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

15 February 2024

Handling Distinct Correlated Effects with CCE

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

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

February 16, 2024

Abstract

The Common Correlated Effects (CCE) estimator is a popular method to estimate panel data regression models with interactive effects. Due to its simplicity in approximating the common factors with cross-section averages of the observables, it lends itself to a wide range of applications. They include static and dynamic models, homogeneous or heterogeneous coefficients or possibly very general types of factor structure. Despite such flexibility, with very few exceptions, CCE properties are usually examined under a restrictive assumption that all the observed variables load on the same set factors, which ensures joint identification of the factor space. In this paper, we explore an empirically relevant scenario when the dependent and explanatory variables are driven by distinct but correlated factors. In doing this, we consider panel dimensions such that $TN^{-1} \to \tau < \infty$ as $(N, T) \to \infty$, which is known to induce an asymptotic bias in CCE setting. We subsequently develop a toolbox to perform asymptotically valid inference in homogeneous and heterogeneous panels.

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

1 Introduction

We consider the following interactive effects model for unit $i = 1, \dots, N$ and period $t = 1, \dots, T$:

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

where $y_{i,t} \in \mathbb{R}$ is the dependent variable, $\mathbf{x}_{i,t} \in \mathbb{R}^k$ represents explanatory variables, and $\boldsymbol{\beta}$ is the parameter vector of interest. Equation (1.1) defines a multi-factor error structure permitting the cross-section units to be affected by common unobserved factors $\mathbf{f}_t \in \mathbb{R}^m$ to which they can respond with heterogeneous intensities (loadings) $\boldsymbol{\gamma}_i \in \mathbb{R}^m$. Factors drive the co-movements in the variable $y_{i,t}$ and induce "strong" cross-section dependence (see e.g. Chudik et al., 2011), while the idiosyncratic mean-zero innovations $\varepsilon_{i,t} \in \mathbb{R}$ are assumed to be covariance stationary and weakly dependent over time. Interactive effects come natural in macroeconomic applications with panel data where both N and T are large (see Westerlund et al., 2019, for small T, or micro, examples). Here, $\mathbf{f}_t$ may represent the unobserved global technological progress (with $\boldsymbol{\gamma}_i$ representing the local absorption intensity) that is relevant for modelling long-run growth (see e.g. Eberhardt and Teal, 2011), or business cycles that mediate the relationship between public debt and economic growth (see e.g. Eberhardt and Presbitero, 2015, or Chudik et al., 2017).

It is natural for $\mathbf{f}_t$ to be correlated with $\mathbf{x}_{i,t}$. Therefore, by following Pesaran (2006) we let

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

where $\mathbf{\Gamma}_i \in \mathbb{R}^{m\times k}$ is the loading matrix and $\mathbf{v}_{i,t} \in \mathbb{R}^k$ is the vector of idiosyncratic errors. Hence, the model in (1.1) - (1.2) exhibits not only strong cross-section dependence, but also endogeneity making it essential to control for $\mathbf{f}_t$. The Common Correlated Effects (CCE) approach by Pesaran (2006) estimates the factor space with the cross-section averages (CAs) of the observables $\widehat{\mathbf{f}}_t = \overline{\mathbf{z}}_t = \frac{1}{N}\sum_{i=1}^{N}\mathbf{z}_{i,t}$, where $\mathbf{z}_{i,t} = [y_{i,t}, \mathbf{x}_{i,t}']' \in \mathbb{R}^{k+1}$, and adds them as regressors to (1.1), which is in turn estimated by Least Squares (LS). CCE enjoys popularity due to its simplicity and excellent small sample performance (see e.g. Westerlund and Urbain, 2015), hence it has been used in various settings such as structural breaks modelling or unit root testing (see Karavias et al., 2023, and Norkutė and Westerlund, 2021).

A necessary condition for informativeness about the factor space is $m \leq k+1$, which means that we have enough CAs to proxy $\mathbf{f}_t$. If the CAs are informative, then $\widehat{\mathbf{f}}_t$ is consistent for (the space spanned by) the factors and LS yields consistent estimates of $\boldsymbol{\beta}$ as $N \to \infty$ for T fixed or growing (see Westerlund et al., 2019). In large T panels, if $TN^{-1} \to 0$, standard normal inference ensues, whereas if $TN^{-1} \to \tau < \infty$, a bias-correction is unavoidably needed due to accumulation of the factor estimation error generated at every $t = 1, \ldots, T$ (see e.g. Westerlund and Urbain, 2015). The structure of the bias depends on whether $m = k+1$ or $m < k+1$, because in the latter case the excess CAs result in nuisance parameters due to asymptotic singularity of $\frac{1}{T}\sum_{t=1}^{T}\widehat{\mathbf{f}}_t\widehat{\mathbf{f}}_t'$ (see Karabiyik et al., 2017; also see Westerlund and Urbain, 2013, or De Vos and Stauskas, 2024, for the potential solutions). Both fixed- and large-T settings need a strong assumption that all the factors are estimable by the CAs. However, if some factors are unattended, then even consistency of CCE might break down. Therefore, in the current paper we put this assumption to test. While there are several situations in which the CAs can be uninformative (e.g. when $\overline{\mathbf{C}}$ is very sparse), we focus on the very estimation mechanics of CCE, where the dependent and explanatory variables load on the same set of factors, and $\overline{y}_t$ alone can proxy only one unique factor at most.

To make the above discussion a little more precise, by inserting (1.2) into (1.1), we get $\mathbf{z}_{i,t} = \mathbf{C}_i'\mathbf{f}_t + \mathbf{u}_{i,t}$ and $\overline{\mathbf{z}}_t = \widehat{\mathbf{f}}_t = \overline{\mathbf{C}}'\mathbf{f}_t + \overline{\mathbf{u}}_t$, where $\overline{\mathbf{C}} \in \mathbb{R}^{m\times(k+1)}$ is a function of $\overline{\mathbf{\Gamma}}$ and $\overline{\gamma}$, and $\overline{\mathbf{u}}_t$ is negligible as $N \to \infty$ under a wide variety of empirically relevant assumptions (see e.g. Pesaran and Tosetti, 2011). If the row rank of $\overline{\mathbf{C}}$ is m for all N including $N \to \infty$, then

$$\mathbf{f}_t = (\overline{\mathbf{C}}')^{+}\left[\widehat{\mathbf{f}}_t - \overline{\mathbf{u}}_t\right] \approx (\overline{\mathbf{C}}')^{+}\widehat{\mathbf{f}}_t, \tag{1.3}$$

where $\overline{\mathbf{C}}^{+} = \overline{\mathbf{C}}'(\overline{\mathbf{C}}\overline{\mathbf{C}}')^{-1}$ is the Moore-Penrose (MP) inverse of $\overline{\mathbf{C}}$. The result in (1.3) implies that the CAs *jointly* span the factor space asymptotically. While the row rank condition is sufficient for the approximation, it is not guaranteed to hold in many empirically relevant scenarios. Indeed, by following Cui et al. (2022), we split the total factors into $\mathbf{f}_t = [\mathbf{f}_{\mathbf{y},t}', \mathbf{f}_{\mathbf{x},t}']'$, such that $m = m_y + m_x$ and $\mathbb{C}ov(\mathbf{f}_{\mathbf{y},t}, \mathbf{f}_{\mathbf{x},t}) \neq \mathbf{0}_{m_y\times m_x}$. Plus, the rank of $\overline{\mathbf{\Gamma}}$ is m_x, so that $\overline{\mathbf{x}}_t$ is informative for $\mathbf{f}_{\mathbf{x},t}$. The cases interesting to us are the following:

1. $\mathbf{f}_{\mathbf{y},t} \cap \mathbf{f}_{\mathbf{x},t} = \varnothing$, so that (1.1) - (1.2) are driven by the "distinct" sets of factors.

2. $\mathbf{f}_{\mathbf{x},t} \subset \mathbf{f}_{\mathbf{y},t}$, so that $y_{i,t}$ contains additional factors.

If $m_y > 1$, it is possible that $m > k+1$, thus $\overline{\mathbf{C}}^{+}$ does not exist and the representation in (1.3) is not valid, which renders the CCE estimator inconsistent. It is so, because $\overline{y}_t$ is a *single* CA, and it contributes to estimation of the total $\mathbf{f}_t$ when $\mathbf{f}_{\mathbf{y},t} = \mathbf{f}_{\mathbf{x},t} = \mathbf{f}_t$, i.e. all CAs estimate the same factors. Otherwise, we need $m_y = 1$, so that $\overline{y}_t$ is informative about its single factor in case 1, or there is at most one additional factor in case 2. Neither of these cases can typically be tested in practice, therefore, a robust approach is needed.

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) in the Principal Components (PC) context assume factors in $e_{i,t}$ from (1.1) only as they use high-level conditions to determine correlation between $\mathbf{f}_{\mathbf{y},t}$ and $\mathbf{x}_{i,t}$. While flexible, such approach relies on a non-linear optimization problem, therefore convergence issues may arise (see e.g. Jiang et al., 2021). For CCE, 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. After the appropriate regularization of the CCE estimator, $\overline{\mathbf{z}}_t$ remains informative about $\mathbf{f}_{1,t}$, which means that $y_{i,t}$ and $\mathbf{x}_{i,t}$ again jointly identify the factor space. 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.

We utilize the latter strategy in the CCE context to solve the problem of a limited informativeness of $\overline{y}_t$. As the rank condition of $\overline{\boldsymbol{\Gamma}}$ holds, it validates the (first stage) de-correlation step. Our central finding is that the remaining $\mathbf{f}_{\mathbf{y},t}$ may still seriously affect the asymptotic distribution of CCE. Moreover, the distribution is non-standard if $m_x < k$, so that the excess CAs still retain a non-trivial effect (see a similar finding in Juodis, 2022). As $\widehat{\mathbf{f}}_{\mathbf{y},t}$ is typically unavailable, we aid the second stage with the cross-section (CS) bootstrap scheme introduced by Kapetanios (2008) in panel data context, and formalized by De Vos and Stauskas (2024) for the CCE regressions when N and T are large. In particular, we demonstrate that the variance of the asymptotic distribution and the bias depend on the unknown covariance between $\mathbf{f}_{\mathbf{x},t}$ and $\mathbf{f}_{\mathbf{y},t}$, and in turn we formulate conditions under which CS bootstrap is able to replicate this distribution. As a result, this enables estimation of the asymptotic variance and allows accommodation of the usual $TN^{-1} \to \tau < \infty$ bias in the spirit of Gonçalves and Perron (2014) or Djogbenou et al. (2015).

In the pursuit of re-establishing asymptotically valid inference, we consider homogeneous and heterogeneous panels. Specifically, we let $\boldsymbol{\beta}$ be either constant or vary across individuals, which leads to the analysis of the so-called Pooled (CCEP) or Mean Group (CCEMG) estimators. Under homogeneous slopes, irrespective if $TN^{-1} \to \tau < \infty$ or $TN^{-1} \to 0$, the asymptotic distribution of the CCEP estimator is non-standard if any redundant CAs ($m_x < k$) are used to estimate the factor space, and this result echoes findings of Juodis et al. (2021) (see their Theorem 1). The reason for this is interaction between the accumulated factor estimation error and covariance between the two sets of factors. In such case, CS bootstrap is not consistent as the non-normal part driven by the redundant averages is exacerbated by the bootstrap-induced randomness. In order to exactly match the number of CAs and m_x, we employ an Information Criterion in the spirit of Margaritella and Westerlund (2023). In case of heterogeneous slopes, both CCEP and CCEMG estimators are asymptotically normal and unbiased irrespective if $m_x = k$ or $m_x < k$, which corresponds to the result in Theorem 2 of Cui et al. (2022) in the PC setting. The intuition lies in the fact that slope heterogeneity dominates the asymptotic distribution of both estimators, meaning that $\mathbf{f}_{\mathbf{y},t}$ plays the same role as the idiosyncratic error component. This result is also captured by the CS bootstrap. The major implication of our proposed toolbox is that as long as the rank of $\overline{\boldsymbol{\Gamma}}$ is m_x, then asymptotically valid inference can ensue irrespective of whether $\overline{y}_t$ is informative or not, $m_x = k$ or $m_x < k$, or the slopes are heterogeneous or not. This significantly boosts applicability of the CCE methods.

The remainder of this paper is organized as follows: in Section 2 we provide our assumptions with the details on the estimators and an explanation of the CS bootstrap scheme. In Section 3, we derive the asymptotic distribution of CCEP and CCEMG in the original and bootstrap samples and discuss inference by exploring the asymptotic variance estimator. Monte Carlo evidence and a comparison 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. Next, $\|\mathbf{A}_n\| = O_p(a_n)$ means that a random vector sequence $\mathbf{A}_n$ is at most of order a_n in probability, where $a_n \in \mathbb{R}_{++}$ is a generic deterministic sequence. $\|\mathbf{A}_n\| = o_p(a_n)$ means it is of smaller order in probability than a_n. Finally, 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 (1.1) - (1.2) in time-stacked notation

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{F}_\mathbf{y}\gamma_i + \boldsymbol{\varepsilon}_i, \qquad \mathbf{X}_i = \mathbf{F}_\mathbf{x}\boldsymbol{\Gamma}_i + \mathbf{V}_i, \tag{2.1}$$

where, $\mathbf{y}_i = [y_{i,1},\dots,y_{i,T}]' \in \mathbb{R}^{T\times 1}$ for $i = 1,\dots,N$, $\mathbf{X}_i = [\mathbf{x}_{i,1},\dots,\mathbf{x}_{i,T}]' \in \mathbb{R}^{T\times k}$, $\mathbf{V}_i = [\mathbf{v}_{i,1},\dots,\mathbf{v}_{i,T}]' \in \mathbb{R}^{T\times k}$ and $\boldsymbol{\varepsilon}_i = [\varepsilon_{i,1},\dots,\varepsilon_{i,T}]' \in \mathbb{R}^{T\times 1}$, $\mathbf{F} = [\mathbf{f}_1,\dots,\mathbf{f}_T]' \in \mathbb{R}^{T\times m}$. Because we focus on a realistic case when $m_y > 1$, $\overline{\mathbf{y}}$ is uninformative. Therefore, the factors $\mathbf{F}_\mathbf{x}$ can be estimated by

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

which implies that

$$\mathbf{F}_\mathbf{x} = (\widehat{\mathbf{F}}_\mathbf{x} - \overline{\mathbf{V}})\overline{\boldsymbol{\Gamma}}^{+}, \tag{2.3}$$

where $\overline{\boldsymbol{\Gamma}}^{+}$ is the MP inverse of $\overline{\boldsymbol{\Gamma}}$ that exists if $\mathrm{rk}(\overline{\boldsymbol{\Gamma}}) = m_x$. Note that (2.1) excludes individual fixed effects (FE). However, this is only for the ease of exposition, because the CCE distributions discussed in the upcoming sections will stay invariant to FE if a column of ones is added to the CAs. That is, $\widehat{\mathbf{F}}_{\iota,\dot{\mathbf{x}}} = [\boldsymbol{\iota}_T, \overline{\mathbf{X}}]$, such that we now use $\mathbf{M}_{\widehat{\mathbf{F}}_{\iota,x}}$. This means that we employ time-demeaned observables $\widetilde{\mathbf{y}}_i = \mathbf{y}_i - \overline{\mathbf{y}}$.

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 $\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},y}} = \mathrm{plim}_{T\to\infty} T^{-1}\mathbf{F}_\mathbf{x}'\mathbf{F}_\mathbf{y}$ 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*

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

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) $\mathrm{rk}(\overline{\boldsymbol{\Gamma}}\mathbf{q}_{\ddot{\mathbf{x}}}) = m$, 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*

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

with $\boldsymbol{\Omega}_{\boldsymbol{v}}$ *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 $\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} = \mathbf{F}_{\mathbf{x}}$, *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 assumptions that we use are very 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 $TN^{-1} \to \tau < \infty$ 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). 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 employed to approximate the factors by introducing a 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 required the CCE specifications. This corresponds to practice where some observables (e.g. the dependent variable, 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 details). Assumption 6 formalizes the slope heterogeneity, while Assumption 7 is sufficient for identification of the mean $\boldsymbol{\beta}$ when the sloeps are heterogeneous.

We further define the CCEP and CCEMG estimators. By letting $\overline{\mathbf{Q}}_{\dot{\mathbf{x}}} = \frac{1}{N}\sum_{i=1}^{N}\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}$, we have

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} &= \overline{\mathbf{Q}}_{\dot{\mathbf{x}}}^{-1}\frac{1}{N}\sum_{i=1}^{N}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i,\\ &= \boldsymbol{\beta} + \overline{\mathbf{Q}}_{\dot{\mathbf{x}}}^{-1}\frac{1}{N}\sum_{i=1}^{N}T^{-1}\left(\mathbb{I}_{\boldsymbol{\nu}\neq\mathbf{0}} \times \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i + \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right) \qquad (2.4)\end{aligned}$$

and

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} &= \frac{1}{N}\sum_{i=1}^{N}\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i\\ &= \boldsymbol{\beta} + \mathbb{I}_{\boldsymbol{\nu}\neq\mathbf{0}} \times \overline{\boldsymbol{\nu}} + \frac{1}{N}\sum_{i=1}^{N}\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}^{-1}T^{-1}(\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + \mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i), \qquad (2.5)\end{aligned}$$

where $\mathbb{I}_{\boldsymbol{\nu}\neq\mathbf{0}}$ is an indicator function which equals to $\mathbf{I}_k$ or $\mathbf{0}_{k\times k}$ depending whether the slopes are heterogeneous or not. The typical estimators of the asymptotic variance suggested by Pesaran (2006) are given by

$$\widehat{\boldsymbol{\Theta}}_{CCEP,\dot{\mathbf{x}}} = \overline{\mathbf{Q}}_{\dot{\mathbf{x}}}^{-1}\left(\frac{1}{N(N-1)}\sum_{i=1}^{N}\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})'\widehat{\mathbf{Q}}_{\dot{\mathbf{x}},i}\right)\overline{\mathbf{Q}}_{\dot{\mathbf{x}}}^{-1}, \qquad (2.6)$$

$$\widehat{\boldsymbol{\Theta}}_{CCEMG,\dot{\mathbf{x}}} = \frac{1}{N(N-1)}\sum_{i=1}^{N}(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})(\widehat{\boldsymbol{\beta}}_{\dot{\mathbf{x}},i} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})' \qquad (2.7)$$

for CCEP and CCEMG estimators, respectively. Clearly, the expansions in (2.4) and (2.5) reveal that $\mathbf{F_y}$ enters the asymptotic analysis of both estimators non-trivially. Intuitively, because $\mathbf{F_y}$ is not projected out as $\overline{\mathbf{y}}$ is uninformative, it should affect the asymptotic distribution by altering the variance and possibly the mean because $\mathbf{F_y}$ is typically not mean-zero. Moreover, because m_y is unknown and it is potentially bigger than 1, we are also running the risk of having more factors than the available CAs. In order to take these major deviations from the classical CCE setup into account, we will employ the cross-section (CS) bootstrap theory by De Vos and Stauskas (2024) built for the CCE estimators. We begin with a general description and a practical implementation of the CS bootstrap algorithm.

Remark 1. *Note that if $\mathbf{q}_{\dot{\mathbf{x}}} = \mathbf{I}_k$, such that the whole $\overline{\mathbf{X}}$ is employed, then (2.4) can be simplified by noticing that*

$$\frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F_y}(\gamma+\boldsymbol{\eta}_{\gamma,i}) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F_y}\boldsymbol{\eta}_{\gamma,i},$$

since $\overline{\mathbf{X}}'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} = \mathbf{0}_{k\times T}$ then. 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.

2.2 Bootstrap Algorithm

The CS bootstrap scheme is straightforward to implement, and has the advantage that factors are replicated in the bootstrap realm without requiring a decision on their number. 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 $\sigma\{\mathbf{F}\}$. To present the resampling scheme, let $\mathcal{B}_b^* = \{\mathbf{Z}_1^*, \dots, \mathbf{Z}_N^*\}$ denote bootstrap sample $b = 1, \dots, B$, obtained as described in Algorithm 1 below. Accordingly, for $s \in \{CCEP, CCEMG\}$, we use $\widehat{\boldsymbol{\beta}}_{s,b}^* = \widehat{\boldsymbol{\beta}}_s(\dot{\mathbf{x}}, \mathcal{B}_b^*)$, denote the estimates in bootstrap sample b following the same specification $\dot{\mathbf{x}}$.

Algorithm 1: Cross-section resampling scheme.

1) Initialization: Choose the combination of the CAs with the appropriate $\mathbf{q}_{\dot{\mathbf{x}}}$. Estimate $\widehat{\boldsymbol{\beta}}_s = \widehat{\boldsymbol{\beta}}_s(\dot{\mathbf{x}}, \mathcal{B})$ based on the original sample.

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

   i) Generate $\mathcal{B}_b^* = \{\mathbf{Z}_1^*, \dots, \mathbf{Z}_N^*\}$ according to

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

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

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

$$CI(\alpha, \widehat{\boldsymbol{\beta}}_s^*) = \left[\, 2\widehat{\boldsymbol{\beta}}_s - \boldsymbol{\theta}_{(1-\alpha/2)}^*(\widehat{\boldsymbol{\beta}}_s^*)\, ,\, 2\widehat{\boldsymbol{\beta}}_s - \boldsymbol{\theta}_{\alpha/2}^*(\widehat{\boldsymbol{\beta}}_s^*)\,\right], \tag{2.8}$$

where $\boldsymbol{\theta}_\alpha^*(\cdot)$ is the empirical α-quantile of the obtained bootstrap distribution for the statistic inside the brackets and the quantiles are understood element-wise.

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:

$$\mathbf{y}_i^* = \mathbf{y}_{i^*} = \mathbf{X}_{i^*}\boldsymbol{\beta} + \mathbf{F}_{\mathbf{y}}\gamma_{i^*} + \boldsymbol{\varepsilon}_{i^*} \tag{2.9}$$
$$\mathbf{X}_i^* = \mathbf{X}_{i^*} = \mathbf{F}_{\mathbf{x}}\boldsymbol{\Gamma}_{i^*} + \mathbf{V}_{i^*} \tag{2.10}$$

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 also are copied in their entirety, but implicitly permuted across units under the assumption that these are cross-sectionally independent. This retains the within-unit correlations and variances of loadings and innovations, as well as their time series properties. The bootstrap factor space estimator can then be expressed as

$$\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^* = \frac{1}{N}\sum_{i=1}^{N}\mathbf{X}_i^*\mathbf{q}_{\ddot{\mathbf{x}}} = \overline{\mathbf{X}}^*\mathbf{q}_{\ddot{\mathbf{x}}} = (\mathbf{F}_{\mathbf{x}}\overline{\boldsymbol{\Gamma}}_w + \overline{\mathbf{V}}_w)\mathbf{q}_{\ddot{\mathbf{x}}}, \tag{2.11}$$

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, and s_i denotes the sampling frequency of unit i in the bootstrap dataset $\mathcal{B}_b^*$, which follows a multinomial distribution. The properties of s_i imply that $\overline{\mathbf{V}}_w \to_{p^*} \mathbf{0}_{T\times k}$ and $\overline{\boldsymbol{\Gamma}}_w\mathbf{q}_{\ddot{\mathbf{x}}} \to_{p^*} \boldsymbol{\Gamma}_{\ddot{\mathbf{x}}}$ as $N\to\infty$, and in turn $(\overline{\boldsymbol{\Gamma}}_w\mathbf{q}_{\ddot{\mathbf{x}}})^+ \to_{p^*} \boldsymbol{\Gamma}_{\ddot{\mathbf{x}}}^+$. This confirms that the asymptotic information content in the cross-section averages 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. To begin with, we assume that $\mathbb{I}_{\boldsymbol{\nu}\neq\mathbf{0}} = \mathbf{0}_{k\times k}$.

3.1 Homogeneous Slopes

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

(a) If $m_x < g$:

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

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

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

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

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

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

Theorem 1 (a) and (b) confirm our prediction that presence of the unaccounted $\mathbf{F_y}$ affects both the mean and the variance of the asymptotic CCEP distribution as $TN^{-1} \to \tau < \infty$. In particular, the variance is affected irrespective of the N, T expansion rate. The bias is also a function of the remaining factors due to the presence of $\mathbf{\Sigma}_{\mathbf{F}_{x,y}}$. Importantly, under part (a) we have a term $\mathbf{h}_2$ which is stochastic and does not follow normal distribution, therefore making the total distribution non-standard. This term is mainly driven by the interaction of two components: the error part of the $g - m_x$ redundant CAs and a covariance between $\mathbf{F_x}$ and $\mathbf{F_y}$. Unlike the deterministic bias components, $\mathbf{h}_2$ cannot be eliminated even if $TN^{-1} \to 0$. The asymptotic variance estimator in (2.6) is inconsistent, because it necessarily requires the restriction of $\mathbf{F_x} = \mathbf{F_y} = \mathbf{F}$; otherwise it cannot identify $\mathbf{\Psi}_f$ (See Proposition 3 in the supplementary material of De Vos and Stauskas, 2024). Due to these reasons, using bootstrap methods is necessary.

On the other hand, part (b) demonstrates that if we have exactly $m_x = g$, then the asymptotic distribution does not include any non-standard terms impeding normal inference. Nevertheless, the bias $\widetilde{\mathbf{h}}_1$ still depends on $\mathbf{\Sigma}_{\mathbf{F}_{x,y}}$. This means that the bias cannot be estimated in spirit of Westerlund and Urbain (2013) even under $m_x = g$, because $\mathbf{F_y}$ is neither observed, nor estimated. Similarly to part (a), the variance estimator in (2.6) is inconsistent for the $\widetilde{\mathbf{\Psi}}_f$ component, which means that it is essential to consider bootstrap distributions for both cases.

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 $m_x < g$*:*

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}^{*}_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \to_{d^*} \mathcal{N}(\mathbf{0}_{k\times1}, \mathbf{\Sigma}^{-1}(\mathbf{\Psi} + \mathbf{\Psi}_f)\mathbf{\Sigma}^{-1}) + \mathbf{\Sigma}^{-1}(\sqrt{\tau}\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}^{+})$$

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 $m_x = g$*:*

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}^{*}_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \to_{d^*} \mathcal{N}\left(\mathbf{0}_{k\times1}, \mathbf{\Sigma}^{-1}(\mathbf{\Psi} + \widetilde{\mathbf{\Psi}}_f)\mathbf{\Sigma}^{-1}\right) + \sqrt{\tau}\mathbf{\Sigma}^{-1}\widetilde{\mathbf{h}}_1,$$

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

$$\sup_{x\in\mathbb{R}^{k\times1}} \left| \mathbb{P}^{*}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^{*}_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \le x] - \mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \le x] \right| \to_p 0,$$

where the inequalities should be interpreted coordinate-wise.

It is evident from part (a) that while the asymptotic variance is replicated in the bootstrap realm, the mean is not due to the presence of an extra noise $\mathbf{h}^{+}_2$. This is so, because the stochastic term $\mathbf{h}_2$ has a very complicated functional form, and it becomes exacerbated in the bootstrap realm once it is augmented with the bootstrap weights. However, if $m_x = g$, then the original sample and bootstrap distributions coincide due to the fact that the excess $m_x - g$ CAs are now absent. Therefore, ensuring the condition of $m_x = g$ is quintessential for the normal asymptotic inference to ensue under the distinct factor case.

In order to asymptotically guarantee that $m_x = g$, we use the following Information Criterion (IC) that was adapted from Margaritella and Westerlund (2023) by De Vos and Stauskas (2024):

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

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 $\text{rk}(\mathbf{\Gamma}\mathbf{q}_{\dot{\mathbf{x}}}) = m_x$, $\text{cols}(\mathbf{q}_{\dot{\mathbf{x}}}) = m_x$, 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 $p_{N,T}C^2_{N,T} \to \infty$ 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, $p_{N,T} = \frac{N+T}{NT}\log(C^2_{N,T})$ 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 set in Proposition 3 of De Vos and Stauskas, 2024).[1] 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. (2023). In summary, the consistency of (3.1) guarantees that the conditions of part (b) of Theorem 2 can be met, so that the asymptotic bias and the variance can be estimated by the means of CS bootstrap.

3.2 Heterogeneous Slopes

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

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

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

where $\mathbf{\Sigma} = \text{plim}_{N\to\infty} \frac{1}{NT}\sum_{i=1}^{N} \mathbf{V}_i'\mathbf{V}_i$ *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 is $\sqrt{N}$-consistent, it also has an asymptotically normal distribution, and the relative N, T expansion rate does not play a role. The theorem 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 if $m_x < g$ or $m_x = g$. Moreover, $\mathbf{F}_{\mathbf{y}}$ does not affect the asymptotic variance. This result coincides with the findings of Stauskas

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

(2022) and the heterogeneous slopes analysis of De Vos and Stauskas (2024), where $\mathbf{F}_{\mathbf{x}} = \mathbf{F}_{\mathbf{y}} = \mathbf{F}$ in both studies. To the best of our knowledge, Theorem 3 is the first to highlight CCE robustness to distinct factors in heterogeneous panels. The intuition behind this result follows from the two facts. Firstly, the slope heterogeneity $\boldsymbol{\nu}_i$ dominates the asymptotic distribution through

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\bar{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i\boldsymbol{\nu}_i + o_p(1), \tag{3.3}$$

which obeys the standard Central Limit Theorem (CLT), and the term driven by the idiosyncratic error $\boldsymbol{\varepsilon}_i$ in the expansion in (2.4) is of the lower order. Secondly, the component driven by $\mathbf{F}_{\mathbf{y}}$ can similarly be treated as an idiosyncratic term. This is so, because $\mathbf{M}_{\widehat{\mathbf{F}}_{\bar{\mathbf{x}}}}\mathbf{X}_i$ is asymptotically uncorrelated with $\mathbf{F}_{\mathbf{y}}$ as $\mathbf{F}_{\mathbf{x}}$ is projected out. Therefore,

$$\left\| \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\bar{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| = o_p(1). \tag{3.4}$$

In effect, $\mathbf{F}_{\mathbf{y}}$ can be asymptotically neglected, as long as $\mathbf{F}_{\mathbf{x}}$ can be consistently estimated in the heterogeneous panels. We further turn to the CCEMG estimator.

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

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

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 the variance unaffected by the presence of $\mathbf{F}_{\mathbf{y}}$ irrespective if $m_x < g$ or $m_x = g$.[2] The rationale behind this outcome is the same as behind Theorem 3, meaning that the slope heterogeneity is dominant:

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

While this result is new in the CCE literature, it also coincides with Proposition 4.1 in Cui et al. (2022) in the PC context. Particularly, their two-stage procedure can now be reduced to the first stage estimation of $\mathbf{F}_{\mathbf{x}}$ only, where the dominance of $\boldsymbol{\nu}_i$ relegates the effect $\mathbf{F}_{\mathbf{y}}$ to the idiosyncratic components.[3] This is also the main message of our Theorem 4 in the CCE context.

Remark 2. *Given that CCE can accommodate a wide variety of factors without changing the rate of convergence (see Westerlund, 2018, or Stauskas, 2022), it is important to see whether it is still possible in a distinct factors case. Turns out, only to a limited degree: $\mathbf{F}_{\mathbf{x}}$ cannot be stochastically trending, and $\mathbf{F}_{\mathbf{y}}$ remains stationary. The limitation can be illustrated with the following example. Let $\mathbf{F}$ be such that $\mathbf{D}_T^{-1}\mathbf{F}'\mathbf{F}\mathbf{D}_T^{-1}$ is asymptotically full-rank and $(\mathbf{I}_{k+1} \otimes \mathbf{D}_T^{-1})vec(\mathbf{F}'\mathbf{U}_i)$ converges weakly for some normalization matrix $\mathbf{D}_T = \mathrm{diag}(\mathbf{D}_{T,\mathbf{x}}, \mathbf{D}_{T,\mathbf{y}})$. Take the CCEP expansion in (2.4) for either homogeneous or heterogeneous slopes (CCEMG case is similar). By using (2.3), we can show that one of the key terms in the asymptotic analysis is*

$$\frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\bar{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i = \frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\gamma_i - \frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{P}_{\mathbf{F}_{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\gamma_i - \frac{1}{NT}\sum_{i=1}^{N}\mathbf{V}_i'(\mathbf{M}_{\mathbf{F}_{\mathbf{x}}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\bar{\mathbf{x}}}})\mathbf{F}_{\mathbf{y}}\gamma_i = \mathbf{I} - \mathbf{II} - \mathbf{III}.$$

[2]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.

[3]Note that according to (2.5) and Theorem 4, under homogeneous $\boldsymbol{\beta}$, we have $\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\bar{\mathbf{x}}} - \boldsymbol{\beta}) = o_p(N^{-1/2})$. This means that we can always consistently estimate the homogeneous $\boldsymbol{\beta}$ by CCEMG , but inference should be based on $\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEMG,\bar{\mathbf{x}}} - \boldsymbol{\beta})$, as suggested by Theorem 1 and 2. We skip such analysis in the interest of space.

We must have $\sqrt{N}\mathbf{I} = o_p(1)$ under heterogeneous slopes. This happens only when $\mathbf{F_y}$ is stationary. In the same fashion, $\sqrt{N}\mathbf{II}$ and $\sqrt{N}\mathbf{III}$ should be negligible under heterogeneous slopes, as well. Again, this is ensured only if $\mathbf{F_y}$ is stationary. Let $\mathbf{D}_{T,\mathbf{y}} = \sqrt{T}\mathbf{I}_{m_y}$. Then, by using self-normalization of the projection matrix:

$$\left\|\sqrt{N}\mathbf{II}\right\| \leq \frac{1}{\sqrt{T}} \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\gamma_i' \otimes \mathbf{V}_i'\mathbf{F_x}\mathbf{D}_{T,\mathbf{x}}^{-1} \right) \right\| \left\| (\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_\mathbf{x}'\mathbf{F_x}\mathbf{D}_{T,\mathbf{x}}^{-1})^{+} \right\| \left\| \mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_\mathbf{x}'\mathbf{F_y}\mathbf{D}_{T,\mathbf{y}}^{-1} \right\| = o_p(1),$$

where $\sqrt{N}\mathbf{III}$ behaves similarly as the expansion of $\mathbf{M}_{\mathbf{F}_{\dot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\mathbf{x}}}$ produces self-normalizing terms. In homogeneous case, $\sqrt{NT}\mathbf{I}$, $\sqrt{NT}\mathbf{II}$ and $\sqrt{NT}\mathbf{III}$ are $O_p(1)$ when $\mathbf{F_y}$ is stationary. Plus, if $\mathbf{F_x}$ is not stochastically trending, the terms involving $\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1/2}vec(\mathbf{F}_\mathbf{y}'\mathbf{V}_i)$ and $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(\mathbf{I}_k \otimes \mathbf{D}_{T,\mathbf{x}}^{-1})vec(\mathbf{F}_\mathbf{x}'\mathbf{V}_i)$ will contribute to asymptotically normal distribution (see Phillips and Moon, 1999), whereas the "covariance" term $\mathbf{D}_{T,\mathbf{x}}^{-1}\mathbf{F}_\mathbf{x}'\mathbf{F_y}\mathbf{D}_{T,\mathbf{y}}^{-1}$ will stay deterministic and contribute to the mean. Overall, we will avoid having a non-standard distribution.

Theorem 3 and Theorem 4 suggest that the variance estimators in (2.6) - (2.7) should be consistent, unlike under Theorem 1. Theorem 5 confirms this.

Theorem 5. *Under Assumptions 1 - 7, for either $m_x < g$ or $m_x = g$ 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, inference does not require to be aided by the means of bootstrap if it is known that $\boldsymbol{\beta}$ is heterogeneous. Additionally, we do not need to take into consideration whether $m_x = k$ or $m_x < k$, which is a major convenience. As it is a priori unclear whether the factors are distinct, in the same fashion, it is unclear if the slopes are heterogeneous or not. While this can be tested[4], the most suitable approach would be to have a robust procedure, which does not require discrimination between the two cases. Indeed, because we can rely on a handy CS bootstrap as long as $m_x = g$ is guaranteed in the homogeneous slopes case, it is natural to attempt the same in the heterogeneous panels. It is especially beneficial, because the asymptotic properties of CCEP and CCEMG are invariant to whether $m_x = g$ or $m_x < g$, according to Theorem 3 and Theorem 4. In Theorem 6 below, we provide the bootstrap consistency results for both estimators under the heterogeneous slopes.

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.

The major practical implication of Theorem 6 is that a researcher does not need to differentiate between homogeneous and heterogeneous panels and whether $\mathbf{y}_i$ and $\mathbf{X}_i$ are driven by the common or distinct factors under the asymptotic configuration of $TN^{-1} \to \tau < \infty$. This holds as long as $\text{rk}(\overline{\mathbf{\Gamma}}\mathbf{q}_{\dot{\mathbf{x}}}) = m_x$, and we can approximate the space spanned by $\mathbf{F_x}$. Even if bootstrap is not strictly needed in heterogeneous panels, Theorem 7 in the supplementary material provides the bootstrap equivalent of Theorem 5 for completeness. Therefore, bootstrap t-statistics can be computed, as well, if a researcher prefers such

[4]See some tests discussed in Pesaran and Yamagata (2008); Blomquist and Westerlund (2013).

inferential methods. In what comes next, we will verify our theoretical predictions in simulations.

Remark 3. *Note that the cross-section independence of $\mathbf{V}_i$ and $\boldsymbol{\varepsilon}_i$ is not required if we know that the slopes are heterogeneous. Independence is needed to implement CS bootstrap, but under heterogeneity it is not strictly needed. We can therefore relax this assumption along the lines of Pesaran and Tosetti (2011) by requiring that $\mathbf{U}_t = (\mathbf{M}_N \otimes \mathbf{I}_{k+1})\boldsymbol{\xi}_t$, 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 time-dependence requirements of Assumption 1. Here, $\mathbf{M}_N$ is an $N \times N$ "network matrix" with bounded row and column norms.*

4 Monte Carlo Simulations

4.1 Design

We utilize the simulation design largely similar to the one in De Vos and Stauskas (2024). Particularly, 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 $\mathbf{F}_\mathbf{y}$ and $\mathbf{F}_\mathbf{x}$ drive $\mathbf{y}_i$ and $\mathbf{X}_i$, respectively. With $m_x < k$ and $m_y > 1$ we reflect our main theoretical insights, where the rank condition of $\overline{\mathbf{C}}$ fails and $\mathbf{F}_\mathbf{y}$ cannot be estimated from a single CA. Hence, only $\text{rk}(\overline{\boldsymbol{\Gamma}}) = m_x$. Moreover, we induce a correlation of $\rho_f = corr(\mathbf{F}_\mathbf{y}, \mathbf{F}_\mathbf{x}) \in (0.3, 0.7)$ between them. 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 simulate the correlated loadings, we let $\widetilde{\mathbf{C}} = [\gamma_i, \boldsymbol{\Gamma}_i] = \widetilde{\mathbf{C}} + \widetilde{\boldsymbol{\eta}}_i \boldsymbol{\iota}_{1+k}'$, 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 $\mathbf{F}_\mathbf{x}$ from the CAs, we also regulate their informativeness through the population mean $\boldsymbol{\Gamma}$. It is controlled through $d = \det(\boldsymbol{\Gamma}\boldsymbol{\Gamma}')$, and we generate given an upper bound d^u the entries in $\boldsymbol{\Gamma}$ independently from $\mathcal{U}[0,2]$ such that $d^u - 0.1 \leq d \leq d^u$. The obtained $\boldsymbol{\Gamma}$ is then fixed over Monte Carlo replications and sample sizes. We take $d = 10$ as our baseline scenario with a standard information content, and study the impact of a less informative setting by lowering d 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 let the slope population mean be $\beta = 1$.

We examine performance of $\text{CCEP}_\mathbf{A}$ and $\text{CCEMG}_\mathbf{A}$ in the experiments with $\mathbf{A}$ subscript referring to the used specification of the CAs. We include 4 different 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) $\mathbf{A} = \widehat{\mathbf{z}}$ is the IC selection based on Margaritella and Westerlund (2023), and 4) $\mathbf{A} = \widehat{\mathbf{x}}$ with the selection from (3.1).

[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 $\mathbb{E}(d) \approx 9.2$.

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

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. Empirical size is at the 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 (2.8). 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 perfectly serves as a specific alternative to CCE method. Clearly, the 2SIV thus achieves the same goal as the CS-bootstrap, therefore comparisons will be informative for practice. 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.

4.2 Results: Homogeneous Slopes

To begin with, it is clear that that standard asymptotic t-tests with CCEP cannot be trusted if the factors are distinct. Particularly, Table 1 reveals the near-zero size for all asymptotic t-tests with CCEP. This occurs, because the standard errors in (2.6) are inconsistent in such case, and inference needs to be aided by the means of bootstrap. Moreover, it is evident from the relatively poor bias and size of CCEP$_{\widehat{\mathbf{z}}}$ that the IC of Margaritella and Westerlund (2023) selects CA that are inconsistent for the $\mathbf{X}$-specific factors. The boot$_{\widehat{\mathbf{z}}}$ correction has equally poor properties, which is explained by the bootstrap inability to save the estimator that is inconsistent in the first place. However, bootstrap inference for the $\mathbf{A} \in \{\mathbf{x}, \mathbf{x}_{inf}, \widehat{\mathbf{x}}\}$ specifications 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 the size closer to the nominal one. As demonstrated in Theorem 2 (a), the size distortions in case of boot$_{\mathbf{x}}$ are caused by $m_x < g$ condition, whereas the bootstrap was shown to be consistent in part (b). This is illustrated by boot$_{\mathbf{x}_{inf}}$ for $m_x = g$ selections. Results suggest, however, that the distortions for $m_x < g$ are not large and they have a fairly minor effect on testing. The IC criterion in (3.1) can also clearly estimate the optimal set of averages for which $m_x = g$, at least given sufficiently large T.[7] Clearly, the boot$_{\widehat{\mathbf{x}}}$ achieves practically the same bias and empirical size as boot$_{\mathbf{x}_{inf}}$ when $T > 100$. This confirms a great effectiveness of the combination of CS-bootstrap and the IC selector in the distinct factor case. Ultimately, we see that the 2SIV estimator achieves a close-to-nominal size for sufficiently large T, but the bootstrap tests are 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 $\mathbf{F}_{\mathbf{x}}$ and $\mathbf{F}_y$ is stronger. As before, performance of the bootstrap is practically unaffected, whereas the 2SIV suffers some size distortions for $T < 100$.

[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 ($[\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2]$).

Table 1: High dependence non-common factors

	$N \backslash T$	$\sqrt{NT} \times bias$				*size*			
		25	50	100	500	25	50	100	500
CCEP$_{\widehat{\mathbf{z}}}$	25	-0.25	-0.27	-0.57	-0.46	0.15	0.11	0.14	0.09
	50	-0.07	-0.10	-0.53	-0.45	0.09	0.07	0.12	0.08
	100	-0.06	-0.09	-0.76	-0.39	0.10	0.08	0.19	0.09
	500	-0.12	0.08	-2.01	-3.82	0.08	0.11	0.37	0.51
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
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
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
boot$_{\widehat{\mathbf{z}}}$	25	-0.28	-0.28	-0.33	-0.54	0.12	0.10	0.07	0.08
	50	-0.10	-0.07	-0.17	-0.34	0.07	0.05	0.05	0.06
	100	-0.08	-0.08	-0.27	-0.18	0.06	0.06	0.06	0.06
	500	-0.09	0.13	-1.54	-2.55	0.04	0.06	0.23	0.34
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
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
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: $(d_u, \beta, \sigma^2, \sigma^2_\eta, \sigma^2_v, \rho_f) = (10, 1, 1, 1, 0, 0.7)$. 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 $\mathbf{F}_\mathbf{x}$. An $\mathbf{A} \in \{\widehat{\mathbf{z}}, \widehat{\mathbf{x}}\}$ subscript denotes CCE specifications with CA selected from the IC criterion of Margaritella and Westerlund (2023) and (3.1), respectively. $\mathbf{A} = \mathbf{x}_{inf}$ is the infeasible CCEP specification with the optimal $g = 2$ averages from $\overline{\mathbf{X}}$ (optimal in terms of their information content on $\mathbf{F}_\mathbf{x}$). These are $[\overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2]$. Size reported for boot$_\mathbf{A}$ estimators are for the bootstrap interval in (2.8).

4.3 Results: Heterogeneous Slopes

We begin with the CCEP estimator. Our immediate focus is on the plain CCEP$_\mathbf{x}$ because the key message of Theorem 3 is its robustness to the distinct factors case. Table 2 corroborates this. We see that the estimator is virtually unbiased for all the combinations of larger N and T, and only for $N = 25$ we obtain a minimal bias. However, even then it is substantially smaller than in the CCEP case under homogeneous slopes documented in Table 1. The bias results essentially carry over when we employ the infeasible selection of CAs (CCEP$_{\mathbf{x}_{inf}}$), where $g = m_x$. For both $\mathbf{A} \in \{\mathbf{x}, \mathbf{x}_{inf}\}$, the empirical size is similar and it revolves closely around the nominal 0.05 level for all the (N, T) combinations with the exception of $N = 25$. This can be partially explained by the large N that CCE generally needs 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 boot$_{\mathbf{x}_{inf}}$ is almost identical to 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 CCEP$_{\mathbf{A}}$ and boot$_{\mathbf{A}}$ for all versions of $\mathbf{A}$ perform very similarly to 2SIV of Cui et al. (2022), which is constructed to accommodate the distinct factor case. In fact, we see that the plain CCEP estimator showcases a better performance in terms of the 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)

		$\sqrt{NT} \times bias$				*size*			
	$N \backslash T$	25	50	100	500	25	50	100	500
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
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
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
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 boot$_{\mathbf{A}}$ estimators are for the bootstrap interval in (2.8).

We further move on to Table 3, where the results for the CCEMG estimator under heterogeneous slopes are depicted. The overall results are fairly similar to the CCEP case, especially when it comes the bias. 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 contributed to the fact that a large N is needed to approximate the factor space, and $\boldsymbol{\nu}_i$ comes from a chi-squared distribution, therefore, the first component of (3.5) is not normal for any finite sample size. This suggests that the results of Theorem 4 are borne out well. Plus, in comparison to the CCEP case, we can see smaller size distortions for $N = 25$ and $T \geq 100$ across the board in both original and bootstrap samples. Moreover, 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 does not exhibit a clear size advantage anymore, at least in small samples.

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

	N \ 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: $(d_u, \beta, \sigma^2, \sigma^2_\eta, \sigma^2_v, \rho_f) = (5, 1, 1, 1, 1, 0.7)$. 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 (2.8).

5 Conclusions

In this study we considered a CCE problem, which is likely to often occur in practice. When the dependent and explanatory variables are driven by two distinct sets of factors, their cross-section averages are not consistent for the space spanned by the factors, unless the number of factors underlying the dependent variable is equal to 1. To circumvent this problem, we develop a toolbox that is a CCE-equivalent 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 the bootstrap consistency and show that the algorithm and the 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 the 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.

Blomquist, J. and Westerlund, J. (2013). Testing slope homogeneity in large panels with serial correlation. *Economics Letters*, 121(3):374–378.

Chudik, A., Mohaddes, K., Pesaran, M. H., and Raissi, M. (2017). Is there a debt-threshold effect on output growth? *Review of Economics and Statistics*, 99(1):135–150.

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. (2023). A method to evaluate the rank condition for cce estimators. *Econometric Reviews*, pages 1–33.

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.

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 Presbitero, A. F. (2015). Public debt and growth: Heterogeneity and non-linearity. *Journal of international Economics*, 97(1):45–58.

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.

Juodis, A., Karabiyik, H., and Westerlund, J. (2021). On the robustness of the pooled cce estimator. *Journal of Econometrics*, 220(2):325–348.

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.

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.

Pesaran, M. H. and Yamagata, T. (2008). Testing slope homogeneity in large panels. *Journal of econometrics*, 142(1):50–93.

Phillips, P. C. and Moon, H. R. (1999). Linear regression limit theory for nonstationary panel data. *Econometrica*, 67(5):1057–1111.

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]

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

February 15, 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.

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 $\mathbf{A}$, $\mathrm{rk}(\mathbf{A})$ for its rank, $\det(\mathbf{A})$ for the determinant and let $\|\mathbf{A}\| = \sqrt{\mathrm{tr}\,(\mathbf{A}'\mathbf{A})}$ 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 $\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x},y}} = \mathrm{plim}_{T\to\infty} T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}$ 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) $\mathrm{rk}(\overline{\boldsymbol{\Gamma}}\mathbf{q}_{\ddot{\mathbf{x}}}) = m$, 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 $\widehat{\mathbf{F}}_{\dot{\mathbf{x}}} = \overline{\mathbf{Z}}\mathbf{q}_{\dot{\mathbf{x}}} = \overline{\mathbf{X}}\mathbf{q}_{\ddot{\mathbf{x}}}$, where $\overline{\mathbf{Z}} = [\overline{\mathbf{y}}, \overline{\mathbf{X}}]$ is the full set of available CAs and let $\mathbf{q}_{\dot{\mathbf{x}}} = [\mathbf{0}_{g\times 1}, \mathbf{q}_{\ddot{\mathbf{x}}}']'$ be a $(1+k) \times g$ selection matrix that picks g cross-section averages determined by $\mathbf{q}_{\ddot{\mathbf{x}}}$ (a $k \times g$ matrix) exclusively from $\overline{\mathbf{X}}$, such that

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

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 $\|\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\| = O_p(N^{-1/2})$ for the fixed T, we have

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

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

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

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

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

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

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

such that $\mathbf{F}_{\dot{\mathbf{x}}}^0 = [\mathbf{F}_{\mathbf{x}}, \mathbf{0}_{T\times(g-m_x)}]$ and $\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}}} = [\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}},m_x}, \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\mathbf{T}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{\dot{\mathbf{x}},-m_x}]$. 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 $\boldsymbol{\Sigma}_{\mathbf{F}_{\mathbf{x}}}$, 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

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

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 $\mathbf{W}_b'\mathbf{W}_b = \mathbf{w}'\mathbf{w} \otimes \mathbf{I}_b = \operatorname{diag}(\mathbf{s} \otimes \boldsymbol{\iota}_b')$, because $\mathbf{w}'\mathbf{w} = \operatorname{diag}(\mathbf{s})$. 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 $\overline{\boldsymbol{\Gamma}}_w = \frac{1}{N}\sum_{i=1}^N s_i \boldsymbol{\Gamma}_i$ and $\overline{\mathbf{V}}_w = \frac{1}{N}\sum_{i=1}^N s_i \mathbf{V}_i$. By implementing the selection of the averages, we get

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

This representation ensures that $\overline{\boldsymbol{\Gamma}}_{w,,\ddot{\mathbf{x}}} \to_{p^*} \boldsymbol{\Gamma}_{\ddot{\mathbf{x}}}$ as $N \to \infty$, and in turn $\overline{\boldsymbol{\Gamma}}^+_{w,\ddot{\mathbf{x}}} \to_{p^*} \boldsymbol{\Gamma}^+_{\ddot{\mathbf{x}}}$. 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 $m_x < g$:

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

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

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

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

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

where $\mathbf{h}_2(\widetilde{\mathbf{I}}_{\ddot{\mathbf{x}}})$ involves the terms depending on $(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\mathbf{x}}\widehat{\mathbf{F}}^0_{\mathbf{x}})^+ - \boldsymbol{\Sigma}^+_{\mathbf{F}^0_{\dot{\mathbf{x}},v}}$, 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

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

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

(b) If $m_x = g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \to_d \mathcal{N}\left(\mathbf{0}_{k\times1}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Psi} + \widetilde{\boldsymbol{\Psi}}_f)\boldsymbol{\Sigma}^{-1}\right) + \sqrt{\tau}\boldsymbol{\Sigma}^{-1}\widetilde{\mathbf{h}}_1,$$

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 $m_x < g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}^{*}_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \to_{d^*} \mathcal{N}(\mathbf{0}_{k\times1}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Psi} + \boldsymbol{\Psi}_f)\boldsymbol{\Sigma}^{-1}) + \boldsymbol{\Sigma}^{-1}(\sqrt{\tau}\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}^{+})$$

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

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

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

(b) If $m_x = g$:

$$\sqrt{NT}(\widehat{\boldsymbol{\beta}}^{*}_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \to_{d^*} \mathcal{N}\left(\mathbf{0}_{k\times1}, \boldsymbol{\Sigma}^{-1}(\boldsymbol{\Psi} + \widetilde{\boldsymbol{\Psi}}_f)\boldsymbol{\Sigma}^{-1}\right) + \sqrt{\tau}\boldsymbol{\Sigma}^{-1}\widetilde{\mathbf{h}}_1,$$

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

$$\sup_{x\in\mathbb{R}^{k\times1}}\left|\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^{*}_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \le x] - \mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \le x]\right| \to_p 0,$$

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

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

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,

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

Note that $\mathbf{IV}$ is algebraically equal to $\mathbf{0}_k$ if $\mathbf{q}_{\ddot{\mathbf{x}}} = \mathbf{I}_k$. 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} \\
&- \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} \\
&- \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 $\left\| \sqrt{N} T^{-1} \overline{\mathbf{V}}_{\dot{\mathbf{x}}}' \mathbf{F}_{\dot{\mathbf{x}}}^{0} \right\| = O_p(T^{-1/2})$ and

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

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

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

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

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

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},
\end{aligned} \tag{3.17}
$$

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

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

since

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

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:

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

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,

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

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. Therefore, in summary

$$
\|\mathbf{C1}\| = \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)' T^{-1}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x}(T^{-1}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{\ddot{\mathbf{x}},-m_x})^{+}\overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},-m_x}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i}\right\| = O_p(T^{-1/2}).
\tag{3.22}
$$

We next move on to **C5**:

$$
\begin{aligned}
\mathbf{C5} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)' T^{-1}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}\left[(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}})^{+}-\boldsymbol{\Sigma}^{+}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right]\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i' T^{-1}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}\left[(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}})^{+}-\boldsymbol{\Sigma}^{+}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right]\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&\quad - \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}' T^{-1}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}\left[(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}})^{+}-\boldsymbol{\Sigma}^{+}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right]\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} \\
&= \frac{1}{\sqrt{N}}\sum_{i=1}^{N} T^{-1}\mathbf{V}_i' T^{-1}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}\left[(T^{-1}\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}})^{+}-\boldsymbol{\Sigma}^{+}_{\mathbf{F}^0_{\ddot{\mathbf{x}},v}}\right]\widehat{\mathbf{F}}^{0\prime}_{\ddot{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\eta}_{\gamma,i} + O_p(T^{-1/2}) + O_p(N^{-1/2})
\end{aligned}
\tag{3.23}
$$

since

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

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

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

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

$$\begin{aligned}
&\mathbb{E} \left(\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\boldsymbol{\eta}_{\gamma,i}' T^{-1} \mathbf{F}_{\mathbf{y}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0} \otimes T^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0} \right) \right\|^2 \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbb{E} \left(\operatorname{tr} \left[\boldsymbol{\eta}_{\gamma,i}' T^{-1} \mathbf{F}_{\mathbf{y}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime} \mathbf{F}_{\mathbf{y}}) \boldsymbol{\eta}_{\gamma,j} \otimes T^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime} \mathbf{V}_j) \right] \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left(\operatorname{tr} \left[\boldsymbol{\eta}_{\gamma,i}' T^{-1} \mathbf{F}_{\mathbf{y}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime} \mathbf{F}_{\mathbf{y}}) \boldsymbol{\eta}_{\gamma,i} \otimes T^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime} \mathbf{V}_i) \right] \right) \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathbb{E} \left(\boldsymbol{\eta}_{\gamma,i}' T^{-1} \mathbf{F}_{\mathbf{y}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime} \mathbf{F}_{\mathbf{y}}) \boldsymbol{\eta}_{\gamma,i} \operatorname{tr} \left[T^{-1} \mathbf{V}_i' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime} \mathbf{V}_i) \right] \right) \\
&= O(N^{-1}) + O(T^{-1}),
\end{aligned} \tag{3.26}$$

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

$$\begin{aligned}
\mathbf{C5} &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\mathbf{\Gamma}}_{\dot{\mathbf{x}}}^{+} \mathbf{\Gamma}_i)' T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0} \left[(T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0})^{+} - \mathbf{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^{0}}^{+} \right] \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}).
\end{aligned} \tag{3.27}$$

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

$$
\begin{aligned}
\|\mathbf{C2}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \overline{\mathbf{V}}^{0}_{\ddot{\mathbf{x}},m_x} (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+} \overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&\leq \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+} \right\| \left\| \sqrt{N} T^{-1} \overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \mathbf{F}_{\mathbf{y}} \right\| \frac{1}{N} \sum_{i=1}^{N} \left\| T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' \overline{\mathbf{V}}^{0}_{\ddot{\mathbf{x}},m_x} \right\| \left\| \boldsymbol{\eta}_{\gamma,i} \right\| \\
&= O_p(T^{-1/2}) \left(O_p(N^{-1}) + O_p((NT)^{-1/2}) \right) \\
&= O_p(N^{-1}T^{-1/2}) + O_p(N^{-1/2}T^{-1}) \qquad (3.28)
\end{aligned}
$$

and

$$
\begin{aligned}
\|\mathbf{C3}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \overline{\mathbf{V}}^{0}_{\ddot{\mathbf{x}},m_x} (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&\leq \left\| T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}} \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+} \right\| \frac{1}{N} \sum_{i=1}^{N} \left\| T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' \sqrt{N} \overline{\mathbf{V}}^{0}_{\ddot{\mathbf{x}},m_x} \right\| \left\| \boldsymbol{\eta}_{\gamma,i} \right\| \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}), \qquad (3.29)
\end{aligned}
$$

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,

$$
\begin{aligned}
\|\mathbf{C4}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \mathbf{F}_{\mathbf{x}} (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+} \overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&\leq \left\| \sqrt{N} T^{-1} \overline{\mathbf{V}}^{0\prime}_{\ddot{\mathbf{x}},m_x} \mathbf{F}_{\mathbf{y}} \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+} \right\| \frac{1}{N} \sum_{i=1}^{N} \left\| T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' \mathbf{F}_{\mathbf{x}} \right\| \left\| \boldsymbol{\eta}_{\gamma,i} \right\| \\
&= O_p(T^{-1/2}) \left(O_p(T^{-1/2}) + O_p((NT)^{-1/2}) \right) \\
&= O_p(T^{-1}). \qquad (3.30)
\end{aligned}
$$

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

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

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

$$
\begin{aligned}
\|\mathbf{II}\| &= \left\| \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\mathbf{x}}} \mathbf{X}_i \right)^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\mathbf{x}}} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\| \\
&\leq \underbrace{\left\| \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\mathbf{x}}} \mathbf{X}_i \right)^{-1} \right\|}_{O_p(1)} \underbrace{\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\mathbf{x}}} \mathbf{F}_{\mathbf{y}} \boldsymbol{\eta}_{\gamma,i} \right\|}_{O_p(N^{-1/2}) + O_p(T^{-1/2})} \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}). \qquad (3.32)
\end{aligned}
$$

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. This gives, respectively,

$$
\begin{aligned}
\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\gamma\right\|^2\right) &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left(T^{-2}\mathbf{tr}\left[\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\gamma\gamma'\mathbf{F}_{\mathbf{y}}'\mathbf{V}_j\right]\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(T^{-2}\mathbf{tr}\left[\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\gamma\gamma'\mathbf{F}_{\mathbf{y}}'\mathbf{V}_i\right]\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(T^{-2}\mathbf{tr}\left[\gamma'\mathbf{F}_{\mathbf{y}}'\mathbf{V}_i\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\gamma\right]\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{tr}\left[\gamma\gamma'\mathbb{E}(T^{-2}\mathbf{F}_{\mathbf{y}}'\mathbf{V}_i\mathbf{V}_i'\mathbf{F}_{\mathbf{y}})\right]\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{tr}\left[\gamma\gamma'\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbb{E}(\mathbf{f}_{\mathbf{y},t}\mathbf{v}_{i,t}'\mathbf{v}_{i,s}\mathbf{f}_{\mathbf{y},s}')\right]\right)\\
&= O(T^{-1}) \qquad (3.33)
\end{aligned}
$$

and similarly by cross-section independence

$$
\begin{aligned}
\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\gamma'\otimes T^{-1}\mathbf{V}_i'\mathbf{F}_{\dot{\mathbf{x}}}^0\right)\right\|^2\right) &= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\mathbf{tr}\left[\gamma'\gamma\otimes T^{-2}\mathbf{V}_i'\mathbf{F}_{\dot{\mathbf{x}}}^0\mathbf{F}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{V}_i\right]\right)\\
&= \frac{1}{N}\sum_{i=1}^{N}\gamma'\gamma\mathbf{tr}\left[\mathbb{E}\left(T^{-2}\mathbf{V}_i'\mathbf{F}_{\dot{\mathbf{x}}}^0\mathbf{F}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{V}_i\right)\right]\\
&= \gamma'\gamma\frac{1}{N}\sum_{i=1}^{N}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbf{tr}\left[\mathbb{E}\left(\mathbf{v}_{i,t}\mathbf{f}_{\dot{\mathbf{x}},t}^{0\prime}\mathbf{f}_{\dot{\mathbf{x}},s}^0\mathbf{v}_{i,s}'\right)\right]\\
&= O(T^{-1}). \qquad (3.34)
\end{aligned}
$$

The two exceptions are (3.25) and (3.21), which slightly change. In particular,

$$
\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0)^+ - \mathbf{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^0}^+\right]\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma\right\|\\
&= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\gamma'T^{-1}\mathbf{F}_{\mathbf{y}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0\otimes T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0\right)\right\|\left\|\mathrm{vec}\left(\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0)^+ - \mathbf{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^0}^+\right]\right)\right\|\\
&\leq \left\|\gamma'T^{-1}\mathbf{F}_{\mathbf{y}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0\otimes\sqrt{N}T^{-1}\overline{\mathbf{V}}'\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0\right\|\left\|\mathrm{vec}\left(\underbrace{\left[(T^{-1}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0)^+ - \mathbf{\Sigma}_{\mathbf{F}_{\dot{\mathbf{x}},v}^0}^+\right]}_{O_p(N^{-1/2})+O_p(T^{-1/2})}\right)\right\|\\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}), \qquad (3.35)
\end{aligned}
$$

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

$$
\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}T^{-1}\mathbf{V}_i'T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^+\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma\right\|\\
&= \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\widehat{\mathbf{\Sigma}}_{\mathbf{v}_{\ddot{\mathbf{x}},-m_x}^0}^+T^{-1}(\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)'\mathbf{F}_y\gamma\right\|\\
&\leq \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\right\|\left\|\widehat{\mathbf{\Sigma}}_{\mathbf{v}_{\ddot{\mathbf{x}},-m_x}^0}^+\right\|\left\|T^{-1}(\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)'\mathbf{F}_y\gamma\right\| = O_p(T^{-1/2}). \qquad (3.36)
\end{aligned}
$$

This means that

$$\begin{aligned}
\|\mathbf{IV}\| &\leq \underbrace{\left\| \left(\frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} \right\|}_{O_p(1)} \underbrace{\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma \right\|}_{O_p(N^{-1/2}) + O_p(T^{-1/2})} \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}). \qquad (3.37)
\end{aligned}$$

By putting the results together, we simplify (3.8) and obtain the asymptotic distribution by standard Lindeberg-Lévy Central Limit Theorem:

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) &= \boldsymbol{\Sigma}^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (T^{-1} \mathbf{V}_i' \mathbf{V}_i) \boldsymbol{\nu}_i + o_p(1) \\
&= \boldsymbol{\Sigma}^{-1} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\Sigma}_i \boldsymbol{\nu}_i + o_p(1) \\
&\to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1} \boldsymbol{\Psi}_{\nu} \boldsymbol{\Sigma}^{-1} \right) \qquad (3.38)
\end{aligned}$$

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

$$\begin{aligned}
\mathbb{E}\left(\left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[(T^{-1} \mathbf{V}_i' \mathbf{V}_i) - \boldsymbol{\Sigma}_i \right] \boldsymbol{\nu}_i \right\|^2 \right) &= \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathrm{tr} \left[\mathbb{E}\left(\left[(T^{-1} \mathbf{V}_i' \mathbf{V}_i) - \boldsymbol{\Sigma}_i \right] \boldsymbol{\nu}_i \boldsymbol{\nu}_j \left[T^{-1} \mathbf{V}_j' \mathbf{V}_j) - \boldsymbol{\Sigma}_j \right]' \right) \right] \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathrm{tr} \left[\mathbb{E}\left(\left[(T^{-1} \mathbf{V}_i' \mathbf{V}_i) - \boldsymbol{\Sigma}_i \right] \boldsymbol{\Omega}_{\nu} \left[T^{-1} \mathbf{V}_i' \mathbf{V}_i) - \boldsymbol{\Sigma}_i \right]' \right) \right] \\
&= \frac{1}{N} \sum_{i=1}^{N} \mathrm{tr} \left[\boldsymbol{\Omega}_{\nu} \mathbb{E}\left(\left[(T^{-1} \mathbf{V}_i' \mathbf{V}_i) - \boldsymbol{\Sigma}_i \right]' \left[(T^{-1} \mathbf{V}_i' \mathbf{V}_i) - \boldsymbol{\Sigma}_i \right] \right) \right] \\
&= O(T^{-1}). \qquad (3.39)
\end{aligned}$$

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

$$\begin{aligned}
\mathbf{M}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}} = \mathbf{M}_{\mathbf{F}_{\mathbf{x}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} &= T^{-1} \overline{\mathbf{V}}_{\dot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+} \overline{\mathbf{V}}_{\dot{\mathbf{x}}}' + T^{-1} \overline{\mathbf{V}}_{\dot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}}' \\
&+ T^{-1} \mathbf{F}_{\mathbf{x}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+} \overline{\mathbf{V}}_{\dot{\mathbf{x}}}' + T^{-1} \mathbf{F}_{\mathbf{x}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}} [(T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+} - (\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}' T^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{x}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}})^{+}] \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}}'. \qquad (3.40)
\end{aligned}$$

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 $\mathbf{C}$ in (3.9) as the analysis of $\mathbf{A}$ and $\mathbf{B}$ 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{3.41}
$$

which is driven by the highest order component

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

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

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

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:

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} &= \frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&= \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}(\mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \boldsymbol{\varepsilon}_i) \\
&= \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\beta}_i + \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \\
&= \boldsymbol{\beta} + \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\nu}_i + \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \frac{1}{N}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i,
\end{aligned} \tag{3.44}
$$

which implies that

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

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

$$
\mathbf{I} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\nu}_i \to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Omega}_{\nu}\right), \tag{3.46}
$$

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, $\bar{\varepsilon}$ 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

$$
\begin{aligned}
&\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\|^2\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left(\mathrm{tr}\left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-2}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_j'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_j\left(T^{-1}\mathbf{X}_j'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_j\right)^{-1}\right]\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\mathrm{tr}\left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-2}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right]\right) = O(T^{-1}),
\end{aligned} \tag{3.47}
$$

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)

$$
\begin{aligned}
T^{-1/2}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i &= T^{-1/2}(\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \\
&= T^{-1/2}(\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\boldsymbol{\varepsilon}_i - T^{-1/2}(\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\mathbf{P}_{\mathbf{F}_{\dot{\mathbf{x}}}^0}\boldsymbol{\varepsilon}_i \\
&\quad - T^{-1/2}(\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}_{\dot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^0})\boldsymbol{\varepsilon}_i,
\end{aligned} \tag{3.48}
$$

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,

$$\left\| T^{-1/2}\mathbf{V}_i'\boldsymbol{\varepsilon}_i \right\| = O_p(1), \tag{3.49}$$

$$\left\| T^{-1/2}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\boldsymbol{\varepsilon}_i \right\| = O_p(N^{-1/2}) \tag{3.50}$$

under our assumptions. Next,

$$\left\| T^{-1/2}\mathbf{V}_i'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\boldsymbol{\varepsilon}_i \right\| \leq \left\| T^{-1/2}\mathbf{V}_i'\mathbf{F}_{\ddot{\mathbf{x}}}^0 \right\| \left\| (T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\ddot{\mathbf{x}}}^0)^+ \right\| \left\| T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\boldsymbol{\varepsilon}_i \right\| = O_p(T^{-1/2}), \tag{3.51}$$

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

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

$$\begin{aligned}
\left\| T^{-1/2}\mathbf{V}_i'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\boldsymbol{\varepsilon}_i \right\| \leq & \left\| T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0 \right\| \left\| (T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^+ \right\| \left\| T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\boldsymbol{\varepsilon}_i \right\| \\
& + \left\| T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\boldsymbol{\varepsilon}_i \right\| \\
& + \left\| T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1/2}\mathbf{F}_{\mathbf{x}}'\boldsymbol{\varepsilon}_i \right\| \\
& + \left\| T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{x}} \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1/2}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\boldsymbol{\varepsilon}_i \right\| \\
& + \left\| T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0 \right\| \left\| \left[(T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0)^+ - \boldsymbol{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^+ \right] \right\| \left\| T^{-1/2}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\boldsymbol{\varepsilon}_i \right\| \\
& = O_p(N^{-1/2}) + O_p(T^{-1/2}),
\end{aligned} \tag{3.53}$$

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{3.54}$$

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,

$$\left\| T^{-1/2}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \right\| = O_p(1), \tag{3.55}$$

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

and hence

$$\|\mathbf{III}\| = \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{X}_i \right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \right\| = O_p(N^{-1/2}) + O_p(T^{-1/2}). \tag{3.57}$$

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

$$\begin{aligned}\mathbf{II} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \boldsymbol{\Sigma}_i^{-1}\right]T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\ &= \mathbf{A} + \mathbf{B},\end{aligned} \tag{3.58}$$

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

$$\begin{aligned}\mathbf{A} &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\ &= \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\mathbf{F}_{\mathbf{y}}\gamma_i - \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\mathbf{F}_{\mathbf{y}}\gamma_i \\ &\quad - \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i \\ &= \mathbf{A_1} - \mathbf{A_2} - \mathbf{A_3},\end{aligned} \tag{3.59}$$

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

$$\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \leq \left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{F}_{\mathbf{y}}\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Sigma}_i^{-1}\right\|\left\|\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i\right\|\|\gamma_i\| = O_p(T^{-1/2}), \tag{3.60}$$

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

$$\begin{aligned}\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\gamma_i\right\|^2\right) &= \frac{1}{N}\sum_{i=1}^{N}\mathrm{tr}\left[\mathbb{E}\left(\gamma_i\gamma_i'\right)\mathbb{E}\left(\boldsymbol{\Sigma}_i^{-1}T^{-2}\mathbf{V}_i'\mathbf{F}_{\mathbf{y}}\mathbf{F}_{\mathbf{y}}'\mathbf{V}_i\boldsymbol{\Sigma}_i^{-1}\right)\right] \\ &= O(T^{-1})\end{aligned} \tag{3.61}$$

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

$$\begin{aligned}&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Sigma}_i^{-1}T^{-1}\boldsymbol{\Gamma}_i'\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'\mathbf{P}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\ &\leq \left\|(T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\ddot{\mathbf{x}}}^0)^{+}\right\|\left\|T^{-1}\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\right\|\left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}'\mathbf{F}_{\ddot{\mathbf{x}}}^0\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Sigma}_i^{-1}\right\|\left\|\overline{\boldsymbol{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\boldsymbol{\Gamma}_i\right\|\|\gamma_i\| = O_p(T^{-1/2})\end{aligned} \tag{3.62}$$

and

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

where the order comes by exactly the same argument as in (3.61) by using the Kronecker properties:

$$\begin{aligned}
\mathbb{E}\left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(\gamma_i'\otimes\mathbf{\Sigma}_i^{-1}T^{-1}\mathbf{V}_i'\mathbf{F}_{\ddot{\mathbf{x}}}^0\right)\right\|^2\right) &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbb{E}\left(\mathrm{tr}\left[\gamma_i'\gamma_j\otimes\mathbf{\Sigma}_i^{-1}T^{-2}\mathbf{V}_i'\mathbf{F}_{\ddot{\mathbf{x}}}^0\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{V}_j\mathbf{\Sigma}_j^{-1}\right]\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\mathrm{tr}\left[\gamma_i'\gamma_i\otimes\mathbf{\Sigma}_i^{-1}T^{-2}\mathbf{V}_i'\mathbf{F}_{\ddot{\mathbf{x}}}^0\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{V}_i\mathbf{\Sigma}_i^{-1}\right]\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\gamma_i'\gamma_i\right)\mathrm{tr}\left[\mathbb{E}\left(\mathbf{\Sigma}_i^{-1}T^{-2}\mathbf{V}_i'\mathbf{F}_{\ddot{\mathbf{x}}}^0\mathbf{F}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{V}_i\mathbf{\Sigma}_i^{-1}\right)\right] \\
&= O(T^{-1}). \qquad (3.64)
\end{aligned}$$

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

$$\begin{aligned}
\mathbf{A3} &= \frac{1}{N}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}\sqrt{N}T^{-1}(\mathbf{V}_i-\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+}\mathbf{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}-\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \frac{1}{N}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}\sqrt{N}T^{-1}\mathbf{V}_i'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}-\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i - \frac{1}{N}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}\sqrt{N}T^{-1}\mathbf{\Gamma}_i'\overline{\mathbf{\Gamma}}_{\ddot{\mathbf{x}}}^{+\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}}}'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0}-\mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i \\
&= \mathbf{A3.1}-\mathbf{A3.2}, \qquad (3.65)
\end{aligned}$$

such that

$$\begin{aligned}
\|\mathbf{A3.1}\| \leq{}& \frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Sigma}_i^{-1}\right\|\left\|\sqrt{N}T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0\right\|\left\|(T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^{+}\right\|\left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&+\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Sigma}_i^{-1}\right\|\left\|\sqrt{N}T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0\right\|\left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+}\right\|\left\|T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&+\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Sigma}_i^{-1}\right\|\left\|\sqrt{N}T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0\right\|\left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+}\right\|\left\|T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&+\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Sigma}_i^{-1}\right\|\left\|T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{x}}\right\|\left\|(T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^{+}\right\|\left\|\sqrt{N}T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&+\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Sigma}_i^{-1}\right\|\left\|T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0\right\|\sqrt{N}\left\|\left[(T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0)^{+}-\mathbf{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^{+}\right]\right\|\left\|T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
={}& O_p(N^{-1/2})+O_p(T^{-1/2}) \qquad (3.66)
\end{aligned}$$

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\|\mathbf{\Sigma}_i^{-1}\left\|\overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}}\mathbf{\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\|\mathbf{\Sigma}_i^{-1}\right\|\left\|\overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}}\mathbf{\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\|\mathbf{\Sigma}_i^{-1}\right\|\left\|\overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}}\mathbf{\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\|\mathbf{\Sigma}_i^{-1}\right\|\left\|\overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}}\mathbf{\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\|\mathbf{\Sigma}_i^{-1}\right\|\left\|\overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}}\mathbf{\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}}})^+ - \mathbf{\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})
\end{aligned}
\tag{3.67}
$$

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

$$
\|\mathbf{A}\| = \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathbf{\Sigma}_i^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| = O_p(N^{-1/2}) + O_p(T^{-1/2}). \tag{3.68}
$$

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

$$
\begin{aligned}
\|\mathbf{B}\| &= \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \mathbf{\Sigma}_i^{-1}\right]T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&\leq \underbrace{\sqrt{N}\sup_i\left\|\left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \mathbf{\Sigma}_i^{-1}\right]\right\|}_{O_p(1)\text{ if } TN^{-1}=O(1)}\frac{1}{N}\sum_{i=1}^{N}\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&= O_p(N^{-1/2}) + O_p(T^{-1/2}),
\end{aligned}
\tag{3.69}
$$

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

$$
\begin{aligned}
\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \leq & \left\|T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}}\mathbf{\Gamma}_i)'\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + \left\|T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}}\mathbf{\Gamma}_i)'\mathbf{P}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
& + \left\|T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}}\mathbf{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}})\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
\leq & \left\|T^{-1}(\mathbf{V}_i - \overline{\mathbf{V}}_{\ddot{\mathbf{x}}}\overline{\mathbf{\Gamma}}^+_{\ddot{\mathbf{x}}}\mathbf{\Gamma}_i)'(\mathbf{M}_{\mathbf{F}^0_{\ddot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\ddot{\mathbf{x}}}})\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + O_p(T^{-1/2}),
\end{aligned}
\tag{3.70}
$$

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 (3.48), we obtain

$$
\begin{aligned}
\left\| T^{-1}\mathbf{V}_i'(\mathbf{M}_{\mathbf{F}_{\ddot{\mathbf{x}}}^0} - \mathbf{M}_{\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0})\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \leq & \underbrace{\left\| T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0 \right\|}_{O_p(N^{-1/2})+O_p(T^{-1/2})} \left\| (T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^0)^+ \right\| \underbrace{\left\| T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},-m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\|}_{O_p(T^{-1/2})} \\
& + \left\| T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& + \underbrace{\left\| T^{-1}\mathbf{V}_i'\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^0 \right\|}_{O_p(N^{-1})+O_p((NT)^{-1/2})} \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& + \left\| T^{-1}\mathbf{V}_i'\mathbf{F}_{\mathbf{x}} \right\| \left\| (T^{-1}\mathbf{F}_{\mathbf{x}}'\mathbf{F}_{\mathbf{x}})^+ \right\| \left\| T^{-1}\overline{\mathbf{V}}_{\ddot{\mathbf{x}},m_x}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& + \left\| T^{-1}\mathbf{V}_i'\widehat{\mathbf{F}}_{\mathbf{x}}^0 \right\| \left\| \left[(T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^0)^+ - \mathbf{\Sigma}_{\mathbf{F}_{\ddot{\mathbf{x}},v}^0}^+ \right] \right\| \left\| T^{-1}\widehat{\mathbf{F}}_{\ddot{\mathbf{x}}}^{0\prime}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& = O_p(T^{-1}) + O_p(N^{-1}) + O_p((NT)^{-1/2}),
\end{aligned}
\tag{3.71}
$$

with the drivers of the order indicated. In summary,

$$
\begin{aligned}
\|\mathbf{B}\| \leq & \sqrt{N}\sup_i \left\| \left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i \right)^{-1} - \mathbf{\Sigma}_i^{-1} \right] \right\| \frac{1}{N}\sum_{i=1}^N \left\| T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| \\
& = O_p(N^{-1}) + O_p(T^{-1/2}),
\end{aligned}
\tag{3.72}
$$

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

$$
\|\mathbf{II}\| = \left\| \frac{1}{\sqrt{N}}\sum_{i=1}^N \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i \right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right\| = O_p(N^{-1/2}) + O_p(T^{-1/2}),
\tag{3.73}
$$

which ultimately leads to

$$
\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) = & \frac{1}{\sqrt{N}}\sum_{i=1}^N \boldsymbol{\nu}_i + O_p(N^{-1/2}) + O_p(T^{-1/2}) \\
& \to_d \mathcal{N}\left(\mathbf{0}_{k\times 1}, \mathbf{\Omega}_{\boldsymbol{\nu}}\right)
\end{aligned}
\tag{3.74}
$$

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 (3.40). Because now the convergence rate will be quicker, (3.57) will hold as well, therefore it is sufficient to check **II** in the expansion (3.44) 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} \boldsymbol{\Sigma}_i^{-1} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' (\mathbf{M}_{\mathbf{F}^0_{\dot{\mathbf{x}}}} - \mathbf{M}_{\widehat{\mathbf{F}}^0_{\dot{\mathbf{x}}}}) \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&\leq \left\| \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\Sigma}_i^{-1} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \overline{\mathbf{V}}_{\dot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+} \overline{\mathbf{V}}_{\dot{\mathbf{x}}}' \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&+ \left\| \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\Sigma}_i^{-1} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \overline{\mathbf{V}}_{\dot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}' \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&+ \left\| \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\Sigma}_i^{-1} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \mathbf{F}_{\mathbf{x}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}} (T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+} \overline{\mathbf{V}}_{\dot{\mathbf{x}}}' \mathbf{F}_{\mathbf{y}} \gamma_i \right\| \\
&+ \left\| \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{\Sigma}_i^{-1} \sqrt{N} T^{-1} (\mathbf{V}_i - \overline{\mathbf{V}}_{\dot{\mathbf{x}}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}^{+} \boldsymbol{\Gamma}_i)' T^{-1} \mathbf{F}_{\mathbf{x}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}} [(T^{-1} \widehat{\mathbf{F}}_{\dot{\mathbf{x}}}' \widehat{\mathbf{F}}_{\dot{\mathbf{x}}})^{+} - (\overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}}' T^{-1} \mathbf{F}_{\mathbf{x}}' \mathbf{F}_{\mathbf{x}} \overline{\boldsymbol{\Gamma}}_{\dot{\mathbf{x}}})^{+}] \overline{\boldsymbol{\Gamma}}_{\dot{\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{3.75}$$

which is driven the highest order term which is almost identical to (3.42). Note that unlike in the case of $m_x < g$, we do not need to impose $TN^{-1} = O(1)$. In fact, such restriction is not needed to demonstrate that $\mathbf{B}$ term is negligible as well, since we only need to split $\mathbf{II}$ differently for $m_x = g$ case. Particularly,

$$\begin{aligned}
\mathbf{II} &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (T^{-1} \mathbf{V}_i' \mathbf{M}_{\mathbf{F}_{\mathbf{x}}} \mathbf{V}_i)^{-1} T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i \\
&+ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left[\left(T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - (T^{-1} \mathbf{V}_i' \mathbf{M}_{\mathbf{F}_{\mathbf{x}}} \mathbf{V}_i)^{-1} \right] T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i = \mathbf{A} + \mathbf{B}
\end{aligned} \tag{3.76}$$

where $\|\mathbf{A}\| = o_p(1)$ still as $\left\|(T^{-1} \mathbf{V}_i' \mathbf{M}_{\mathbf{F}_{\mathbf{x}}} \mathbf{V}_i)^{-1}\right\| = O_p(1)$ and under $m_x = g$ we have that

$$\left\| \left(T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - (T^{-1} \mathbf{V}_i' \mathbf{M}_{\mathbf{F}_{\mathbf{x}}} \mathbf{V}_i)^{-1} \right\| = o_p(N^{-1/2}) \tag{3.77}$$

and hence

$$\|\mathbf{B}\| \leq \sqrt{N} \sup_i \left\| \left(T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{X}_i \right)^{-1} - (T^{-1} \mathbf{V}_i' \mathbf{M}_{\mathbf{F}_{\mathbf{x}}} \mathbf{V}_i)^{-1} \right\| \frac{1}{N} \sum_{i=1}^{N} \left\| T^{-1} \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}} \mathbf{F}_{\mathbf{y}} \gamma_i \right\| = o_p(1). \tag{3.78}$$

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:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} &= \left(\mathbf{X}^{*\prime}\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}^*\right)^{-1}\mathbf{X}^{*\prime}\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{y}^* \\
&= \left(\mathbf{X}'\mathbf{W}_T'\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{W}_T\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}_T'\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{W}_T\mathbf{y} \\
&= \left(\mathbf{X}'\mathbf{W}_T'\mathbf{W}_T\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{W}_T'\mathbf{W}_T\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{y} \\
&= \left(\mathbf{X}'\mathrm{diag}(\mathbf{s}\otimes\boldsymbol{\iota}_T')\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}\right)^{-1}\mathbf{X}'\mathrm{diag}(\mathbf{s}\otimes\boldsymbol{\iota}_T')\underline{\mathbf{M}}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{y} \\
&= \left(\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*}\mathbf{y}_i \\
&= \boldsymbol{\beta} + \left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i + \frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right),
\end{aligned} \tag{3.79}$$

which implies that

$$\begin{aligned}
\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) &= \left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} \\
&\quad\times\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right).
\end{aligned} \tag{3.80}$$

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_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 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_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 T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\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_iT^{-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}}_{\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}}_{\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}}_{\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}}_{\mathbf{x}}}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\mathbf{x}}}\boldsymbol{\varepsilon}_i\right).
\end{aligned} \tag{3.81}$$

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 $\mathbf{Z}_1, \ldots \mathbf{Z}_n$ and multinomial weights $s_1, \ldots, s_n$ (independent from model primitives), then for a deterministic sequence a_n we have

$\triangle_n = O_{p^*}(a_n)$ *in probability* $\Leftrightarrow$ $\triangle_n = O_p(a_n)$ *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 (3.82)
\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,

$$
\begin{aligned}
\|\mathbf{III}\| &\leq \left\|\left(\frac{1}{NT}\sum_{i=1}^N s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i\right)^{-1}\right\| \left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N s_i T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\| + \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\|\right) \\
&= o_{p^*}(1) \qquad (3.83)
\end{aligned}
$$

in probability and

$$
\begin{aligned}
\mathbf{I} &= \left(\frac{1}{NT}\sum_{i=1}^N s_i \mathbf{X}_i' \mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}} \mathbf{X}_i\right)^{-1} \left(\frac{1}{\sqrt{N}}\sum_{i=1}^N s_i T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i - \frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\boldsymbol{\nu}_i\right) \\
&= \boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^N (s_i - 1)\boldsymbol{\Sigma}_i\boldsymbol{\nu}_i + O_{p^*}(T^{-1/2}) \\
&\to_{d^*} \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_{\nu}\boldsymbol{\Sigma}^{-1}\right) \qquad (3.84)
\end{aligned}
$$

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

$$\overline{\mathbf{H}}_{w,\dot{\mathbf{x}}} = [\overline{\mathbf{H}}_{w,\dot{\mathbf{x}},m_x}, \overline{\mathbf{H}}_{w,\dot{\mathbf{x}},-m_x}] = \begin{bmatrix} \overline{\mathbf{\Gamma}}^{-1}_{w,\ddot{\mathbf{x}},m_x} & -\overline{\mathbf{\Gamma}}^{-1}_{w,\ddot{\mathbf{x}},m_x}\overline{\mathbf{\Gamma}}_{w,\ddot{\mathbf{x}}-m_x} \\ \mathbf{0}_{(g-m_x)\times m_x} & \mathbf{I}_{g-m_x} \end{bmatrix}, \qquad \mathbf{D}_N = \begin{bmatrix} \mathbf{I}_{m_x} & \mathbf{0}_{m_x\times(g-m_x)} \\ \mathbf{0}_{(g-m_x)\times m_x} & \sqrt{N}\mathbf{I}_{g-m_x} \end{bmatrix} \tag{3.85}$$

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

$$\widehat{\mathbf{F}}^{0*}_{\dot{\mathbf{x}}} = \widehat{\mathbf{F}}^{*}_{\dot{\mathbf{x}}}\overline{\mathbf{H}}_{w,\dot{\mathbf{x}}}\mathbf{D}_N = \mathbf{F}^0_{\dot{\mathbf{x}}} + [\overline{\mathbf{V}}_{w,\dot{\mathbf{x}}}\overline{\mathbf{H}}_{w,\dot{\mathbf{x}},m_x}, \sqrt{N}\overline{\mathbf{V}}_{w,\dot{\mathbf{x}}}\overline{\mathbf{H}}_{w,\dot{\mathbf{x}},-m_x}] = \mathbf{F}^0_{\dot{\mathbf{x}}} + [\overline{\mathbf{V}}^0_{w,\dot{\mathbf{x}},m_x}, \overline{\mathbf{V}}^0_{w,\ddot{\mathbf{x}},-m_x}]. \tag{3.86}$$

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

$$(1)\quad \left\|\overline{\mathbf{V}}_{w,\ddot{\mathbf{x}}}\right\| = O_{p^*}(N^{-1/2}), \tag{3.87}$$

$$(2)\quad \left\|T^{-1}\overline{\mathbf{V}}'_{w,\ddot{\mathbf{x}}}\mathbf{V}_i\right\| = O_{p^*}(N^{-1}) + O_{p^*}((NT)^{-1/2}), \tag{3.88}$$

$$(3)\quad \left\|\left(T^{-1}\widehat{\mathbf{F}}^{0*\prime}_{\dot{\mathbf{x}}}\widehat{\mathbf{F}}^{0*}_{\dot{\mathbf{x}}}\right)^+ - \mathbf{\Sigma}^+_{w,\mathbf{F}^0_{\dot{\mathbf{x}},v}}\right\| = O_{p^*}(N^{-1/2}) + O_{p^*}(T^{-1/2}), \tag{3.89}$$

$$(4)\quad \mathbb{E}(s_i) = 1, \tag{3.90}$$

$$(5)\quad \mathrm{V}ar(s_i) = \mathbb{E}[(s_i-1)^2] = 1 - N^{-1} \ \ \textit{(multinomial variance)} \tag{3.91}$$

where

$$\mathbf{\Sigma}_{w,\mathbf{F}^0_{\dot{\mathbf{x}}},v} = \operatorname{diag}\left[\mathbf{\Sigma}_{\mathbf{F}_{\mathbf{x}}}, (T^{-1}\overline{\mathbf{V}}^{0\prime}_{w,\dot{\mathbf{x}},-m_x}\overline{\mathbf{V}}^0_{w,\dot{\mathbf{x}},-m_x})\right]. \tag{3.92}$$

Therefore,

$$\begin{aligned} \|\mathbf{II}\| &= \left\| \left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} \left(\frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i - \frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right)\right\| \\ &\leq \left\| \left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\| \left\|\left(\frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i - \frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right)\right\| \\ &\leq \underbrace{\left\| \left(\frac{1}{NT}\sum_{i=1}^N s_i\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\|}_{O_{p^*}(1)} \left(\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|\right) \\ &= O_{p^*}(N^{-1/2}) + O_{p^*}(T^{-1/2}). \end{aligned} \tag{3.93}$$

Note how (3.91) 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,

$$\begin{aligned} \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N s_iT^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| &\leq \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N (s_i-1)T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| + \left\|\frac{1}{\sqrt{N}}\sum_{i=1}^N T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\ &= O_{p^*}(N^{-1/2}) + O_{p^*}(T^{-1/2}). \end{aligned} \tag{3.94}$$

In summary, we obtain

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) = \boldsymbol{\Sigma}^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(s_i-1)\boldsymbol{\Sigma}_i\boldsymbol{\nu}_i + o_{p^*}(1)$$
$$\rightarrow_{d^*} \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_\nu\boldsymbol{\Sigma}^{-1}\right) \tag{3.95}$$

as $(N,T)\rightarrow\infty$ 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 $\sqrt{N}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \rightarrow_d \mathcal{N}(\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_\nu\boldsymbol{\Sigma}^{-1})$ (proven in Theorem 1), then

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

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

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

which is in turn guaranteed by Polya's Theorem because (3.95) holds in probability. Hence, uniform consistency follows:

$$\begin{aligned}
&\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \leq x] - \mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \leq x]\right| \\
&= \sup_{x\in\mathbb{R}^{k\times 1}}\Big|\left(\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \leq x] - \boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_\nu\boldsymbol{\Sigma}^{-1}\right) \\
&- \left(\mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \leq x] - \boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_\nu\boldsymbol{\Sigma}^{-1})\right)\Big| \\
&\leq \sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEP,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}) \leq x] - \boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_\nu\boldsymbol{\Sigma}^{-1}\right| \\
&+ \sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}} - \boldsymbol{\beta}) \leq x] - \boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1}, \boldsymbol{\Sigma}^{-1}\boldsymbol{\Psi}_\nu\boldsymbol{\Sigma}^{-1})\right| \\
&= o_p(1),
\end{aligned} \tag{3.96}$$

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

$$\begin{aligned}
\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} &= \frac{1}{N}\sum_{i=1}^{N}s_i\left(\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{y}_i \\
&= \frac{1}{N}\sum_{i=1}^{N}s_i\boldsymbol{\beta}_i + \frac{1}{N}\sum_{i=1}^{N}s_i\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + \frac{1}{N}\sum_{i=1}^{N}s_i\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \\
&= \boldsymbol{\beta}\frac{1}{N}\underbrace{\sum_{i=1}^{N}s_i}_{N} + \frac{1}{N}\sum_{i=1}^{N}s_i\boldsymbol{\nu}_i + \frac{1}{N}\sum_{i=1}^{N}s_i\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \\
&+ \frac{1}{N}\sum_{i=1}^{N}s_i\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i,
\end{aligned} \tag{3.97}$$

hence

$$\sqrt{N}(\widehat{\boldsymbol{\beta}}^{*}_{CCEMG,\dot{\mathbf{x}}} - \boldsymbol{\beta}) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \boldsymbol{\nu}_i + \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^{*}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^{*}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i$$
$$+ \frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^{*}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^{*}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i, \tag{3.98}$$

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) \\
&= \mathbf{I} + \mathbf{II} + \mathbf{III} + o_{p^*}(1)
\end{aligned} \tag{3.99}$$

in probability, because

$$\begin{aligned}
&\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N} s_i \left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^{*}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right]\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i + T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right)\right\| \\
&\leq \sqrt{N}\sup_i \left\|\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^{*}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\| \frac{1}{N}\sum_{i=1}^{N}|s_i|\left(\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\| + \left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\|\right) \\
&= \sqrt{N}\sup_i \left\|\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^{*}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\| \frac{1}{N}\sum_{i=1}^{N}|s_i|\left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i\right\| \\
&+ \sqrt{N}\sup_i \left\|\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^{*}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1} - \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{X}_i\right)^{-1}\right\| \left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i\right\| \\
&= o_{p^*}(1)
\end{aligned} \tag{3.100}$$

as $TN^{-1} = O(1)$ in analogy to (3.69). Then

$$\begin{aligned}\|\mathbf{III}\| &= \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i \right)^{-1} \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \right) \right\| \\ &\leq \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i \right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \right\| + \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i \right)^{-1} T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\boldsymbol{\varepsilon}_i \right\| \\ &= o_{p^*}(1) \end{aligned} \tag{3.101}$$

in analogy to (3.47) by using the fact that bootstrap weights are independent from the model primitives and the results in (3.87) - (3.89). Further,

$$\begin{aligned}\mathbf{II} &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i \right)^{-1} \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right) \\ &= \underbrace{\frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \boldsymbol{\Sigma}_i^{-1} \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right)}_{o_{p^*}(1) \text{ in analogy to (3.68)}} \\ &\quad \underbrace{+ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} s_i \left[\left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{X}_i \right)^{-1} - \boldsymbol{\Sigma}_i^{-1} \right] \left(T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}^*_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}}\mathbf{F}_{\mathbf{y}}\gamma_i \right)}_{o_{p^*}(1) \text{ in analogy to (3.69)}} \\ &= o_{p^*}(1) \end{aligned} \tag{3.102}$$

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

$$\begin{aligned}\sqrt{N}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}} - \widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}) &= \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (s_i - 1)\boldsymbol{\nu}_i + o_{p^*}(1) \\ &\rightarrow_{d^*} \mathcal{N}\left(\mathbf{0}_{k\times 1}, \boldsymbol{\Omega}_{\boldsymbol{\nu}}\right) \end{aligned} \tag{3.103}$$

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

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

Hence, uniformity follows if also

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

which is in turn guaranteed by Polya's Theorem because (3.103) holds in probability. Hence, uniform consistency follows:

$$
\begin{aligned}
&\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}}-\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})\le x]-\mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEP,\dot{\mathbf{x}}}-\boldsymbol{\beta})\le x]\right|\\
&=\sup_{x\in\mathbb{R}^{k\times 1}}\Bigg|\left(\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}}-\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})\le x]-\boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1},\boldsymbol{\Omega}_{\boldsymbol{\nu}}\right)\\
&-\left(\mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}-\boldsymbol{\beta})\le x]-\boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1},\boldsymbol{\Omega}_{\boldsymbol{\nu}})\right)\Bigg|\\
&\le\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}^*[\sqrt{NT}(\widehat{\boldsymbol{\beta}}^*_{CCEMG,\dot{\mathbf{x}}}-\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}})\le x]-\boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1},\boldsymbol{\Omega}_{\boldsymbol{\nu}})\right|\\
&+\sup_{x\in\mathbb{R}^{k\times 1}}\left|\mathbb{P}[\sqrt{NT}(\widehat{\boldsymbol{\beta}}_{CCEMG,\dot{\mathbf{x}}}-\boldsymbol{\beta})\le x]-\boldsymbol{\Phi}(x;\mathbf{0}_{k\times 1},\boldsymbol{\Omega}_{\nu})\right|\\
&=o_p(1),
\end{aligned}
\tag{3.104}
$$

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

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

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 (3.48) and (3.70), respectively. Also,

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

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 $\left\| T^{-1}\mathbf{X}_i'\mathbf{M}_{\widehat{\mathbf{F}}_{\dot{\mathbf{x}}}^*}\mathbf{F}_{\mathbf{y}}\boldsymbol{\gamma}_i \right\| = o_{p^*}(1)$, which come from the proof of Theorem 6. Also,

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

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

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

Based on these arguments, we again obtain

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

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

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

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

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

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.