ \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}-\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&= O_p(N^{-1/2})+O_p(T^{-\kappa/2}),
\end{aligned} \tag{5.36}$$

as the simpler second term is bounded by

$$
\begin{aligned}
&\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\Sigma}_i^{-1} T^{-1} \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' (\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0}) \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&\leq \frac{1}{\sqrt{T}} \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Sigma}_i^{-1} \right\| \left\| \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| \sqrt{N} T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0 \right\| \left\| (T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^+ \right\| \left\| T^{-1/2} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&+ \frac{1}{N\sqrt{T}} \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Sigma}_i^{-1} \right\| \left\| \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| N T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\|^2 \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^+ \right\| \left\| T^{-1/2} \sqrt{N} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&+ \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Sigma}_i^{-1} \right\| \left\| \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| T^{-1} \sqrt{N} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^+ \right\| \left\| \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \left\| \gamma_i \right\| \\
&+ \frac{1}{T} \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Sigma}_i^{-1} \right\| \left\| \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| \sqrt{N} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^+ \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| T^{-1/2} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&+ \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Sigma}_i^{-1} \right\| \left\| \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| \sqrt{N} T^{-1/2} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0 \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right\| \left\| (\mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0 \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1})^+ - \boldsymbol{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^+ \right\| \\
&\times \left\| \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime} \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \left\| \gamma_i \right\| \\
&+ \frac{1}{\sqrt{T}} \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Sigma}_i^{-1} \right\| \left\| \boldsymbol{\Gamma}_i' \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime} \right\| \left\| \sqrt{N} \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^+ - (\mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}' \mathbf{F} \mathbf{D}_{T,\mathbf{x}}^{-1})^+ \right\| \\
&\times \left\| \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \left\| \gamma_i \right\| \\
&+ O_p(N^{-1/2}) + O_p(T^{-\kappa/2}),
\end{aligned} \tag{5.37}
$$

which is driven by the second-to-last term, whereas the first term is bounded by

$$
\begin{aligned}
&\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\Sigma}_i^{-1} T^{-1} \mathbf{V}_i' (\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0}) \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&\leq \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Sigma}_i^{-1} \right\| \left\| \sqrt{N} T^{-1} \mathbf{V}_i' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0 \right\| \left\| (T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^+ \right\| \left\| T^{-1} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&+ \frac{1}{\sqrt{T}} \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Sigma}_i^{-1} \right\| \left\| T^{-1} \mathbf{V}_i' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\|^2 \left\| \sqrt{N} T^{-1/2} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}} \right\| \\
&+ \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Sigma}_i^{-1} \right\| \left\| \sqrt{N} T^{-1} \mathbf{V}_i' \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}} \right\| \left\| \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \left\| \gamma_i \right\| \\
&+ \frac{1}{T} \frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Sigma}_i^{-1} \right\| \left\| \mathbf{V}_i' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \sqrt{T} \mathbf{D}_{T,\mathbf{x}}^{-1} \right\| \left\| \sqrt{N} T^{-1/2} \overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^+ \right\| \\
&+ \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1/2} \boldsymbol{\Sigma}_i^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0 \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right\| \left\| (\mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0 \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1})^+ - \boldsymbol{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^+ \right\| \left\| \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime} \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \left\| \gamma \right\| \\
&+ \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\boldsymbol{\eta}_{\gamma,i}' \otimes T^{-1/2} \boldsymbol{\Sigma}_i^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0 \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right) \right\| \left\| (\mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0 \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1})^+ - \boldsymbol{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^+ \right\| \left\| \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime} \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \\
&+ \frac{1}{\sqrt{T}} \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\gamma_i' \otimes \boldsymbol{\Sigma}_i^{-1} \mathbf{V}_i' \mathbf{F}_{\mathbf{x}} \mathbf{D}_{T,\mathbf{x}}^{-1} \right) \right\| \left\| \boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}^+ - (\mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}' \mathbf{F} \mathbf{D}_{T,\mathbf{x}}^{-1})^+ \right\| \left\| \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right\| \\
&= O_p(N^{-1/2}) + O_p(T^{-\kappa/2}),
\end{aligned} \tag{5.38}
$$

which is driven by the third- and second-to-last terms due to

$$\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1/2} \mathbf{\Sigma}_i^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0 \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right\| \le \frac{1}{N} \sum_{i=1}^{N} \left\| \mathbf{\Sigma}_i^{-1} \right\| \left\| T^{-1} \sqrt{N} \mathbf{V}_i' \overline{\mathbf{V}}_{\dot{\mathbf{x}}}^0 \right\| \left\| \sqrt{T} \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right\| + \frac{\sqrt{N}}{\sqrt{T}} \frac{1}{N} \sum_{i=1}^{N} \left\| \mathbf{\Sigma}_i^{-1} \right\| \left\| \mathbf{V}_i' \mathbf{F}_{\dot{\mathbf{x}}}^0 \mathbf{D}_{T,\widehat{\mathbf{x}}}^{-1} \right\| = O_p(1) \tag{5.39}$$

under $TN^{-1} = O(1)$.

(b) The proof is almost identical to the proof of Proposition 1 of De Vos and Stauskas (2024), where the expansion in (5.8) is used instead and by assuming $\mathbf{F_x}$ that is trending non-stochastically to ensure that $\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x},\mathbf{y}}}$ is deterministic. An example of this would be $\mathbf{f}_{\mathbf{x},t} = (1, t, t^2, \ldots, t^{m_x - 1})' \in \mathbb{R}^{m_x}$ and $\mathbf{f}_{\mathbf{y},t}$ is covariance stationary with absolute summable autocovariances. Let $\mathbf{D}_{T,\mathbf{x}} = \text{diag}(T^{1/2}, T^{3/2}, \ldots, T^{(m_x - 1/2)})$ and $\mathbf{D}_{T,\mathbf{y}} = \sqrt{T} \mathbf{I}_{m_y}$. Then

$$\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x},\mathbf{y}}} = \underset{T\to\infty}{\text{plim}}\, \mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} = \underset{T\to\infty}{\text{plim}}\, \mathbf{D}_{T,\mathbf{x}}^{-1} \sum_{t=1}^{T} \mathbf{f}_{\mathbf{x},t} \mathbf{f}_{\mathbf{y},t}' \mathbf{D}_{T,\mathbf{y}}^{-1} = \int_{s=0}^{1} \left(\mathbf{s} \times \boldsymbol{\mu}_{\mathbf{f}_{\mathbf{y}}}' \right) ds, \tag{5.40}$$

where $\mathbf{s} = (1, s, s^2, \ldots, s^{m_x})'$, $s \in [0,1]$ and $\boldsymbol{\mu}_{\mathbf{f}_{\mathbf{y}}} = \mathbb{E}(\mathbf{f}_{\mathbf{y},t})$. To see how this result comes about, we can examine a typical element $\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x},\mathbf{y}}}^{j,l} = \mathbb{E}(f_{\mathbf{y},l,t}) \int_{s=0}^{1} s^j ds$ for $j = 0, \ldots, m_x - 1$ and $l = 1, \ldots, m_y$. Note how

$$\begin{aligned} \left(\mathbf{D}_{T,\mathbf{x}}^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \mathbf{D}_{T,\mathbf{y}}^{-1} \right)^{j,l} = \frac{1}{T^{j+1/2} \sqrt{T}} \sum_{t=1}^{T} t^j f_{\mathbf{y},l,t} &= \mathbb{E}(f_{\mathbf{y},l,t}) \frac{1}{T} \sum_{t=1}^{T} (t/T)^j + \frac{1}{T^{j+1/2} \sqrt{T}} \sum_{t=1}^{T} t^j (f_{\mathbf{y},l,t} - \mathbb{E}(f_{\mathbf{y},l,t})) \\ &= \mathbb{E}(f_{\mathbf{y},l,t}) \frac{1}{T} \sum_{t=1}^{T} (t/T)^j + O_p(T^{-1/2}) \\ &\to_p \mathbb{E}(f_{\mathbf{y},l,t}) \int_{s=0}^{1} s^j ds \end{aligned} \tag{5.41}$$

as $T \to \infty$ and

$$\begin{aligned} &\mathbb{E} \left[\left(\frac{1}{T^{j+1/2} \sqrt{T}} \sum_{t=1}^{T} t^j (f_{\mathbf{y},l,t} - \mathbb{E}(f_{\mathbf{y},l,t})) \right) \left(\frac{1}{T^{j+1/2} \sqrt{T}} \sum_{r=1}^{T} r^j (f_{\mathbf{y},l,r} - \mathbb{E}(f_{\mathbf{y},l,r})) \right) \right] \\ &= \frac{1}{T^{2j} T^2} \sum_{t=1}^{T} t^{2j} \mathbb{V}ar(f_{\mathbf{y},l,t}) + \frac{1}{T^{2j} T^2} \sum_{t=1}^{T} \sum_{r \neq t}^{T} t^j r^j \mathbb{C}ov(f_{\mathbf{y},l,t}, f_{\mathbf{y},l,r}) \\ &= O(T^{-1}), \end{aligned}$$

since $0 < (t/T)^j \le 1 = O(1)$ and hence

$$\begin{aligned} \left| \frac{1}{T^{2j} T^2} \sum_{t=1}^{T} t^{2j} \mathbb{V}ar(f_{\mathbf{y},l,t}) \right| &= \frac{1}{T^2} \sum_{t=1}^{T} (t/T)^{2j} \left| \mathbb{V}ar(f_{\mathbf{y},l,t}) \right| \\ &\le \frac{1}{T^2} \sum_{t=1}^{T} \left| \mathbb{V}ar(f_{\mathbf{y},l,t}) \right| \\ &= O(T^{-1}) \end{aligned} \tag{5.42}$$

and

$$\begin{aligned}
\left| \frac{1}{T^{2j}T^2} \sum_{t=1}^{T} \sum_{r \neq t}^{T} t^j r^j \mathbb{C}ov(f_{\mathbf{y},l,t}, f_{\mathbf{y},l,r}) \right| &\leq \frac{1}{T^{2j}T^2} \sum_{t=1}^{T} \sum_{r \neq t}^{T} t^j r^j \left| \mathbb{C}ov(f_{\mathbf{y},l,t}, f_{\mathbf{y},l,r}) \right| \\
&= \frac{1}{T^2} \sum_{t=1}^{T} \sum_{r \neq t}^{T} (t/T)^j (r/T)^j \left| \mathbb{C}ov(f_{\mathbf{y},l,t}, f_{\mathbf{y},l,r}) \right| \\
&\leq \frac{1}{T^2} \sum_{t=1}^{T} \sum_{r \neq t}^{T} \left| \mathbb{C}ov(f_{\mathbf{y},l,t}, f_{\mathbf{y},l,r}) \right| \\
&= O(T^{-1}) \qquad (5.43)
\end{aligned}$$

due to absolute summable autocovariances.

References

Andrews, D. W. K. (1987). Asymptotic results for generalized wald tests. *Econometric Theory*, 3(3):348–358.

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

De Vos, I. and Stauskas, O. (2024). Cross-section bootstrap for cce regressions. *Journal of Econometrics*, 240(1):105648.

Galvao, A. F. and Kato, K. (2014). Estimation and inference for linear panel data models under misspecification when both n and t are large. *Journal of Business & Economic Statistics*, 32(2):285–309.

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

Karabiyik, H., Reese, S., and Westerlund, J. (2017). On the role of the rank condition in CCE estimation of factor-augmented panel regressions. *Journal of Econometrics*, 197(1):60 – 64.

Stauskas, O. (2022). Complete theory for cce under heterogeneous slopes and general unknown factors. *Oxford Bulletin of Economics and Statistics*.

Westerlund, J. (2018). Cce in panels with general unknown factors. *The Econometrics Journal*, 21(3):264–276.