10-2022

Low-rank panel quantile regression: Estimation and inference

Yiren WANG
Singapore Management University, yiren.wang.2018@phdecons.smu.edu.sg

Yichong ZHANG
Singapore Management University, yczhang@smu.edu.sg

Yichong ZHANG

Low-rank Panel Quantile Regression: Estimation and Inference*

Yiren Wang[a], Liangjun Su[b] and Yichong Zhang[a]
[a]School of Economics, Singapore Management University, Singapore
[b]School of Economics and Management, Tsinghua University, China

October 21, 2022

Abstract

In this paper, we propose a class of low-rank panel quantile regression models which allow for unobserved slope heterogeneity over both individuals and time. We estimate the heterogeneous intercept and slope matrices via nuclear norm regularization followed by sample splitting, row- and column-wise quantile regressions and debiasing. We show that the estimators of the factors and factor loadings associated with the intercept and slope matrices are asymptotically normally distributed. In addition, we develop two specification tests: one for the null hypothesis that the slope coefficient is a constant over time and/or individuals under the case that true rank of slope matrix equals one, and the other for the null hypothesis that the slope coefficient exhibits an additive structure under the case that the true rank of slope matrix equals two. We illustrate the finite sample performance of estimation and inference via Monte Carlo simulations and real datasets.

Key words: Debiasing, heterogeneity, nuclear norm regularization, panel quantile regression, sample splitting, specification test.

JEL Classification: C23, C31, C32, C52

*Su gratefully acknowledges the support from the National Natural Science Foundation of China under Grant No. 72133002. Zhang acknowledges the financial support from a Lee Kong Chian fellowship. Any and all errors are our own.

1 Introduction

Panel quantile regressions are widely used to estimate the conditional quantiles, which can capture the heterogeneous effects that may vary across the distribution of the outcomes. Such effects are usually assumed to be homogeneous across individuals and over time periods. However, in empirical analyses, it is usually unknown whether the slope coefficients are homogeneous across individuals and/or time. Mistakenly forcing slopes to be homogeneous across time and individuals may lead to inconsistent estimation and misleading inferences. This prompts two questions to be answered: how can we estimate the true model at different quantiles when we allow for heterogeneous slopes across individuals and time at the same time? How to conduct specification tests for homogeneous effects over individuals or time and tests for the additive structure of the slope coefficients?

To answer the first question, we propose an estimation procedure for heterogeneous panel quantile regression models where we allow the fixed effects to be either additive or interactive, and the slope coefficients to be heterogeneous over both individuals and time. We impose a low-rank structure for both the intercept and slope coefficient matrices and estimate them via nuclear norm regularization (NNR) followed by the sample splitting, row- and column-wise quantile regressions and debiasing steps. The estimation algorithm is inspired by Chernozhukov et al. (2019), where the main difference is that we split the full sample into three subsamples rather than two because we need certain uniform results which require independence of regressors and regressand used in the debiasing step, and we do not have the closed form for the quantile regression estimates. At last, we derive the asymptotic distributions for the estimators of the factors and factor loadings associated with slope coefficient matrices.

To answer the second question, under the case when the rank of slope coefficient matrix equals one, we conduct sup-type specification tests for homogeneous effects over individuals or time following the lead of Castagnetti et al. (2015) and Lu and Su (2021). We show that our sup-test statistics follow the Gumbel distribution under the null, and the tests have non-trivial power against certain classes of local alternatives. Under the case when the rank of slope matrix equals two, our sup-type test statistic is also shown to follow the Gumbel distribution under the null that the slope coefficient exhibits an additive structure.

This paper relates to three bunches of literature. First, we contribute to the large literature on panel quantile regressions (PQRs). Since Koenker (2004) studied the PQRs with individual fixed effects, there has been an increasing number of papers on PQRs. Galvao and Montes-Rojas (2010), Kato et al. (2012), Galvao and Wang (2015), Galvao and Kato (2016), Machado and Silva (2019), and Galvao et al. (2020) study the asymptotics for PQRs with individual fixed effects. Chen et al. (2021) study quantile factor models and Chen (2022) considers PQRs with interactive fixed effects (IFEs). We complement the literature by allowing for unobserved heterogeneity in the slope coefficients of PQRs.

Second, our paper also pertains to slope heterogeneity in panel data models. Latent group structures across individuals and structural changes over time are two common types of slope heterogeneity that have received vast attention in the literature. To recover the unobserved group structures, various methods have been proposed. For example, Lin and Ng (2012), Bonhomme and Manresa (2015) and Ando and Bai (2016) use the K-means algorithm; Su et al. (2016) propose the C-lasso algorithm which is further studied and extended by Su and Ju (2018), Su et al. (2019) and Wang et al. (2019); Wang et al. (2018) propose an clustering algorithm in regression via data-driven segmentation called CARDS; Wang and Su (2021) propose a sequential binary segmentation algorithm to identify the latent group structures in nonlinear panels. Recent literature on the estimation with structural changes in panel data models includes, but is not limited to, Chen (2015), Cheng et al. (2016), Ma and Su (2018), Baltagi et al. (2021). In addition, Galvao et al. (2018) and Zhang et al. (2019) consider individual heterogeneity in PQRs while they assume homogeneity across time. To allow for both latent groups and structural breaks, Okui and Wang (2021) study a linear panel data model with individual fixed effects where each latent group has common breaks and the breaking points can be different across different groups, and they propose a grouped adaptive group fused lasso (GAGFL) approach to estimate slope coefficients. Lumsdaine et al. (2021) consider a linear panel data model with a grouped pattern of heterogeneity where the latent group membership structure and/or the values of slope coefficients can change at a breaking point, and they propose a K-means-type estimation algorithm and establish the asymptotic properties of the resulting estimators. Compared with the models studied above, our model combines both individual and time heterogeneity and only requires certain low-rank structure in the slope coefficient matrix. So the unobserved heterogeneity takes a more flexible form in our model than those in the literature such as Okui and Wang (2021) and Lumsdaine et al. (2021).

Last, our paper also connects with the burgeoning literature on nuclear norm regularization. Such a method has been widely adopted to study panel and network models. See, Alidaee et al. (2020), Athey et al. (2021), Bai and Ng (2019), Belloni et al. (2022), Chen et al. (2020), Chernozhukov et al. (2019), Feng (2019), Hong et al. (2022), Miao et al. (2022), among others. In the least squares panel framework, Moon and Weidner (2018) consider a homogeneous panel with IFEs by using NNR-based estimator as an initial estimator to construct iterative estimators that are asymptotically equivalent to the least squares estimators; Chernozhukov et al. (2019) study a heterogenous panel where both the intercept and slope coefficient matrices exhibit a low-rank structure and establish the asymptotic distribution theory based on NNR. In the presence of endogeneity, Hong et al. (2022) proposes a profile GMM method to estimate panel data models with IFEs. In the panel quantile regression setting, Feng (2019) develops error bounds for the low-rank estimates in terms of Frobenius norms under independence assumption; Belloni et al. (2022) relaxes the independence assumption to the β-mixing condition along the time dimension. Our paper extends Chernozhukov et al. (2019) from the least squares framework to the PQR framework, derives the asymptotic distribution theory and develops various specification tests under some strong mixing conditions along the time dimension that is weaker than the β-mixing condition. We also rely on the sequential symmetrization technique developed by Rakhlin et al. (2015) to obtain the convergence rates of the nuclear norm regularized estimators.

The rest of the paper is organized as follows. We first introduce the low-rank structure PQR model and the estimation algorithm in Section 2. We study the asymptotic properties of our estimators in Section 3. In Section 4, we propose two specification tests: one for the no-factor structure and one for the additive structure, and study the asymptotic properties of the test statistics. In Section 5, we show the finite sample performance of our method via Monte Carlo simulations. In Section 6, we apply our method to two datasets: one is to study how Tobin's q and cash flows affect corporate investment and whether firm's external investment to its internal financing exhibits heterogeneity structure, and the other is to study the relationship between economics growth, foreign direct investment and unemployment. Section 7 concludes. All proofs are related to the online supplement.

Notation. $\|\cdot\|_1$, $\|\cdot\|_{op}$, $\|\cdot\|_\infty$, $\|\cdot\|_{\max}$ $\|\cdot\|_2$, $\|\cdot\|_F$, $\|\cdot\|_*$ denote the matrix norm induced by 1-norms, the matrix norm induced by 2-norms, the matrix norm induced by ∞-norms, the maximum norm, the Euclidean norm, the Frobenius norm and the nuclear norm. $\odot$ is the element-wise product. $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ denote the floor and ceiling functions, respectively. $a \vee b$ and $a \wedge b$ return the max and the min of a and b, respectively. The symbol $\lesssim$ means "the left is bounded by a positive constant times the right". Let $A = \{A_{it}\}_{i\in[n],t\in[T]}$ be a matrix with its (i,t)-th entry denoted as A_{it}, where $[n]$ to denote the set $\{1,\cdots,n\}$ for any positive integer n. Let $\{A_j\}_{j=0}^{p}$ denote the collection of matrices A_j for all $j \in \{0,\cdots,p\}$. When A is symmetric, $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ denote its largest and smallest eigenvalues, respectively. The operators $\rightsquigarrow$ and $\xrightarrow{p}$ denote convergence in distribution and in probability, respectively. Besides, we use w.p.a.1 and a.s. to abbreviate "with probability approaching 1" and "almost surely", respectively.

2 Model and Estimation

In this section, we introduce the PQR model and estimation algorithm.

2.1 Model

Consider the PQR model

$$\mathscr{Q}_\tau\left(Y_{it}\middle|\left\{X_{j,it}\right\}_{j\in[p],t\in[T]},\left\{\Theta^0_{j,it}\left(\tau\right)\right\}_{j\in[p]\cup\{0\},t\in[T]}\right) = \Theta^0_{0,it}(\tau) + \sum_{j=1}^{p} X_{j,it}\Theta^0_{j,it}(\tau), \tag{2.1}$$

where $i \in [N]$, $t \in [T]$, $\tau \in (0,1)$ is the quantile index, Y_{it} is the dependent variable, $X_{j,it}$ is the j-th regressor for individual i at time t, $\{\Theta^0_{j,it}\}_{j\in[p]}$ is the corresponding slope coefficient, $\Theta^0_{0,it}$ is the intercept, and $\mathscr{Q}_\tau\left(Y_{it}\middle|\left\{X_{j,it}\right\}_{j\in[p],t\in[T]},\left\{\Theta^0_{j,it}\left(\tau\right)\right\}_{j\in[p]\cup\{0\},t\in[T]}\right)$ denotes the conditional

τ-quantile of Y_{it} given the regressors $\{X_{j,it}\}_{j\in[p],t\in[T]}$ and the parameters $\left\{\Theta^0_{j,it}(\tau)\right\}_{j\in[p]\cup 0,t\in[T]}$.[1] Alternatively, we can rewrite the above model as

$$Y = \Theta^0_0(\tau) + \sum_{j=1}^{p} X_j \odot \Theta^0_j(\tau) + \epsilon(\tau) \quad \text{and}$$
$$\mathscr{Q}_\tau\left(\epsilon_{it}(\tau) \middle| \{X_{j,it}\}_{j\in[p],t\in[T]}, \{\Theta^0_{j,it}(\tau)\}_{j\in[p]\cup\{0\},t\in[T]}\right) = 0, \tag{2.2}$$

where $\epsilon(\tau)$ is the idiosyncratic error matrix with the (i,t)-th entry being $\epsilon_{it}(\tau)$. Similarly, X_j, $\Theta_j(\tau)$, and Y are matrices with the (i,t)-th entry being $X_{j,it}$, $\Theta_{j,it}(\tau)$, and Y_{it}, respectively. In this model, we assume p, the number of regressors, is fixed and both N and T pass to infinity. In Assumption 1 below, we characterize the dependence of the data, under which (2.1) holds.

In the paper, we focus on the panel quantile regression for a fixed τ and thus suppress the dependence of $\Theta^0_j(\tau)$ and $\epsilon(\tau)$ on τ for notation simplicity. In addition, we impose low-rank structures for the intercept and slope matrices, i.e., $\text{rank}(\Theta^0_j) = K_j$ for some positive constant K_j and for each $j \in \{0, \cdots, p\}$. By the singular value decomposition (SVD), we have

$$\Theta^0_j = \sqrt{NT}\mathcal{U}^0_j\Sigma^0_j\mathcal{V}^{0\prime}_j = U^0_j V^{0\prime}_j \;\; \forall\, j = 0, \cdots, p,$$

where $\mathcal{U}^0_j \in \mathbb{R}^{N\times K_j}$, $\mathcal{V}^0_j \in \mathbb{R}^{T\times K_j}$, $\Sigma^0_j = \text{diag}(\sigma_{1,j}, \cdots, \sigma_{K_j,j})$, $U^0_j = \sqrt{N}\mathcal{U}^0_j\Sigma^0_j$ with each row being $u^{0\prime}_{i,j}$, and $V^0_j = \sqrt{T}\mathcal{V}^0_j$ with each row being $v^{0\prime}_{t,j}$.

The low-rank structure assumption includes several popular cases. For the intercept term, one commonly assumes that $\Theta^0_{0,it}$ to take the forms α^0_i, μ^0_t, or $\alpha^0_i + \mu^0_t$ in classical PQRs. Then the matrix Θ^0_0 has rank 1, 1, and 2, respectively. It is also possible to assume $\Theta^0_{0,it}$ to take an interactive form, say, $\Theta^0_{0,it} = \lambda^{0\prime}_{0,i} f^0_{0,t}$, where both $\lambda^0_{0,i}$ and $f^0_{0,t}$ are K_0-vectors. For the slope matrix Θ^0_j, $j \in [p]$, the early PQR models frequently assume that $\Theta^0_{j,it}$ is a constant across (i,t) to yield a homogenous PQR model. Obviously, such a model is very restrictive by assuming homogenous slope coefficients. It is possible to allow the slope coefficients to change over either i, or t, or both. See the following examples for different low-rank structures.

Example 1. When $\Theta^0_{j,it} = \Theta^0_{j,i}\ \forall t \in [T]$, or $\Theta^0_{j,it} = \Theta^0_{j,t}\ \forall i \in [N]$, or $\Theta^0_{j,it} = \Theta^0_j\ \forall\,(i,t) \in [N]\times[T]$, and this holds for all $j \in [p]$, we have the PQR models with only individual heterogeneity, with only time heterogeneity, and with homogeneity, respectively. We observe that $K_j = 1$ for these three cases.

[1] We will assume that both the intercept term $\Theta^0_{0,it}$ and the slope coefficients $\{\Theta^0_{j,it}\}_{j\in[p]}$ have low-rank structures, and follow the convention in the panel data literature by treating the factors to be random. Therefore, $\{\Theta^0_{j,it}\}_{j\in[p]\cup\{0\}}$ are random as well.

Example 2. When $\Theta^0_{j,it} = \lambda^0_{j,i} + f^0_{j,t}$, we notice that

$$\frac{\Theta^0_j}{\sqrt{NT}} = \begin{bmatrix} \frac{1}{\sqrt{N}} & \frac{\lambda^0_{j,1}}{\sqrt{N}} \\ \vdots & \vdots \\ \frac{1}{\sqrt{N}} & \frac{\lambda^0_{j,N}}{\sqrt{N}} \end{bmatrix} \begin{bmatrix} \frac{f^0_{j,1}}{\sqrt{T}} & \cdots & \frac{f^0_{j,T}}{\sqrt{T}} \\ \frac{1}{\sqrt{T}} & \cdots & \frac{1}{\sqrt{T}} \end{bmatrix} := A_j B'_j.$$

Let $\Sigma_{A,j} := A'_j A_j$ and $\Sigma_{B,j} := B'_j B_j$. Let $\Sigma^{\frac{1}{2}}_{A,j}$ (resp. $\Sigma^{\frac{1}{2}}_{B,j}$) be the symmetric square root of $\Sigma_{A,j}$ (resp. $\Sigma_{B,j}$). By eigendecomposition, we have $\Sigma^{\frac{1}{2}}_{A,j} = P_{j,1} S_{j,1} P'_{j,1}$ and $\Sigma^{\frac{1}{2}}_{B,j} = P_{j,2} S_{j,2} P'_{j,2}$. Besides, we apply singular value decomposition to matrix $S_{j,1} P'_{j,1} P_{j,2} S_{j,2}$: $S_{j,1} P'_{j,1} P_{j,2} S_{j,2} = Q_{j,1} R_j Q'_{j,2}$. Then it follows that

$$\begin{aligned} \frac{\Theta^0_j}{\sqrt{NT}} &= A_j B'_j = A_j \Sigma^{-\frac{1}{2}}_{A,j} P_{j,1} S_{j,1} P'_{j,1} P_{j,2} S_{j,2} P'_{j,2} \Sigma^{-\frac{1}{2}}_{B,j} B'_j \\ &= A_j \Sigma^{-\frac{1}{2}}_{A,j} P_{j,1} Q_{j,1} R_j Q'_{j,2} P'_{j,2} \Sigma^{-\frac{1}{2}}_{B,j} B'_j := \mathcal{U}^0_j \Sigma^0_j \mathcal{V}^{0\prime}_j, \end{aligned}$$

where $\mathcal{U}^0_j = A_j \Sigma^{-\frac{1}{2}}_{A,j} P_{j,1} Q_{j,1}$, $\Sigma^0_j = R_j$ and $\mathcal{V}^0_j = B_j \Sigma^{-\frac{1}{2}}_{B,j} P_{j,2} Q_{j,2}$. Given $P_{j,1}$, $P_{j,2}$, $Q_{j,1}$ and $Q_{j,2}$ are orthonormal matrices, it's easy to that $\mathcal{U}^0_j$ and $\mathcal{V}^0_j$ are also orthonormal so that $\mathcal{U}^{0\prime}_j \mathcal{U}^0_j = \mathcal{V}^{0\prime}_j \mathcal{V}^0_j = I_2$. When $j = 0$, $\left\{\lambda^0_{0,i}\right\}_{i=1}^N$ and $\left\{f^0_{0,t}\right\}_{t=1}^T$ are usually referred to as the individual and time fixed effects, respectively, so that the intercept term exhibits an additive fixed effects structure.

Example 3. Let $\Theta^0_{j,it} = \sum_{k \in [K_{j,t}]} \alpha_{j,kt} \mathbf{1}\{i \in G_{j,kt}\}$, where $\{G_{j,kt}\}$ forms a partition of $[N]$ for each specific time t and $K_{j,t}$ is the number of groups at time t. Moreover, let

$$\alpha_{j,kt} = \begin{cases} \alpha^{(1)}_{j,k}, & \text{for} \quad t = 1, \ldots, T_b, \\ \alpha^{(2)}_{j,k}, & \text{for} \quad t = T_b + 1, \ldots, T, \end{cases}$$

$$G_{j,kt} = \begin{cases} G^{(1)}_{j,k}, & \text{for} \quad t = 1, \ldots, T_b, k = 1, \ldots, K^{(1)}_j, \\ G^{(2)}_{j,k}, & \text{for} \quad t = T_b + 1, ..., T, k = 1, ..., K^{(2)}_j, \end{cases}$$

where $K^{(1)}_j$ and $K^{(2)}_j$ are the number of groups before and after the break point T_b. If $K^{(1)}_j = K^{(2)}_j$, it is clear that $rank(\Theta^0_j) = 1$. If the group structure does not change after the break but $\alpha^{(1)}_{j,k} = c\alpha^{(2)}_{j,k}$ for some constant c, we also have $rank(\Theta^0_j) = 1$. Except for these two cases, we can show that

$$\Theta^0_j = \begin{bmatrix} \sum\limits_{k \in [K^{(1)}]} \alpha^{(1)}_{j,k} \mathbf{1}\left\{1 \in G^{(1)}_{j,k}\right\}, & \sum\limits_{k \in [K^{(2)}]} \alpha^{(2)}_{j,k} \mathbf{1}\left\{1 \in G^{(2)}_{j,k}\right\} \\ \vdots & \vdots \\ \sum\limits_{k \in [K^{(1)}]} \alpha^{(1)}_{j,k} \mathbf{1}\left\{i \in G^{(1)}_{j,k}\right\}, & \sum\limits_{k \in [K^{(2)}]} \alpha^{(2)}_{j,k} \mathbf{1}\left\{i \in G^{(2)}_{j,k}\right\} \\ \vdots & \vdots \\ \sum\limits_{k \in [K^{(1)}]} \alpha^{(1)}_{j,k} \mathbf{1}\left\{N \in G^{(1)}_{j,k}\right\}, & \sum\limits_{k \in [K^{(2)}]} \alpha^{(2)}_{j,k} \mathbf{1}\left\{N \in G^{(2)}_{j,k}\right\} \end{bmatrix} \begin{bmatrix} \iota_{T_b} & \mathbf{0}_{T_b} \\ \mathbf{0}_{T-T_b} & \iota_{T-T_b} \end{bmatrix}'$$

where ι_{T_b} is a $T_b \times 1$ vector of ones and $\mathbf{0}_{T_b}$ is a $T_b \times 1$ vector of zeros. In this case, we notice that $rank(\Theta_j^0) = 2$.

Example 4. When $\Theta_{j,it}^0 = \lambda_{j,i}^{0\prime} f_{j,t}^0$ with $\lambda_{j,i}^0$ and $f_{j,t}^0$ being two K_j-vectors, we have the IFEs structure. This is the most general example without further restrictions.

Like Chernozhukov et al. (2019), we assume that for each $j \in [p]$, $X_{j,it}$ exhibits a factor structure: $X_{j,it} = \mu_{j,it} + e_{j,it} = l_{j,i}^{0\prime} w_{j,t}^0 + e_{j,it}$, where $w_{j,t}^0$ and $l_{j,i}^0$ are the factors and factor loadings of dimension r_j.

2.2 Estimation Algorithm

In this subsection we provide the estimation algorithm by assuming that K_j are all known for all j. In the next subsection, we will introduce a rank estimation method to estimate K_j consistently.

Define the check function $\rho_\tau(u) = u\,(\tau - \mathbf{1}\{u \leq 0\})$. The estimation procedure goes as follows:

Step 1: **Sample Splitting and Nuclear Norm Regularization.** Along the cross-section span, randomly split the sample into three subsets denoted as I_1, I_2 and I_3, where I_ℓ has N_ℓ individuals such that $N_1 \approx N_2 \approx N_3 \approx N/3$. Using the data with $(i,t) \in I_1 \times [T]$, we run the nuclear norm regularized quantile regression (QR) and obtain $\{\tilde{\Theta}_j^{(1)}\}_{j \in \{0,\cdots,p\}}$, i.e.,

$$\{\tilde{\Theta}_j^{(1)}\}_{j=0}^p = \underset{\{\Theta_j\}_{j=0}^p}{\arg\min} \frac{1}{N_1 T} \sum_{i \in I_1} \sum_{t=1}^T \rho_\tau(Y_{it} - \sum_{j=1}^p X_{j,it}\Theta_{j,it} - \Theta_{0,it}) + \sum_{j=0}^p \nu_j \left\|\Theta_j\right\|_*, \quad (2.3)$$

where ν_j is a tuning parameter. For each j, conduct the SVD: $\frac{1}{\sqrt{N_1 T}}\tilde{\Theta}_j^{(1)} = \hat{\tilde{\mathcal{U}}}_j^{(1)} \hat{\tilde{\Sigma}}_j^{(1)} \hat{\tilde{\mathcal{V}}}_j^{(1)\prime}$, where $\hat{\tilde{\Sigma}}_j^{(1)}$ is the diagonal matrix with the diagonal elements being the descending singular values of $\tilde{\Theta}_j^{(1)}$. Let $\tilde{\mathcal{V}}_j^{(1)}$ consists the first K_j columns of $\hat{\tilde{\mathcal{V}}}_j^{(1)}$. Let $\tilde{V}_j^{(1)} = \sqrt{T}\tilde{\mathcal{V}}_j^{(1)}$ and $\tilde{v}_{t,j}^{(1)\prime}$ be the t-th row of $\tilde{V}_j^{(1)}$ $\forall t \in [T]$.

Step 2: **Row- and Column-Wise Quantile Regression.** Using the data with $(i,t) \in I_2 \times [T]$, we first run the row-wise QR of Y_{it} on $\left(\tilde{v}_{t,0}^{(1)}, \left\{\tilde{v}_{t,j}^{(1)} X_{j,it}\right\}_{j \in [p]}\right)$ to obtain $\{\dot{u}_{i,j}^{(1)}\}_{j=0}^p$ for $i \in I_2$, and then run the column-wise QR of Y_{it} on $(\dot{u}_{i,0}^{(1)}, \{\dot{u}_{i,j}^{(1)} X_{j,it}\}_{j \in [p]})$ to obtain $\{\dot{v}_{t,j}^{(1)}\}_{j=0}^p$ for $t \in [T]$. That is,

$$\{\dot{u}_{i,j}^{(1)}\}_{j=0}^p = \underset{\{u_{i,j}\}_{j \in [p] \cup \{0\}}}{\arg\min} \frac{1}{T} \sum_{t \in [T]} \rho_\tau \left(Y_{it} - u_{i,0}'\tilde{v}_{t,0}^{(1)} - \sum_{j=1}^p u_{i,j}'\tilde{v}_{t,j}^{(1)} X_{j,it} \right), \forall i \in I_2, \quad (2.4)$$

$$\{\dot{v}_{t,j}^{(1)}\}_{j=0}^p = \underset{\{v_{t,j}\}_{j \in [p] \cup \{0\}}}{\arg\min} \frac{1}{N_2} \sum_{i \in I_2} \rho_\tau \left(Y_{it} - v_{t,0}'\dot{u}_{i,0}^{(1)} - \sum_{j=1}^p v_{t,j}'\dot{u}_{i,j}^{(1)} X_{j,it} \right), \forall t \in [T]. \quad (2.5)$$

Similarly, we run the row-wise QR of Y_{it} on $(\dot{v}_{t,0}^{(1)}, \{\dot{v}_{t,j}^{(1)} X_{j,it}\}_{j\in[p]})$ to obtain $\{\dot{u}_{i,j}^{(1)}\}_{j=0}^{p}$ for $i \in I_3$, i.e.,

$$\{\dot{u}_{i,j}^{(1)}\}_{j=0}^{p} = \underset{\{u_{i,j}\}_{j\in[p]\cup\{0\}}}{\arg\min} \frac{1}{T} \sum_{t\in[T]} \rho_\tau \left(Y_{it} - u_{i,0}'\dot{v}_{t,0}^{(1)} - \sum_{j=1}^{p} u_{i,j}'\dot{v}_{t,j}^{(1)} X_{j,it} \right), \forall i \in I_3.$$

Step 3: **Debiasing.**

Step 3.1: For each $j \in [p]$, we conduct the principle component analysis (PCA) for $X_{j,it}$ with $(i,t) \in [N] \times [T]$ to obtain the factor and factor loading estimates as

$$\left\{\hat{l}_{j,i}, \hat{w}_{j,t}\right\}_{i\in[N],t\in[T]} = \underset{\{l_{j,i},w_{j,t}\}_{i\in[N],t\in[T]}}{\arg\min} \frac{1}{NT} \sum_{i\in[N]} \sum_{t\in[T]} \left(X_{j,it} - l_{j,i}' w_{j,t}\right)^2, \quad (2.6)$$

subject to the normalizations: $\frac{1}{N}\sum_{i=1}^{N} l_{i,j} l_{i,j}' = I_{r_j}$ and $\frac{1}{T}\sum_{t=1}^{T} w_{j,t} w_{j,t}'$ is a diagonal matrix with descending diagonal elements. Then we define $\hat{\mu}_{j,it} = \hat{l}_{j,i}' \hat{w}_{j,t}$ and $\hat{e}_{j,it} = X_{j,it} - \hat{\mu}_{j,it}$.

Step 3.2: For $(i,t) \in I_3 \times [T]$, let $\tilde{Y}_{it} = Y_{it} - \sum\limits_{j=1}^{p} \hat{\mu}_{j,it} \dot{u}_{i,j}^{(1)\prime} \dot{v}_{t,j}^{(1)}$. We run the row-wise QR of $\tilde{Y}_{it}$ on $(\dot{v}_{t,0}^{(1)}, \{\dot{v}_{t,j}^{(1)} \hat{e}_{j,it}\}_{j\in[p]})$ to obtain the final estimates $\hat{u}_{i,j}^{(3,1)}$, i.e.,

$$\{\hat{u}_{i,j}^{(3,1)}\}_{j=0}^{p} = \underset{\{u_{i,j}\}_{j=0}^{p}}{\arg\min} \frac{1}{T} \sum_{t\in[T]} \rho_\tau \left(\tilde{Y}_{it} - u_{i,0}'\dot{v}_{t,0}^{(1)} - \sum_{j=1}^{p} u_{i,j}'\dot{v}_{t,j}^{(1)} \hat{e}_{j,it} \right), \forall i \in I_3. \quad (2.7)$$

Updating $\hat{Y}_{it} = Y_{it} - \sum_{j=1}^{p} \hat{\mu}_{j,it} \hat{u}_{i,j}^{(3,1)\prime} \dot{v}_{t,j}^{(1)}$, we run the column-wise QR of $\hat{Y}_{it}$ on $(\hat{u}_{i,0}^{(3,1)}, \{\hat{u}_{i,j}^{(3,1)} \hat{e}_{j,it}\}_{j\in[p]})$ to obtain $\hat{v}_{t,j}^{(3,1)}$, i.e.,

$$\{\hat{v}_{t,j}^{(3,1)}\}_{j=0}^{p} = \underset{\{v_{t,j}\}_{j=0}^{p}}{\arg\min} \frac{1}{N_3} \sum_{i\in I_3} \rho_\tau \left(\hat{Y}_{it} - v_{t,0}'\hat{u}_{i,0}^{(3,1)} - \sum_{j=1}^{p} v_{t,j}'\hat{u}_{i,j}^{(3,1)} \hat{e}_{j,it} \right), \forall t \in [T]. \quad (2.8)$$

In order to obtain the final estimators for the full sample, we propose to switch the role of each subsample for the low-rank estimation, row- and column-wise QR and debiasing, then repeat Steps 1-3 to obtain $\left\{\hat{u}_{i,j}^{(a,b)}\right\}_{j=0}^{p}$ and $\left\{\hat{v}_{t,j}^{(a,b)}\right\}_{j=0}^{p}$ for $a \in [3]$ and $b \in [3] \setminus \{a\}$. Here (a,b) denotes the final estimates for subsample I_a obtained from the first step NNR estimates with subsample I_b. Table 1 shows the final estimators we obtain by using different combination of subsamples.

Several remarks are in order. First, we randomly split the full sample into three subsamples, each playing a significant role in the algorithm. We use the first subsample for the low-rank estimation to obtain the preliminary NNR estimators of the submatrices of the intercept and slope matrices. But these estimators are only consistent in terms of Frobenius norm, and one

Table 1: Estimators using different subsamples at different steps in algorithm.

Step 1 (b)	Step 2	Step 3 (a)	estimators (a, b)
I_1	I_2	I_3	$\hat{u}_{i,j}^{(3,1)}$, $\hat{v}_{t,j}^{(3,1)}$
I_2	I_1	I_3	$\hat{u}_{i,j}^{(3,2)}$, $\hat{v}_{t,j}^{(3,2)}$
I_1	I_3	I_2	$\hat{u}_{i,j}^{(2,1)}$, $\hat{v}_{t,j}^{(2,1)}$
I_3	I_1	I_2	$\hat{u}_{i,j}^{(2,3)}$, $\hat{v}_{t,j}^{(2,3)}$
I_2	I_3	I_1	$\hat{u}_{i,j}^{(1,2)}$, $\hat{v}_{t,j}^{(1,2)}$
I_3	I_2	I_1	$\hat{u}_{i,j}^{(1,3)}$, $\hat{v}_{t,j}^{(1,3)}$

cannot derive the pointwise or uniform convergence rates for them. With the low-rank estimates, we use the second subsample to do the row- and column-wise QRs and can now establish the uniform convergence rates for each row of factor and factor loading estimators. Then we use the remaining subsample to debias the second-stage estimator and to obtain the final estimators that have the desirable asymptotic properties.

Second, to reduce the randomness of sample splitting, one can run the estimation algorithm several times with different splittings in practice. Once one obtains factor and factor loading estimates, one can construct estimators for Θ_j^0 under different splittings and then choose the one specific splitting which yields the minimum quantile objective function.

Third, the bias in the second-stage estimator is inherent from the first-stage NNR estimator. We follow the lead of Chernozhukov et al. (2019) to assume that $X_{j,it}$ has a factor structure with an additive idiosyncratic term, and remove the bias by a QR with the demeaned $X_{j,it}$ as regressors. In the least squares panel regression framework, the objective function is smooth and one has closed-form solutions in the last stage so that Chernozhukov et al. (2019) only need to split the sample into two subsamples. In contrast, in the PQR framework, the objective function is non-smooth, we do not have closed-form solutions in any stage. In order to remove the bias from the early stage estimation and to derive the distributional results, we need to split the sample into three subsamples.

To save space, we relegate the detailed algorithm for the nuclear norm regularization to the online supplement.

2.3 Rank Estimation

In this subsection we discuss how to estimate the ranks K_j consistently. To estimate the ranks, we consider the full sample NNR QR estimation:

$$\{\tilde{\Theta}_j\}_{j=0}^p = \underset{\{\Theta_j\}_{j=0}^p}{\arg\min} \ \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \rho_\tau(Y_{it} - \sum_{j=1}^p X_{j,it}\Theta_{j,it} - \Theta_{0,it}) + \sum_{j=0}^p \nu_j \left\|\Theta_j\right\|_* . \tag{2.9}$$

For $j \in \{0, \cdots, p\}$, we estimate K_j by the popular singular value thresholding (SVT) as follows

$$\hat{K}_j = \sum_m \mathbf{1}\left\{\lambda_m\left(\tilde{\Theta}_j\right) \geq 0.6\left(NT\nu_j\left\|\tilde{\Theta}_j\right\|_{op}\right)^{1/2}\right\}.$$

It is standard to show that $\mathbb{P}(\hat{K}_j = K_j) \to 1$ as $(N, T) \to \infty$ under some regularity conditions given in the next section; see also Proposition D.1 in Chernozhukov et al. (2019) and Theorem 2 in Hong et al. (2022). Since the ranks can be estimated consistently, we assume that they are known in the asymptotic theory below.

3 Asymptotic Theory

In this section, we study the asymptotic properties of the estimators introduced in the last section.

3.1 First Stage Estimator

Recall that $X_{j,it} = \mu_{j,it} + e_{j,it} = l_{j,i}^{0\prime} w_{j,t}^0 + e_{j,it}$ for each $j \in [p]$. Let $X_{it} = (X_{1,it}, ..., X_{p,it})'$ and $e_{it} = (e_{1,it}, ..., e_{p,it})'$. Define $\epsilon_i = (\epsilon_{i1}, \cdots, \epsilon_{it})'$, $e_{j,i} = (e_{j,i1}, \cdots, e_{j,iT})'$, W_j^0 as the $T \times r_j$ matrix with each row being $w_{j,t}^{0\prime}$, and V_j^0 as the $T \times K_j$ matrix with each row being $v_{t,j}^{0\prime}$. Further define $a_{it} = \tau - \mathbf{1}\{\epsilon_{it} \leq 0\}$ with $a_i = (a_{i1}, \cdots, a_{iT})'$ and $a = (a_1, \cdots, a_N)'$. Throughout the paper, we treat the factors $\{v_{t,j}^0\}_{t\in[T], j\in[p]\cup\{0\}}$ and $\{w_{j,t}^0\}_{t\in[T], j\in[p]}$ as random and their loadings $\{u_{i,j}^0\}_{i\in[N], j\in[p]\cup\{0\}}$ and $\{l_{j,i}^0\}_{i\in[N], j\in[p]}$ as deterministic.

Table 2 defines several σ-fields. We use $\mathscr{D}$ to denote the minimal σ-field generated by $\left\{V_j^0\right\}_{j\in[p]\cup\{0\}} \bigcup \left\{W_j^0\right\}_{j\in[p]}$; the superscripts I_1 and $I_1 \cup I_2$ are associated with the first subsample and the first two subsamples, respectively. For example, $\mathscr{D}_{e_i}^{I_1}$ denotes the minimal σ-field generated by $\mathscr{D}$, $\{e_{it}\}_{t\in[T]}$ and $\{\epsilon_{it}, e_{it}\}_{i\in I_1, t\in[T]}$.

Let M denote a generic bounded constant that may vary across places. Let $\mathscr{G}_{i,t-1}$ denote the minimal σ-field generated by $\mathscr{D} \cup \{e_{ls}\}_{l\leq i-1, s\in[T]} \cup \{e_{is}\}_{s\leq t} \cup \{\epsilon_{ls}\}_{l\leq i-1, s\in[T]} \cup \{\epsilon_{is}\}_{s\leq t-1}$. Let $\mathsf{F}_{it}(\cdot)$ and $\mathsf{f}_{it}(\cdot)$ be the conditional cumulative distribution function (CDF) and probability density function (PDF) of ϵ_{it} given $\mathscr{G}_{i,t-1}$, respectively. Similarly, let $\mathfrak{F}_{it}(\cdot)$ and $\mathfrak{f}_{it}(\cdot)$ denote the conditional CDF and PDF of ϵ_{it} given $\mathscr{D}_{e_i}$; $F_{it}(\cdot)$ and $f_{it}(\cdot)$ denote the conditional CDF and PDF of ϵ_{it} given $\mathscr{D}_e$. Let $\mathsf{f}_{it}'(\cdot)$, $\mathfrak{f}_{it}'(\cdot)$, and $f_{it}'(\cdot)$ denotes the first derivative of the density $\mathsf{f}_{it}(\cdot)$, $\mathfrak{f}_{it}(\cdot)$, and $f_{it}(\cdot)$, respectively.

We make the following assumptions.

Assumption 1 *(i)* $\{\epsilon_{it}, e_{it}\}_{t\in[T]}$ *are conditionally independent across* i *given* $\mathscr{D}$.

(ii) $\mathbb{E}\left(a_{it}\middle|\mathscr{D}_e\right) = 0.$

Table 2: Definition of various σ-fields

Notation	σ-fields generated by
$\mathscr{D}$	$\left\{V_j^0\right\}_{j\in[p]\cup\{0\}} \bigcup \left\{W_j^0\right\}_{j\in[p]}$
$\mathscr{D}_{e_{it}}$	$\mathscr{D} \bigcup e_{it}$
$\mathscr{D}_{e_i}$	$\mathscr{D} \bigcup \{e_{it}\}_{t\in[T]}$
$\mathscr{D}_e$	$\mathscr{D} \bigcup \{e_{it}\}_{i\in[N],t\in[T]}$
$\mathscr{D}^{I_1\cup I_2}$	$\mathscr{D} \bigcup \{\epsilon_{it}, e_{it}\}_{i\in I_1\cup I_2,t\in[T]}$
$\mathscr{D}^{I_1}_{\{e_{is}\}_{s<t}}$	$\mathscr{D} \bigcup \{e_{is}\}_{s<t} \bigcup \{\epsilon_{i^*t^*}, e_{i^*t^*}\}_{i^*\in I_1,t^*\in[T]}$
$\mathscr{D}^{I_1}_{e_i}$	$\mathscr{D} \bigcup \{e_{it}\}_{t\in[T]} \bigcup \{\epsilon_{i^*t^*}, e_{i^*t^*}\}_{i^*\in I_1,t^*\in[T]}$
$\mathscr{D}^{I_1\cup I_2}_{e_i}$	$\mathscr{D} \bigcup \{e_{it}\}_{t\in[T]} \bigcup \{\epsilon_{i^*t^*}, e_{i^*t^*}\}_{i^*\in I_1\cup I_2,t^*\in[T]}$
$\mathscr{D}^{I_1\cup I_2}_{e}$	$\mathscr{D} \bigcup \{e_{it}\}_{i\in[N,]t\in[T]} \bigcup \{\epsilon_{it}, e_{it}\}_{i\in I_1\cup I_2,t\in[T]}$

(iii) For each i, $\{\epsilon_{it}, t\geq 1\}$ is strong mixing conditional on $\mathscr{D}_{e_i}$, and $\{(\epsilon_{it}, e_{it}), t\geq 1\}$ is strong mixing conditional on $\mathscr{D}$. Both mixing coefficients are upper bounded by $\alpha_i(\cdot)$ such that $\max_{i\in[N]}\alpha_i(z)\leq M\alpha^z$ for some constant $\alpha\in(0,1)$.

(iv) $\max_{i\in[N]}\frac{1}{T}\sum_{t\in[T]}\|X_{it}\|_2^3\leq M$ *a.s.*, $\max_{t\in[T]}\frac{1}{N_2}\sum_{i\in I_2}\|X_{it}\|_2^4\leq M$ *a.s.*, $\max_{(i,t)\in[N]\times[T]}\mathbb{E}\left[\|X_{it}\|_2^3\middle|\mathscr{D}\right]\leq M$ *a.s.*, $\max_{i\in[N]}\sqrt{\frac{1}{T}\sum_{t\in[T]}\left[\mathbb{E}\left(\epsilon_{it}^2\middle|\mathscr{D}_{e_i}\right)\right]^2}\leq M$ *a.s., and* $\max_{(i,t)\in[N]\times[T]}\mathbb{E}\left(\|X_{it}\|_2^2\middle|\mathscr{D}_{\{e_{is}\}_{s<t}}\right)\leq M$ *a.s.*

(v) For $j\in[p]$, there exists a positive sequence ξ_N such that $\max_{(i,t)\in[N]\times[T]}|X_{j,it}|\leq\xi_N$ a.s.

(vi) $\min_{(i,t)\in[N]\times[T]}\mathsf{f}_{it}(0)\geq\underline{\mathsf{f}}>0$ *and* $\max_{(i,t)\in[N]\times[T]}\sup_\epsilon|\mathsf{f}'_{it}(\epsilon)|\leq\bar{\mathsf{f}}'$.

(vii) $\min_{(i,t)\in[N]\times[T]}\mathfrak{f}_{it}(0)\geq\underline{\mathfrak{f}}>0$ *and* $\max_{(i,t)\in[N]\times[T]}\sup_\epsilon|\mathfrak{f}'_{it}(\epsilon)|\leq\bar{\mathfrak{f}}'$.

(viii) $\min_{(i,t)\in[N]\times[T]}f_{it}(0)\geq\underline{f}>0$ *and* $\max_{(i,t)\in[N]\times[T]}\sup_\epsilon|f'_{it}(\epsilon)|\leq\bar{f}'$.

(ix) $\frac{\xi_N^4\log(N\vee T)\sqrt{N\vee T}}{N\wedge T}=o(1)$ *and* $\frac{\left(\frac{N}{T}\vee 1\right)^{1/2}}{(N\wedge T)^{\frac{1}{4+2\vartheta}}}(\log(N\vee T))^{\frac{3+\vartheta}{4+2\vartheta}}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}=o(1)$ *for any $\vartheta>0$.*

Assumptions 1(i) imposes conditional independence of the error terms and covariates $X_{j,it}$ given the fixed effects. Assumptions 1(ii) imposes the moment condition for QR. Assumptions 1(iii) imposes the weak dependence assumption along the time dimension via the use of the notion of conditional strong mixing. See Prakasa Rao (2009) for the definition of conditional strong mixing and Su and Chen (2013) for an application in the panel setup. Assumptions 1(iv)-(v) essentially imposes some conditions on the moments and tail behavior of the both covariates and errors. Note that we allow $X_{j,it}$ to have an infinite support. Assumptions 1(vi)-(viii), which are used in the proofs of Theorems 3.1, 3.2 and 3.3, respectively, specify conditions on the conditional density of ϵ_{it} given different σ-fields. Assumption 1(ix) imposes some restrictions on N, T and ξ_N in order to obtain the error bound of NNR estimators and to achieve the unbiasedness. It allows not only the case that N and T diverge to infinity at the the same rate, but also the case that N diverges to infinity not too faster than T, and vice versa.

Assumption 2 *Θ_0^0 is the fixed effect matrix with fixed rank K_0 and $\|\Theta_0^0\|_{\max} \leq M$. For each $j \in [p]$, Θ_j^0 is the slope matrix of regressor j with rank being K_j such that $\max_{j\in[p]} \|\Theta_j\|_{\max} \leq M$ and $\max_{j\in[p]} K_j \leq \bar{K}$ for some fixed finite $\bar{K}$.*

Assumption 2 is the low-rank assumption for the intercept and slope matrices, which is the key assumption for the NNR. The uniform boundedness of elements of these matrices facilitates the asymptotic analysis, but can be relaxed at the cost of more lengthy argument. See Ma et al. (2020) for a similar condition.

Assumption 3 *There exist some constants C_σ and c_σ such that*

$$\infty > C_\sigma \geq \limsup_{N,T} \max_{j\in[p]\cup\{0\}} \sigma_{1,j} \geq \liminf_{N,T} \min_{j\in[p]\cup\{0\}} \sigma_{K_j,j} \geq c_\sigma > 0.$$

Assumption 3 imposes some conditions on the singular values of the coefficient matrices. It implies that we only allow pervasive factors when these matrices are written as a factor structure. Such an assumption is common in the literature; see, e.g., Assumption 3 in Ma et al. (2020).

To introduce the next assumption, we need some notation. Let $\Theta_j^0 = R_j \Sigma_j S_j'$ be the SVD for Θ_j^0. Further decompose $R_j = (R_{j,r}, R_{j,0})$ with $R_{j,r}$ being the singular vectors corresponding to the nonzero singular values, $R_{j,0}$ being the singular vectors corresponding to the zero singular values. Decompose $S_j = (S_{j,r}, S_{j,0})$ with $S_{j,r}$ and $S_{j,0}$ defined analogously. For any matrix $W \in \mathbb{R}^{N\times T}$, we define

$$\mathcal{P}_j^{\perp}(W) = R_{j,0}R_{j,0}' W S_{j,0} S_{j,0}', \quad \mathcal{P}_j(W) = W - \mathcal{P}_j^{\perp}(W),$$

where $\mathcal{P}_j(W)$ and $\mathcal{P}_j^{\perp}(W)$ are the linear projection of matrix W onto the low-rank space and its orthogonal space, respectively. Let $\Delta_{\Theta_j} = \Theta_j - \Theta_j^0$ for any Θ_j. With some positive constants C_1 and C_2, we define the following cone-like restricted set:

$$\mathcal{R}(C_1, C_2) := \left\{ \left(\{\Delta_{\Theta_j}\}_{j=0}^p\right) : \sum_{j=0}^p \left\|\mathcal{P}_j^{\perp}(\Delta_{\Theta_j})\right\|_* \leq C_1 \sum_{j=0}^p \left\|\mathcal{P}_j(\Delta_{\Theta_j})\right\|_*, \sum_{j=0}^p \left\|\Delta_{\Theta_j}\right\|_F^2 \geq C_2\sqrt{NT} \right\}.$$

Assumption 4 *Let $C_2 > 0$ be a sufficiently large but fixed constant. There are constants C_3, C_4, such that, uniformly over $(\{\Delta_{\Theta_j}\}_{j=0}^p) \in \mathcal{R}(3, C_2)$, we have*

$$\left\| \Delta_{\Theta_0} + \sum_{j=1}^p \Delta_{\Theta_j} \odot X_j \right\|_F^2 \geq C_3 \sum_{j=0}^p \left\| \Delta_{\Theta_j} \right\|_F^2 - C_4(N+T) \; w.p.a.1.$$

The same condition holds when Θ_j^0 is replaced by $\{\Theta_{j,it}^0\}_{i \in I_a, t \in [T]}$ for $a = 1, 2, 3$.

Assumption 4 parallels the restricted strong convexity (RSC) condition in Assumption 3.1 of Chernozhukov et al. (2019) who also provide some sufficient primitive conditions.

For any $j \in \{0, \cdots, p\}$, define $\tilde{\Delta}_{\Theta_j} = \tilde{\Theta}_j - \Theta_j^0$ and $\tilde{\Delta}_{\Theta_j}^{(1)} = \tilde{\Theta}_j^{(1)} - \Theta_j^{0,(1)}$, where $\Theta_j^{0,(1)} = \left\{\Theta_{j,it}^0\right\}_{i \in I_1, t \in [T]}$. The following theorem establishes the convergence rates of the NNR estimators of the coefficient matrices.

Theorem 3.1 *If Assumptions 1-4 hold, for $\forall j \in \{0, \cdots, p\}$, we have*

(i) $\frac{1}{\sqrt{NT}} \left\| \tilde{\Delta}_{\Theta_j} \right\|_F = O_p\left(\sqrt{\frac{\log(N \vee T)}{N \wedge T}} \xi_N^2\right)$, $\frac{1}{\sqrt{NT}} \left\| \tilde{\Delta}_{\Theta_j}^{(1)} \right\|_F = O_p\left(\sqrt{\frac{\log(N \vee T)}{N \wedge T}} \xi_N^2\right)$,

(ii) $\max_{k \in [K_j]} |\tilde{\sigma}_{k,j} - \sigma_{k,j}| = O_p\left(\sqrt{\frac{\log(N \vee T)}{N \wedge T}} \xi_N^2\right)$, $\max_{k \in [K_j]} \left|\tilde{\sigma}_{k,j}^{(1)} - \sigma_{k,j}\right| = O_p\left(\sqrt{\frac{\log(N \vee T)}{N \wedge T}} \xi_N^2\right)$,

(iii) $\frac{1}{\sqrt{T}} \left\| V_j^0 - \tilde{V}_j O_j \right\|_F = O_p\left(\sqrt{\frac{\log(N \vee T)}{N \wedge T}} \xi_N^2\right)$, $\frac{1}{\sqrt{T}} \left\| V_j^0 - \tilde{V}_j^{(1)} O_j^{(1)} \right\|_F = O_p\left(\sqrt{\frac{\log N \vee T}{N \wedge T}} \xi_N^2\right)$,

where O_j and $O_j^{(1)}$ are some orthogonal rotation matrices defined in the proof.

Remark 1. Theorem 3.1(i) reports the "rough" convergence rates of the NNR estimators of the coefficient matrices in terms of Frobenius norm for both the full-sample and sub-sample estimators. Unlike the traditional $(N \wedge T)^{-1/2}$-rate in the least squares framework, NNR estimators' convergence rates in the PQR framework usually have an additional $\sqrt{\log(N \vee T)}$ term due to the use of some exponential inequalities. The extra term ξ_N^2 in our rate is due to the upper bound of $|X_{j,it}|$, and it disappears in case $X_{j,it}$'s are uniformly bounded. Theorem 3.1(ii)-(iii) report the convergence rates for the estimators of the factors and factor loadings of Θ_j^0, which are inherited from those in Theorem 3.1(i). To derive these results, we establish the symmetrization inequality and contraction principle for the sequential symmetrization developed by Rakhlin et al. (2015). See Lemmas B.8 and B.9 in the online supplement for more detail.

3.2 Second Stage Estimator

To study the asymptotic properties of the second-stage estimators, we add some notation. Define

$$\Phi_i = \frac{1}{T} \sum_{t=1}^T \Phi_{it}^0 \Phi_{it}^{0\prime} \quad \text{and} \quad \Psi_t = \frac{1}{N_2} \sum_{i \in I_2} \Psi_{it}^0 \Psi_{it}^{0\prime},$$

where $\Phi_{it}^0 = (v_{t,0}^{0\prime}, v_{t,1}^{0\prime} X_{1,it}, \cdots, v_{t,p}^{0\prime} X_{p,it})'$ and $\Psi_{it}^0 = (u_{i,0}^{0\prime}, u_{i,1}^{0\prime} X_{1,it}, \cdots, u_{i,p}^{0\prime} X_{p,it})'$. Let $K = \sum_{j=0}^p K_j$. Note that Φ_i and Ψ_t are $K \times K$ matrices. We add the following two assumptions.

Assumption 5 *There exist constants C_ϕ and c_ϕ such that a.s.*

$$\infty > C_\psi \geq \limsup_T \max_{t\in[T]} \lambda_{\max}(\Psi_t) \geq \liminf_T \min_{t\in[T]} \lambda_{\min}(\Psi_t) \geq c_\psi > 0,$$
$$\infty > C_\phi \geq \limsup_N \max_{i\in I_2} \lambda_{\max}(\Phi_i) \geq \liminf_N \min_{i\in I_2} \lambda_{\min}(\Phi_i) \geq c_\phi > 0.$$

Assumption 5 is similar to Assumption 8 in Ma et al. (2020). To introduce Theorem 3.2, we define

$$\begin{aligned}
&\dot{\varpi}_{it} = \left(\dot{v}_{t,0}^{(1)\prime}, \dot{v}_{t,1}^{(1)\prime} X_{1,it}, \cdots, \dot{v}_{t,p}^{(1)\prime} X_{p,it}\right)', \\
&\varpi_{it}^0 = \left(\left(O_0^{(1)} v_{t,0}^0\right)', \left(O_1^{(1)} v_{t,1}^0\right)' X_{1,it}, \cdots, \left(O_p^{(1)} v_{t,p}^0\right)' X_{p,it}\right)', \\
&u_i^0 = \left(u_{i,0}^{0\prime}, \cdots, u_{i,p}^{0\prime}\right)', \quad \dot{\Delta}_{t,j} = O_j^{(1)\prime} \dot{v}_{t,j}^{(1)} - v_{t,j}^0, \quad \dot{\Delta}_{t,v} = \left(\dot{\Delta}_{t,0}', \cdots, \dot{\Delta}_{t,p}'\right)', \\
&\dot{\Delta}_{i,j} = O_j^{(1)\prime} \dot{u}_{i,j}^{(1)} - u_{i,j}^0, \quad \dot{\Delta}_{i,u} = \left(\dot{\Delta}_{i,0}', \cdots, \dot{\Delta}_{i,p}'\right)', \\
&D_i^I = \frac{1}{T} \sum_{t=1}^T \mathfrak{f}_{it}(0) \varpi_{it}^0 \varpi_{it}^{0\prime}, \quad D_i^{II} = \frac{1}{T} \sum_{t=1}^T \left[\tau - \mathbf{1}\left\{\epsilon_{it} \leq 0\right\}\right] \varpi_{it}^0, \\
&\mathbb{J}_i\left(\left\{\dot{\Delta}_{t,v}\right\}_{t\in[T]}\right) = \frac{1}{T} \sum_{t=1}^T \left[\mathbf{1}\left\{\epsilon_{it} \leq 0\right\} - \mathbf{1}\left\{\epsilon_{it} \leq \dot{\Delta}_{t,v}' \Psi_{it}^0\right\}\right] \varpi_{it}^0.
\end{aligned}$$

Theorem 3.2 below gives the uniform convergence rate and linear expansion of the factor loading estimators from second stage estimation.

Theorem 3.2 *Suppose Assumptions 1-5 hold. Then for each $j \in \{0, \cdots, p\}$, we have*

(i) $\max_{i\in I_2\cup I_3} \left\|\dot{u}_{i,j}^{(1)} - O_j^{(1)} u_{i,j}^0\right\|_2 = O_p\left(\sqrt{\frac{\log(N\vee T)}{N\wedge T}} \xi_N^2\right)$,

(ii) $\max_{t\in[T]} \left\|\dot{v}_{t,j}^{(1)} - O_j^{(1)} v_{t,j}^0\right\|_2 = O_p\left(\sqrt{\frac{\log(N\vee T)}{N\wedge T}} \xi_N^2\right)$,

(iii) $\dot{\Delta}_{i,u} = \left[D_i^I\right]^{-1} \left[D_i^{II} + \mathbb{J}_i\left(\left\{\dot{\Delta}_{t,v}\right\}_{t\in[T]}\right)\right] + o_p\left((N\vee T)^{-1/2}\right)$ *uniformly over $i \in I_3$.*

Remark 2. Theorem 3.2(i) reports the uniform convergence rate for the factor loading estimators of Θ_j^0 for $i \in I_2 \cup I_3$; Theorem 3.2(ii) reports the uniform convergence rate for the factor estimators of Θ_j^0 for $t \in [T]$; Theorem 3.2(iii) reports the linear expansion for the factor loading estimators of Θ_j^0 for $i \in I_3$. However, the $\mathbb{J}_i\left(\left\{\dot{\Delta}_{t,v}\right\}_{t\in[T]}\right)$ term is not mean-zero and represents the bias induced by the first stage NNR. In the third stage below, we aim to remove such a bias from the linear expansion.

3.3 Third Stage Estimator

In the debiasing stage, we first apply PCA to all independent variables $X_{j,it}$, and then run the row- and column-wise QRs to obtain the final estimators. Below we give Assumptions 6-8 for the PCA procedure and establish the asymptotic linear expansions of PCA estimates in the online supplement. Theorem 3.3 below gives the asymptotic distribution of our final factor and factor loading estimates.

Assumption 6 *For all $j \in [p]$, there exists a constant $M > 0$ such that*

(i) $\mathbb{E}\left(e_{j,it}|\mu_{j,it}\right) = 0$,

(ii) $\mathbb{E}\left[\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left[e_{j,it}e_{j,is} - \mathbb{E}(e_{j,it}e_{j,is})\right]\right]^2 \leq M$,

(iii) for all $i \in [N]$, $\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left|\mathbb{E}(e_{j,it}e_{j,is})\right| \leq M$,

(iv) $\max_{t\in[T]}\frac{1}{N\sqrt{T}}\left\|e_{j,t}'E_j\right\|_2 = O_p\left(\frac{\log N\vee T}{N\wedge T}\right)$, *and for* $\max_{i\in[N]}\frac{1}{T\sqrt{N}}\left\|e_{j,i}'E_j'\right\|_2 = O_p\left(\frac{\log N\vee T}{N\wedge T}\right)$, *where* $e_{j,i} = (e_{j,i1}, ..., e_{j,iT})'$, $e_{j,t} = (e_{j,1t}, ..., e_{j,Nt})'$, *and* $E_j = \{e_{j,it}\}_{i\in[N],t\in[T]}$.

Assumption 7 *For all $j \in [p]$,*

(i) recall that $L_j^0 = \left(l_{j,1}^0, \cdots, l_{j,N}^0\right)'$ *and* $W_j^0 = \left(w_{j,1}^0, \cdots, w_{j,T}^0\right)'$. $\lim_{N\to\infty}\frac{L_j^{0\prime}L_j^0}{N} = \Sigma_{L_j} > 0$ *and* $\lim_{T\to\infty}\frac{W_j^{0\prime}W_j^0}{T} = \Sigma_{W_j} > 0$,

(ii) the r_j eigenvalues of $\Sigma_{L_j}\Sigma_{W_j}$ are distinct.

Assumption 8 *For all $j \in [p]$, there exists a constant $M > 0$ such that*

(i) $\max_{t\in[T]}\mathbb{E}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}l_{j,i}^0 e_{j,it}\right\|_2^2 \leq M$ *and* $\max_{t\in[T]}\frac{1}{NT}e_{j,t}'E_j'L_j^0 = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right)$,

(ii) $\max_{i\in[N]}\mathbb{E}\left\|\frac{1}{\sqrt{T}}\sum_{t=1}^{T}w_{j,t}^0 e_{j,it}\right\|_2^2 \leq M$ *and* $\max_{i\in[N]}\frac{1}{NT}e_{j,t}'E_j'W_j^0 = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right)$.

Assumption 9 *For $\forall j \in [p]$,*

(i) $\mathbb{E}\left[f_{it}(0)e_{j,it}\middle|\mathscr{D}\right] = 0$,

(ii) for each $i \in [N]$ *and* $j \in [p]$, $\{f_{it}(0), f_{it}(0)e_{j,it}\}$ *is stationary strong mixing across t conditional on* $\mathscr{D}$.

Assumptions 6-8 are stronger than those in Bai and Ng (2020) because we strengthen their Assumptions A1(c) and A3 to hold uniformly. Assumption 9 imposes some moment and mixing conditions. Even though $f_{it}(\cdot)$ (the PDF of ϵ_{it} given $\mathscr{D}_e$) is a function of $\{e_{j,it}\}_{j\in[p],i\in[N],t\in[T]}$, we can show that Assumption 9 holds under some reasonable conditions. For example, we consider the location scale model:

$$Y_{it} = \beta_{0,it} + \sum_{j\in[p]} X_{j,it}\beta_{j,it} + \left(\gamma_{0,it} + \sum_{j\in[p]} X_{j,it}\gamma_{j,it}\right) u_{it}, \quad \text{with} \quad X_{j,it} = l_{j,i}^{0\prime} w_{j,t}^0 + e_{j,it},$$

where u_{it} is independent of $\{w_{j,t}^0, e_{j,it}\}_{j\in[p],t\in[T]}$ and $l_{j,i}^0$ and $\beta_{j,it}$ are nonrandom. In this case, $\Theta_{j,it}^0 = \beta_{j,it} + \gamma_{j,it}\mathscr{Q}_\tau(u_{it})$, $\epsilon_{it} = \left(\gamma_{0,it} + \sum_{j\in[p]} X_{j,it}\gamma_{j,it}\right)[u_{it} - \mathscr{Q}_\tau(u_{it})]$, where $\mathscr{Q}_\tau(u_{it})$ is the τ-quantile of u_{it}. It is clear that $f_{it}(\cdot)$ is the function of $\{e_{j,it}\}_{j\in[p],i\in[N],t\in[T]}$ and all factors. However, if u_{it} is independent of sequence $\{e_{j,it}\}_{j\in[p],t\in[T]}$, we observe that $f_{it}(0)$ is the PDF of $u_{it} - \mathscr{Q}_\tau(u_{it})$ evaluated at zero point, which is independent of $\{e_{j,it}\}_{j\in[p],t\in[T]}$. Therefore, Assumption 9(i) holds under mild conditions that u_{it} is independent of the sequence $\{e_{it}\}_{t\in[T]}$ and $\mathbb{E}\left(e_{it}\middle|\mathscr{D}\right) = 0$.

Define

$$\hat{V}_{u_j,i} = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[f_{it}(0)e_{j,it}^2\middle|\mathscr{D}\right] v_{t,j}^0 v_{t,j}^{0\prime}, \quad V_{u_j,i} = \mathbb{E}\left(\hat{V}_{u_j}\right)$$

$$\Omega_{u_j,i} = Var\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_{j,it}v_{t,j}^0(\tau - \mathbf{1}\{\epsilon_{it} \le 0\})\right],$$

$$\hat{V}_{v_j,t}^{(3)} = \frac{1}{N_3}\sum_{i\in I_3} f_{it}(0)e_{j,it}^2 u_{i,j}^0 u_{i,j}^{0\prime}, \quad V_{v_j} = \mathbb{E}\left(\hat{V}_{v_j,t}^{(3)}\right), \quad \Omega_{v_j}^{(3)} = \tau(1-\tau)\frac{1}{N_3}\sum_{i\in I_3}\mathbb{E}\left(e_{j,it}^2 u_{i,j}^0 u_{i,j}^{0\prime}\right).$$

Let $\Sigma_{u_j,i} = O_j^{(1)} V_{u_j,i}^{-1}\Omega_{u_j,i}V_{u_j,i}^{-1}O_j^{(1)\prime}$, $\Sigma_{v_j}^{(3)} = O_j^{(1)}\left(V_{v_j}^{(3)}\right)^{-1}\Omega_{v_j}\left(V_{v_j}^{(3)}\right)^{-1}O_j^{(1)\prime}$, $b_{j,it}^0 = e_{j,it}v_{t,j}^0(\tau - \mathbf{1}\{\epsilon_{it}\le 0\})$ and $\xi_{j,it}^0 = e_{j,it}u_{i,j}^0(\tau - \mathbf{1}\{\epsilon_{it}\le 0\})$. The following theorem establishes the asymptotic properties of the third-stage estimators.

Theorem 3.3 *Suppose that Assumptions 1-9 hold. Suppose that Assumption 13 in Appendix B.3 of the online supplement hold. Let $O_{u,j}^{(1)}$ be the bounded matrix defined in the appendix that is related to rotation matrix $O_j^{(1)}$. Then we have that $\forall j \in [p]$,*

(i) $\hat{u}_{i,j}^{(3,1)} - O_{u,j}^{(1)}u_{i,j}^0 = O_j^{(1)}\hat{V}_{u_j,i}^{-1}\frac{1}{T}\sum_{t=1}^{T} b_{j,it}^0 + \mathcal{R}_{i,u}^j$ *and* $\sqrt{T}\left(\hat{u}_{i,j}^{(3,1)} - O_{u,j}^{(1)}u_{i,j}^0\right) \rightsquigarrow \mathcal{N}\left(0, \Sigma_{u_j,i}\right)$ $\forall i \in I_3$,

(ii) $\hat{v}_{t,j}^{(3,1)} - \left(O_{u,j}^{(1)\prime}\right)^{-1}v_{t,j}^0 = O_j^{(1)}\left(\hat{V}_{v_j,t}^{(3)}\right)^{-1}\frac{1}{N_3}\sum_{i\in I_3}\xi_{j,it}^0 + \mathcal{R}_{t,v}^j$ *and* $\sqrt{N_3}\left(\hat{v}_{t,j}^{(3,1)} - O_{v,j}^{(1)}v_{t,j}^0\right) \rightsquigarrow \mathcal{N}\left(0, \Sigma_{v_j}^{(3)}\right)$ $\forall t \in [T]$,

where $\max_{i\in I_3}\left|\mathcal{R}_{i,u}^j\right| = o_p\left((N\vee T)^{-1/2}\right)$, *and* $\max_{t\in[T]}\left|\mathcal{R}_{t,v}^j\right| = o_p\left((N\vee T)^{-\frac{1}{2}}\right)$.

Remark 3. Theorem 3.3 reports the linear expansions for the factor and factor loading estimators for each slope matrix obtained in Step 3. Compared with Chernozhukov et al. (2019), Theorem 3.3 obtains the uniform convergence rate rather than the point-wise result for the reminder terms $\mathcal{R}_{i,u}^j$ and $\mathcal{R}_{t,v}^j$. In addition, since the regressors in the debiasing step are obtained from Step 2 instead of Step 1, we don't have independence between the regressors and error terms, which makes the proof more complex than that in Chernozhukov et al. (2019). See the proof in the appendix on how to handle the dependence. Assumption 13 in the online supplement is a regularity condition on the density of ϵ_{it}.

Following Theorem 3.3 and estimators defined in Table 1, we have that $\forall j \in [p]$, $\forall i \in [N]$ and $\forall t \in [T]$,

$$\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)} u_{i,j}^{0} = O_{j}^{(b)} \hat{V}_{u_j,i}^{-1} \frac{1}{T} \sum_{t=1}^{T} v_{t,j}^{0} b_{j,it}^{0} + \mathcal{R}_{i,u}^{j},$$
$$\hat{v}_{t,j}^{(a,b)} - \left(O_{u,j}^{(b)\prime}\right)^{-1} v_{t,j}^{0} = O_{j}^{(b)} \left(\hat{V}_{v_j,t}^{(a)}\right)^{-1} \frac{1}{N_a} \sum_{i \in I_a} u_{i,j}^{0} \xi_{j,it}^{0} + \mathcal{R}_{t,v}^{j},$$

where $\hat{V}_{v_j,t}^{(a)} = \frac{1}{N_a} \sum_{i \in I_a} f_{it}(0) e_{j,it}^2 u_{i,j}^{0} u_{i,j}^{0\prime}$, $a \in [3]$ and $b \in [3] \setminus \{a\}$.

Given the above estimates for the factors and factor loadings, we can estimate $\Theta_{j,it}^{0}$ by

$$\hat{\Theta}_{j,it} = \frac{1}{2} \sum_{a \in [3]} \sum_{b \in [3] \setminus \{a\}} \left\{ \hat{u}_{i,j}^{(a,b)\prime} \hat{v}_{t,j}^{(a,b)} \right\} \mathbf{1}_{ia}$$

where $\mathbf{1}_{ia} = \mathbf{1}\{i \in I_a\}$ for $i \in [N]$. Let $\Xi_{j,it}^{0} = \frac{1}{T} v_{t,j}^{0\prime} \Sigma_{u_j,i} v_{t,j}^{0} + \sum_{a=1}^{3} \frac{1}{N_a} \mathbf{1}_{ia} u_{i,j}^{0\prime} \Sigma_{v_j}^{a} u_{i,j}^{0}$. The following proposition studies the asymptotic properties of $\hat{\Theta}_{j,it}$.

Proposition 3.4 *Under Assumptions 1-9 and Assumption 13, $\forall j \in [p]$ we have*

(i) $\hat{\Theta}_{j,it} - \Theta_{j,it}^{0} = \sum_{a=1}^{3} u_{i,j}^{0\prime} \left(\hat{V}_{v_j,t}^{(a)}\right)^{-1} \frac{1}{N_a} \sum_{i^* \in I_a} \xi_{j,i^*t} \mathbf{1}_{i^*a} + v_{t,j}^{0\prime} \hat{V}_{u_j}^{-1} \frac{1}{T} \sum_{t^*=1}^{T} b_{j,it^*}^{0} + \mathcal{R}_{it}^{j}$, *where* $\max_{i \in I_3, t \in [T]} \left|\mathcal{R}_{it}^{j}\right| = o_p\left((N \vee T)^{-1/2}\right)$,

(ii) $\max_{i \in [N], t \in [T]} \left|\hat{\Theta}_{j,it} - \Theta_{j,it}^{0}\right| = O_p\left(\sqrt{\frac{\log N \vee T}{N \wedge T}}\right)$,

(iii) $\left(\Xi_{j,it}^{0}\right)^{-1/2} \left(\hat{\Theta}_{j,it} - \Theta_{j,it}^{0}\right) \rightsquigarrow \mathcal{N}(0, 1)$.

Remark 4. Proposition 3.4 establishes the distribution theory for the slope estimators. Recall that we remove the principle component from the independent variables $X_{j,it}$ which is the key point in the debiasing step and why we don't have the distribution theory result for the intercept estimates $\hat{\Theta}_{0,it}$ in the current framework. However, once we have the distribution theory for the slope estimates, we can follow Chen et al. (2021) and obtain a new estimator for $\Theta_{0,it}^{0}$ from the smoothed quantile regression and establish its distribution theory. We leave this for the further research.

To make inference for $u_{i,j}^{0}$, $v_{t,j}^{0}$, and $\Theta_{j,it}^{0}$, one needs to estimate their asymptotic variances $\Sigma_{u_j,i}$, Σ_{v_j} and $\Xi_{j,it}^{0}$ consistently. Let $k(\cdot)$ be a PDF-type kernel function and $K(\cdot)$ be its survival function such that $\int k(u)du = 1$ and $K(u) := \int_{u}^{\infty} k(v)dv$. Let h_N be the bandwidth such that $h_N \to 0$

with $N \to \infty$. Define $K_{h_N}(\cdot) = K(\frac{\cdot}{h_N})$, $k_{h_N}(\cdot) = \frac{1}{h_N}k(\frac{\cdot}{h_N})$. Let $\hat{\epsilon}_{it} = Y_{it} - \hat{\Theta}_{0,it} - \sum_{j\in[p]} X_{j,it}\hat{\Theta}_{j,it}$, $\hat{\mathrm{v}}_{t,s,j} = \frac{1}{6}\sum_{a\in[3]}\sum_{b\in[3]\setminus\{a\}} \hat{v}_{t,j}^{(a,b)}\hat{v}_{s,j}^{(a,b)\prime}$, and $\hat{\mathrm{u}}_{i,i,j} = \frac{1}{2}\sum_{a\in[3]}\sum_{b\in[3]\setminus\{a\}} \hat{u}_{i,j}^{(a,b)}\hat{u}_{i,j}^{(a,b)\prime}\mathbf{1}_{ia}$. Define

$$
\begin{aligned}
\hat{\mathbb{V}}_{u_j} &= \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} k_{h_N}(\hat{\epsilon}_{it})\hat{e}_{j,it}^2\hat{\mathrm{v}}_{t,t,j}, \quad \hat{\mathbb{V}}_{v_j} = \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} k_{h_N}(\hat{\epsilon}_{it})\hat{e}_{j,it}^2\hat{\mathrm{u}}_{i,i,j}, \\
\hat{\Omega}_{u_j} &= \frac{1}{NT}\sum_{i\in[N]}\left\{\sum_{t\in[T]}\tau(1-\tau)\hat{e}_{j,it}^2\hat{\mathrm{v}}_{t,t,j} + \sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1} S_{j,its} + \sum_{t=1+T_1}^{T}\sum_{s=t-T_1}^{t-1} S_{j,its}\right\}, \\
\hat{\Omega}_{v_j} &= \frac{\tau(1-\tau)}{NT}\sum_{i\in[N]}\sum_{t\in[T]}\hat{e}_{j,it}^2\hat{\mathrm{u}}_{i,i,j}, \quad \hat{\Sigma}_{u_j} = \hat{\mathbb{V}}_{u_j}^{-1}\hat{\Omega}_{u_j}\hat{\mathbb{V}}_{u_j}^{-1}, \quad \hat{\Sigma}_{v_j} = \hat{\mathbb{V}}_{v_j}^{-1}\hat{\Omega}_{v_j}\hat{\mathbb{V}}_{v_j}^{-1},
\end{aligned}
$$

where $S_{j,its} = \hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}\left[\tau - K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)\right]\left[\tau - K\left(\frac{\hat{\epsilon}_{is}}{h_N}\right)\right]$. We further define

$$
\hat{\Xi}_{j,it} = \frac{1}{2}\sum_{a\in[3]}\sum_{b\in[3]\setminus\{a\}}\left(\frac{1}{T}\hat{v}_{t,j}^{(a,b)\prime}\hat{\Sigma}_{u_j}\hat{v}_{t,j}^{(a,b)} + \frac{1}{N_a}\mathbf{1}_{ia}\hat{u}_{i,j}^{(a,b)\prime}\hat{\Sigma}_{v_j}\hat{u}_{i,j}^{(a,b)}\right).
$$

Let $F_{i,ts}(\cdot,\cdot)$ and $f_{i,ts}(\cdot,\cdot)$ denote the joint CDF and PDF of $(\epsilon_{it},\epsilon_{is})$ given $\mathscr{D}_e$, respectively. To justify the consistency of the variance estimators, we add the following assumption.

Assumption 10 *(i)* $\int_{-\infty}^{+\infty} k(u)du = 1$, $\int_{-\infty}^{+\infty} k(u)u^j du = 0$ *for* $j \in \{1,\cdots,m-1\}$ *and* $\int_{-\infty}^{+\infty} k(u)u^m du \neq 0$ *for* $m \geq 1$.

(ii) $h_N \to 0$ *and* $\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{1/4}\frac{\xi_N^2}{h_N} \to 0$.

(iii) $T_1 \to \infty$ *and* $\sqrt{\frac{\log(N\vee T)}{N\wedge T}}\frac{\xi_N^2 T_1}{h_N^2} \to 0$.

(iv) $f_{it}(c)$ *is* m *times continuously differentiable with respect to* c *and* $f_{i,ts}(c_1,c_2)$ *is* m *times continuously differentiable with respect to* (c_1,c_2).

(v) $\forall i \in [N]$, $V_{u_j,i} = V_{u_j}$ *and* $\Omega_{u_j,i} = \Omega_{u_j}$.

(vi) $\forall a \in [3]$, $V_{v_j}^{(a)} = \frac{1}{N}\sum_{i\in[N]}\mathbb{E}\left[f_{it}(0)e_{j,it}^2 u_{i,j}^0 u_{i,j}^{0\prime}\right] + o_p(1)$ *and* $\Omega_{v_j}^{(a)} = \frac{\tau(1-\tau)}{N}\sum_{i\in[N]}\mathbb{E}\left(e_{j,it}^2 u_{i,j}^0 u_{i,j}^{0\prime}\right) + o_p(1)$.

Assumption 10(i)-(iv) are standard for consistent estimation of the asymptotic variance matrix; see, e.g., Chen (2022) and Galvao and Kato (2016). Assumption 10(v) imposes the homogeneity moment condition across individuals, and Assumption 10(vi) assumes the moments calculated from subsamples are close to those from the full sample given the random splitting. Under Assumption 10, following the idea of Chen (2022), we establish in Lemma B.33 of the online supplement the consistency of $\hat{\Sigma}_{u_j}$ and $\hat{\Sigma}_{v_j}$. Similar conclusions hold for the other estimates.

4 Specification Tests

In this section, we consider two specification tests under different rank conditions.

4.1 Testing for Homogeneity across Individuals or Time

When $K_j = 1$ for some $j \in [p]$, it is interesting to test whether the matrix Θ_j^0 is homogeneous across individuals (i.e., row-wise) or across time (i.e., column-wise). For these two cases, we can write factors and factor loadings as

$$u_{i,j}^0 = u_j + c_{i,j}^u \quad \text{and} \quad v_{t,j}^0 = v_j + c_{t,j}^v, \quad \text{respectively,}$$

where $u_j = \frac{1}{N}\sum_{i=1}^N u_{i,j}^0$ and $v_j = \frac{1}{T}\sum_{t=1}^T v_{t,j}^0$. For the homogeneity across individuals, the null and alternative hypotheses can be written as

$$H_0^I : c_{i,j}^u = 0 \quad \forall i \in [N] \quad v.s. \quad H_1^I : c_{i,j}^u \neq 0 \text{ for some } i \in [N]. \tag{4.1}$$

Similarly, for the homogeneity across time, the null and alternative hypotheses can be written as

$$H_0^{II} : c_{t,j}^v = 0 \quad \forall t \in [T] \quad v.s. \quad H_1^{II} : c_{t,j}^v \neq 0 \text{ for some } t \in [T]. \tag{4.2}$$

Note that we aim to test the two null hypotheses separately. That is, we can test for homogeneous slope across individuals while allowing for heterogeneous slopes across time and vice versa. This is different from the majority of the literature which either tests for slope homogeneity across individuals while assuming the slopes are homogeneous across time or tests for structural breaks across time while assuming the slopes are homogeneous across individuals.

We first consider testing H_0^I. Following the lead of Castagnetti et al. (2015), we define[2]

$$\begin{aligned} S_{u_j}^{(a,b)} &= \max_{i \in I_a} T(\hat{u}_{i,j}^{(a,b)} - \hat{\bar{u}}_j^{(a,b)})'\hat{\Sigma}_{u_j}^{-1}(\hat{u}_{i,j}^{(a,b)} - \hat{\bar{u}}_j^{(a,b)}) \quad \text{and} \\ S_{u_j} &= \max\left(S_{u_j}^{(3,1)}, S_{u_j}^{(2,3)}, S_{u_j}^{(1,2)}\right), \end{aligned} \tag{4.3}$$

where $\hat{\bar{u}}_j^{(a,b)} = \frac{1}{N_a}\sum_{i \in I_a} \hat{u}_{i,j}^{(a,b)}$. Similarly, to test for H_0^{II}, we construct

$$S_{v_j} = \max\left(\tilde{S}_{v_j}^{(3,1)}, \tilde{S}_{v_j}^{(2,3)}, \tilde{S}_{v_j}^{(1,3)}\right),$$

where

$$S_{v_j}^{(a,b)} = \max_{t \in [T]} N(\hat{v}_{t,j}^{(a,b)} - \hat{\bar{v}}_j^{(a,b)})'\hat{\Sigma}_{v_j}^{-1}(\hat{v}_{t,j}^{(a,b)} - \hat{\bar{v}}_j^{(a,b)}), \quad \tilde{S}_{v_j}^{(a,b)} = \frac{1}{2}S_{v_j}^{(a,b)} - \mathsf{b}(T),$$

$\hat{\bar{v}}_j^{(a,b)} = \frac{1}{T}\sum_{t=1}^T \hat{v}_{t,j}^{(a,b)}$, and $\mathsf{b}(n) = \log n - \frac{1}{2}\log\log n - \log\Gamma(\frac{1}{2})$ for $n \in \{N, T, NT\}$.

[2] Alternatively, we can also define $S_{u_j}^o = \max\left(S_{u_j}^{(3,2)}, S_{u_j}^{(2,1)}, S_{u_j}^{(1,3)}\right)$. It is easy to show that this statistic shares the same asymptotic null distribution as S_{u_j}. But due to the unknown dependence structure between the two, we cannot take the maximum or the other continuous function of S_{u_j} and $S_{u_j}^o$ as a new test statistic.

To proceed, we introduce some notation. Recall that $b^0_{j,it} = e_{j,it} v^0_{t,j} (\tau - \mathbf{1}\{\epsilon_{it} \leq 0\})$ and $\xi^0_{j,it} = e_{j,it} u^0_{i,j} (\tau - \mathbf{1}\{\epsilon_{it} \leq 0\})$. Define

$$\mathfrak{b}^{(1)}_{j,it} = \hat{V}^{-1}_{u_j} b^0_{j,it}, \quad \mathfrak{b}^{(2)}_{j,it} = \left(\hat{V}^{(3)}_{v_j,t}\right)^{-1} \xi^0_{j,it}, \ \mathfrak{b}^{(3)}_{j,it} = \left(\hat{V}^{(2)}_{v_j,t}\right)^{-1} \xi^0_{j,it}, \text{ and } \mathfrak{b}^{(4)}_{j,it} = \left(\hat{V}^{(1)}_{v_j,t}\right)^{-1} \xi^0_{j,it}.$$

Let $\mathfrak{B}^{(\ell)}_{j,t} = \left(\mathfrak{b}^{(\ell)\prime}_{j,1t}, \cdots, \mathfrak{b}^{(\ell)\prime}_{j,Nt}\right)'$ for $\ell \in [4]$. Define

$$\Sigma^{(1)}_{\mathfrak{B},j} = \frac{1}{T} \sum_{t=1}^{T} \sum_{s=1}^{T} \mathbb{E}\left(\mathfrak{B}^{(1)}_{j,t} \mathfrak{B}^{(1)\prime}_{j,s}\right), \quad \Sigma^{(2)}_{\mathfrak{B},j} = \frac{1}{N_3} \sum_{i \in I_3} \mathbb{E}\left(\mathfrak{B}^{(2)}_{j,i} \mathfrak{B}^{(2)\prime}_{j,i}\right),$$
$$\Sigma^{(3)}_{\mathfrak{B},j} = \frac{1}{N_2} \sum_{i \in I_2} \mathbb{E}\left(\mathfrak{B}^{(3)}_{j,i} \mathfrak{B}^{(3)\prime}_{j,i}\right), \quad \text{and } \Sigma^{(4)}_{\mathfrak{B},j} = \frac{1}{N_1} \sum_{i \in I_1} \mathbb{E}\left(\mathfrak{B}^{(4)}_{j,i} \mathfrak{B}^{(4)\prime}_{j,i}\right).$$

We add the following two assumptions.

Assumption 11 *$\forall j \in [p]$, we assume*

$$\bar{\lambda} \geq \lambda_{\max}\left(\Sigma_{u_j}\right) \geq \lambda_{\min}\left(\Sigma_{u_j}\right) \geq \underline{\lambda} > 0, \quad \bar{\lambda} \geq \lambda_{\max}\left(\Sigma_{v_j}\right) \geq \lambda_{\min}\left(\Sigma_{v_j}\right) \geq \underline{\lambda} > 0.$$

Assumption 12 *(i) There exists a high dimensional Gaussian vector $\mathbb{Z}^{(1)}_{\mathfrak{B}} \sim N\left(0, \Sigma^{(1)}_{\mathfrak{B},j}\right)$ such that $\left\|\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \mathfrak{B}^{(1)}_{j,t} - \mathbb{Z}^{(1)}_{\mathfrak{B}}\right\|_{\max} = o_p(1)$.*

(ii) There exists high dimensional Gaussian vectors $\mathbb{Z}^{(\ell)}_{\mathfrak{B}} \sim N\left(0, \Sigma^{(\ell)}_{\mathfrak{B},j}\right)$ for $\ell = 2, 3, 4$ such that $\left(\mathbb{Z}^{(2)}_{\mathfrak{B}}, \mathbb{Z}^{(3)}_{\mathfrak{B}}, \mathbb{Z}^{(4)}_{\mathfrak{B}}\right)$ are independent,

$$\left\|\frac{1}{\sqrt{N_3}} \sum_{i \in I_3} \mathfrak{B}^{(2)}_{j,i} - \mathbb{Z}^{(2)}_{\mathfrak{B}}\right\|_{\max} = o_p(1), \quad \left\|\frac{1}{\sqrt{N_2}} \sum_{i \in I_2} \mathfrak{B}^{(3)}_{j,i} - \mathbb{Z}^{(3)}_{\mathfrak{B}}\right\|_{\max} = o_p(1), \text{ and}$$
$$\left\|\frac{1}{\sqrt{N_1}} \sum_{i \in I_1} \mathfrak{B}^{(4)}_{j,i} - \mathbb{Z}^{(4)}_{\mathfrak{B}}\right\|_{\max} = o_p(1).$$

Assumption 11 implies that both Σ_{u_j} and Σ_{v_j} are well behaved. Assumption 12 imposes that we can approximate high dimensional vectors $\frac{1}{\sqrt{T}} \sum_{t \in [T]} \mathfrak{B}^{(1)}_{j,t}$, $\frac{1}{\sqrt{N_3}} \sum_{i \in I_3} \mathfrak{B}^{(2)}_{j,i}$, $\frac{1}{\sqrt{N_2}} \sum_{i \in I_2} \mathfrak{B}^{(3)}_{j,i}$ and $\frac{1}{\sqrt{N_1}} \sum_{i \in I_1} \mathfrak{B}^{(4)}_{j,i}$ by four Gaussian vectors. Similar conditions have been imposed in the literature; see, e.g., Assumption SA3 Lu and Su (2021).

The following theorem reports the asymptotic properties of S_{u_j} and S_{v_j} under the respective null and alternative hypotheses.

Theorem 4.1 *Suppose that Assumptions 1-12 and Assumptions 13 in the online supplement hold and $(N, T) \to \infty$. Then*

(i) *Under* H_0^I, *we have* $\mathbb{P}\left(\frac{1}{2}S_{u_j} \le x + \mathsf{b}(N)\right) \to e^{-e^{-x}}$; *and under* H_0^{II}, *we have* $\mathbb{P}\left(S_{v_j} \le x\right) \to e^{-3e^{-x}}$.

(ii) *Under* H_1^I, *if* $\frac{T}{\log N}\max_{i\in[N]}\left\|c_{i,j}^u\right\|_2^2 \to \infty$, *we have* $\mathbb{P}\left(S_{u_j} > c_{\alpha,1\cdot N}\right) \to 1$ *with* $c_{\alpha,1\cdot N} = 2\mathsf{b}(N) - \log\left|\log(1-\alpha)\right|^2$ *and* α *is the significance level. Under* H_1^{II}, *if* $\frac{N}{\log T}\max_{t\in[T]}\left\|c_{t,j}^v\right\|_2^2 \to \infty$, *we have* $\mathbb{P}\left(S_{v_j} > c_{\alpha,2}\right) \to 1$ *with* $c_{\alpha,2} = -\log\left(-\frac{1}{3}\log(1-\alpha)\right)$.

Remark 5. Theorem 4.1 implies that our test statistics follow the Gumbel distributions asymptotically under the null, are consistent under the global alternatives, and have non-trivial power against the local alternatives. The power function of S_{u_j} approaches 1 as long as $\frac{T}{\log N}\max_{i\in[N]}\left\|c_{i,j}^u\right\|_2^2$ diverges to infinity as $(N,T)\to\infty$.

4.2 Test for an Additive Structure

When $K_j = 2$ for some $j \in [p]$, it is interesting to test whether $\Theta_{j,it}^0$ exhibits the additive structure which is widely assumed in a two-way fixed effects model. That is, one may test the following null hypothesis

$$H_0^{III} : \Theta_{j,it}^0 = \lambda_{j,i} + f_{j,t},\ \forall (i,t) \in [N]\times[T], \tag{4.4}$$

The alternative hypothesis H_1^{III} is the negation of H_0^{III}.

Let $\bar{\Theta}_{j,i\cdot} = \frac{1}{T}\sum_{t\in[T]}\Theta_{j,it}^0$, $\bar{\Theta}_{j,\cdot t}^{I_a} = \frac{1}{N_a}\sum_{i\in I_a}\Theta_{j,it}^0$, and $\bar{\Theta}_j^{I_a} = \frac{1}{N_aT}\sum_{i\in I_a}\sum_{t\in[T]}\Theta_{j,it}^0$ for $a \in [3]$. Define

$$\Theta_{j,it}^* = \Theta_{j,it}^0 - \bar{\Theta}_{j,i\cdot} - \bar{\Theta}_{j,\cdot t}^{I_a} + \bar{\Theta}_j^{I_a}, \quad \forall i \in I_a, t\in[T], j\in[p].$$

Note that $\Theta_{j,it}^* = 0\ \forall (i,t) \in [N]\times[T]$ under H_0^{III}. So we can propose a test for H_0^{III} based on estimates of $\Theta_{j,it}^*$. Define

$$\bar{\hat{\Theta}}_{j,i\cdot} = \frac{1}{T}\sum_{t\in[T]}\hat{\Theta}_{j,it},\ \bar{\hat{\Theta}}_{j,\cdot t}^{I_a} = \frac{1}{N_a}\sum_{i\in I_a}\hat{\Theta}_{j,it},\ \text{and}\ \bar{\hat{\Theta}}_j^{I_a} = \frac{1}{N_aT}\sum_{i\in I_a}\sum_{t\in[T]}\hat{\Theta}_{j,it}$$

for $a \in [3]$. Then, we can define the sample analogue of $\hat{\Theta}_{j,it}^*$ as

$$\hat{\Theta}_{j,it}^{*a} = \hat{\Theta}_{j,it} - \bar{\hat{\Theta}}_{j,i\cdot} - \bar{\hat{\Theta}}_{j,\cdot t}^{I_a} + \bar{\hat{\Theta}}_j^{I_a}, \quad \forall i\in I_a, t\in[T], j\in[p].$$

Its corresponding asymptotic variance can be estimated by $\hat{\Sigma}_{j,it}^*$ defined as

$$\begin{aligned}\hat{\Sigma}_{j,it}^* = &\frac{1}{2}\sum_{a\in[3]}\sum_{b\in[3]\setminus\{a\}}\frac{1}{N_a}\left(\hat{u}_{i,j}^{(a,b)} - \bar{\hat{u}}_j^{(a,b)}\right)'\hat{\Sigma}_{v_j}\left(\hat{u}_{i,j}^{(a,b)} - \bar{u}_j^{(a,b)}\right)\mathbf{1}_{ia}\\ &+\frac{1}{6}\sum_{a\in[3]}\sum_{b\in[3]\setminus\{a\}}\frac{1}{T}\left(\hat{v}_{t,j}^{(a,b)} - \bar{\hat{v}}_j^{(a,b)}\right)'\hat{\Sigma}_{u_j}\left(\hat{v}_{t,j}^{(a,b)} - \bar{\hat{v}}_j^{(a,b)}\right),\end{aligned}$$

where $\bar{\hat{u}}_j^{(a,b)} = \frac{1}{N_a}\sum_{i\in I_a}\hat{u}_{i,j}^{(a,b)}$ and $\bar{\hat{v}}_j^{(a,b)} = \frac{1}{T}\sum_{t\in[T]}\hat{v}_{t,j}^{(a,b)}$. Then, the final test statistic is

$$S_{NT} = \max_{i\in[N],t\in[T]}\left(\hat{\Theta}_{j,it}^*\right)^2/\hat{\Sigma}_{j,it}^*.$$

The following theorem studies the asymptotic properties of S_{NT} under the null and alternatives.

Theorem 4.2 *Suppose Assumptions 1-13 hold and* $(N,T)\to\infty$. *Under* H_0^{III},

$$\mathbb{P}\left(\frac{1}{2}S_{NT}\leq x+\mathsf{b}(NT)\right)\to e^{-e^{-x}};$$

under H_1^{III}, *if* $\frac{N\wedge T}{\log NT}\max_{i\in[N],t\in[T]}\left|\Theta_{j,it}^{*}\right|^{2}\to\infty$, *then we have* $\mathbb{P}\left(S_{NT}>c_{\alpha,3\cdot NT}\right)\to 1$ *with* $c_{\alpha,3\cdot NT}=2\mathsf{b}(NT)-\log\left|\log\left(1-\alpha\right)\right|^{2}$.

Similar remark after Theorem 4.1 holds here. In particular, S_{NT} has the desired asymptotic Gumbel distribution under the null and is consistent under the global alternative.

5 Monte Carlo Simulations

In this section, we conduct a set of Monte Carlo simulations to show the finite sample performance of our low-rank quantile regression estimates and specification tests.

5.1 Data Generating Processes

Below we will consider the following data generating process (DGP):

$$Y_{it}=\Theta_{0,it}+X_{it}'\Theta_{it}+(1+0.1X_{1,it}+0.1X_{2,it})u_{it},$$

where $X_{it}=(X_{1,it},X_{2,it})'$, $\Theta_{it}=(\Theta_{1,it},\Theta_{2,it})'$, $\Theta_{0,it}$ is the intercept term which will be specified via the IFEs.

First, we consider four DGPs where the rank of each slope matrix is 1:

DGP 1: **Constant slope with i.i.d. error.** Let $\Theta_{0,it}=\lambda_i f_t$, where $\lambda_i, f_t\sim N(2,5)$. Then let $\Theta_{1,it}=\Theta_{2,it}=2\ \forall\,(i,t)\in[N]\times[T]$, and $X_{j,it}=l_{j,i}^{0}w_{j,t}^{0}+U(0,1)$ for $j\in\{1,2\}$ with $l_{1,i}^{0}$, $l_{2,i}^{0}$, $w_{1,t}^{0}$ and $w_{2,t}\sim U(0,1)$. $u_{it}\overset{i.i.d}{\sim}\frac{t(3)}{\sqrt{3}}$.

DGP 2: **Factor slope with rank 1 and i.i.d. error.** Same as DGP 1 except that the slope coefficients follow the factor structure with one factor rather than homogeneous across both individuals and time, i.e., $\Theta_{1,it}=a_{1,i}g_{1,t}$, $\Theta_{2,it}=a_{2,i}g_{2,t}$, where $a_{1,i}$, $g_{1,t}$, $a_{2,i}$ and $g_{2,t}\sim N(0,2)$. Except these, all other settings remain the same as in DGP 1.

DGP 3: **Constant slope with serial correlation.** Same as DGP 1 except that we set $u_{it}=0.2u_{i,t-1}+\varepsilon_{it}$, $\varepsilon_{it}\overset{i.i.d}{\sim}\frac{t(3)}{\sqrt{3}}$ and all other settings remain the same.

DGP 4: **Factor slope with rank 1 and serial correlation.** Same as DGP 2 except that we set $u_{it}=0.2u_{i,t-1}+\varepsilon_{it}$, $\varepsilon_{it}\overset{i.i.d}{\sim}\frac{t(3)}{\sqrt{3}}$ and all other settings remain the same.

For the case that the rank of the slope matrix is 2, we consider two DGPs which have the additive structure for the slope coefficient of one regressor and the factor structure with two factors for the slope coefficient of another regressor. Specifically,

DGP 5: **Additive and factor slopes with i.i.d. error.** $\Theta_{0,it} = \lambda_i f_t$, $\Theta_{1,it} = a_{1,i} + g_{1,t}$ and $\Theta_{2,it} = a_{2,i}' g_{2,t}$ such that $a_{2,i} = (a_{2,i,1}, a_{2,i,2})'$, $g_{2,t} = (g_{2,t,1}, g_{2,t,2})'$, $\lambda_i, f_t, a_{1,i}, g_{1,i} \sim N(2,5)$ and $a_{2,i,1}, a_{2,i,2}, g_{2,i,1}, g_{2,i,2} \sim N(0,5)$. Moreover, $X_{1,it} = l_{1,i}^0 w_{1,t}^0 + U(0,4)$, $X_{2,it} = l_{2,i}^0 w_{2,t}^0 + Beta(2,5)$ with $l_{1,i}^0, w_{1,t}^0 \sim U(0,4)$ and $l_{2,i}^0, w_{2,t}^0 \sim Beta(2,5)$. $u_{it} \overset{i.i.d}{\sim} \frac{t(3)}{\sqrt{3}}$.

DGP 6: **Additive and factor slopes with serial correlation.** Same as DGP 5 except that the error u_{it} follows AR(1) process like in DGPs 3 and 4.

5.2 Estimation Results

For $\Theta \in \mathbb{R}^{N \times T}$, define $RMSE(\Theta) = \frac{1}{\sqrt{NT}} \left\| \Theta - \Theta^0 \right\|_F$. Table 3 shows the RMSEs of the full-sample low rank matrix estimates under different quantiles for each DGP. As Theorem 3.1(i) predicts, the RMSEs decrease as both N and T increase. Given the fact that $N \wedge T = T$ in the simulations, the decrease of the RMSEs is largely driven by the increase of T.

Table 4 reports the frequency of correct rank estimation by the singular value thresholding (SVT) approach based on 1000 replications. Note that the true ranks of the intercept and slope matrices in DGPs 1-4 and 5-6 are 1 and 2, respectively. The results show that the SVT can accurately determine the correct rank of the coefficient matrices in all DGPs for all three quantile indices under investigation.

5.3 Test Results

In Section 4, we define S_{u_j} and S_{v_j} as the sup-type test statistics. Table 5 reports the empirical size and power at the 5% nominal level for the null hypothesis that the slope coefficient is homogeneous across either i or t. The results in DGPs 1 and 3 give the empirical size, and those in DGPs 2 and 4 give the empirical power. As the results in Table 5 indicate, our tests have reasonable size despite the fact that they are slightly conservative like most extreme-value based sup-tests in the literature. In terms of power, out tests have superb power in both DGPs across all three quantile indices.

Table 6 shows the empirical size and power of our test for DGPs 5 and 6. The findings are similar to those in Table 5. In particular, our tests are a bit conservative under the null. The empirical power tends to 1 quickly as T increases.

6 Empirical Study

In this section we consider two empirical applications: the heterogeneous investment equation and the heterogeneous quantile effect of foreign direct investment on unemployment.

Table 3: RMSEs of low rank estimates in the full sample

DGP	N	T	$\tau = 0.25$			$\tau = 0.50$			$\tau = 0.75$		
			$\tilde{\Theta}_0$	$\tilde{\Theta}_1$	$\tilde{\Theta}_2$	$\tilde{\Theta}_0$	$\tilde{\Theta}_1$	$\tilde{\Theta}_2$	$\tilde{\Theta}_0$	$\tilde{\Theta}_1$	$\tilde{\Theta}_2$
1	75	35	0.922	0.324	0.329	1.242	0.288	0.297	1.839	0.609	0.658
		70	0.707	0.280	0.275	0.819	0.220	0.203	1.266	0.519	0.523
	150	35	1.012	0.337	0.340	1.099	0.258	0.262	1.932	0.661	0.623
		70	0.745	0.272	0.265	0.825	0.205	0.206	1.324	0.522	0.504
2	75	35	0.871	0.521	0.505	0.881	0.704	0.680	1.278	1.055	0.970
		70	0.692	0.401	0.373	0.672	0.553	0.537	1.057	0.744	0.768
	150	35	0.877	0.507	0.480	1.022	0.790	0.815	1.334	1.018	1.040
		70	0.703	0.374	0.373	0.689	0.531	0.538	1.059	0.829	0.787
3	75	35	0.945	0.334	0.329	1.115	0.280	0.265	1.876	0.630	0.627
		70	0.682	0.286	0.279	0.809	0.230	0.214	1.244	0.486	0.492
	150	35	0.973	0.334	0.331	1.211	0.287	0.291	1.771	0.590	0.612
		70	0.757	0.274	0.272	0.801	0.208	0.195	1.360	0.494	0.527
4	75	35	0.885	0.515	0.519	0.915	0.693	0.723	1.382	1.125	1.037
		70	0.669	0.393	0.384	0.652	0.511	0.520	1.053	0.812	0.774
	150	35	0.889	0.513	0.483	0.905	0.761	0.686	1.409	1.118	1.133
		70	0.725	0.376	0.377	0.717	0.547	0.565	1.058	0.724	0.775
5	75	35	0.218	0.268	0.450	0.307	0.308	0.606	0.844	0.466	0.936
		70	0.174	0.226	0.414	0.213	0.200	0.493	0.610	0.388	0.838
	150	35	0.236	0.245	0.458	0.299	0.291	0.634	1.299	0.863	1.778
		70	0.174	0.214	0.423	0.216	0.203	0.450	0.629	0.377	0.679
6	75	35	0.253	0.267	0.293	0.382	0.227	0.421	1.293	0.609	0.892
		70	0.207	0.239	0.278	0.261	0.192	0.366	0.576	0.287	0.415
	150	35	0.225	0.254	0.269	0.363	0.225	0.422	1.486	0.695	0.992
		70	0.193	0.254	0.263	0.254	0.171	0.379	0.797	0.391	0.551

6.1 Investment Equation

In this subsection, we revisit the investment equation. Fazzari et al. (1988) point out that investment may show sensitivity to movements in cash flow when firms face constraints for external finance. Since Fazzari et al. (1988), there has been a large literature on the effect of cash flow on the corporate investment; see Devereux and Schiantarelli (1990), Gilchrist and Himmelberg (1995), Kaplan and Zingales (1995), Cleary (1999), Rauh (2006), and Almeida and Campello (2007), among others. Using the panel dataset, we consider the scaled version of the investment equation as follows:

$$\frac{I_{it}}{K_{i,t-1}} = \Theta_{0,it} + \Theta_{1,it}\frac{CF_{it}}{K_{i,t-1}} + \Theta_{2,it}q_{i,t-1} + u_{it},$$

where I is the corporate investment, CF is the cash flow, q is the Tobin's q, K is the capital stock and u is the innovation. $\Theta_{0,it}$ refers to the fixed effects (FEs). Rather than the mean estimation, Galvao and Wang (2015) estimate the effects of the firm's cash flow and Tobin's q on investment at different quantiles. By using the panel quantile regression with individual FEs, they show that the slope estimates change across τ. However, they do not allow the slope coefficients, Θ_1 and Θ_2, to change either over i or t. Inspired by Galvao and Wang (2015), we estimate the following

Table 4: Frequency of correct rank estimation via the SVT approach

DGP	N	T	$\tau=0.25$ $\hat{K}_0$	$\hat{K}_1$	$\hat{K}_2$	$\tau=0.50$ $\hat{K}_0$	$\hat{K}_1$	$\hat{K}_2$	$\tau=0.75$ $\hat{K}_0$	$\hat{K}_1$	$\hat{K}_2$
1	75	35	1.00	0.996	0.996	1.00	0.999	1.00	1.00	0.999	0.999
		70	1.00	0.994	0.996	1.00	1.00	1.00	1.00	1.00	1.00
	150	35	1.00	0.994	0.995	1.00	1.00	0.999	1.00	1.00	0.999
		70	1.00	0.995	0.996	1.00	0.999	1.00	1.00	0.999	1.00
2	75	35	1.00	0.993	0.999	1.00	0.999	1.00	1.00	1.00	1.00
		70	1.00	0.997	0.998	1.00	1.00	1.00	1.00	1.00	1.00
	150	35	1.00	0.995	0.996	1.00	0.998	1.00	1.00	1.00	1.00
		70	1.00	1.00	1.00	1.00	1.00	0.998	1.00	1.00	1.00
3	75	35	1.00	0.990	0.997	1.00	0.997	0.997	1.00	0.999	0.999
		70	1.00	0.994	0.994	1.00	0.999	0.999	1.00	0.999	0.998
	150	35	1.00	0.999	0.992	1.00	1.00	1.00	1.00	1.00	1.00
		70	1.00	0.996	0.994	1.00	0.998	1.00	1.00	1.00	1.00
4	75	35	1.00	0.992	0.991	1.00	0.999	0.999	1.00	0.999	1.00
		70	1.00	0.995	0.995	1.00	0.999	0.999	1.00	1.00	1.00
	150	35	1.00	0.996	0.997	1.00	0.999	1.00	1.00	1.00	1.00
		70	1.00	0.997	0.999	1.00	0.999	1.00	1.00	1.00	1.00
5	75	35	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	150	35	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
6	75	35	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	150	35	1.00	1.00	0.999	1.00	1.00	1.00	1.00	1.00	1.00
		70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

model

$$\mathscr{Q}_\tau\left(IK_{it}\middle|\left\{CFK_{it},q_{i,t-1}\right\}_{t\in[T]},\left\{\Theta_{j,it}\right\}_{t\in[T],j\in\{0,1,2\}}\right)=\Theta_{0,it}(\tau)+\Theta_{1,it}(\tau)CFK_{it}+\Theta_{2,it}(\tau)q_{i,t-1}, \tag{6.1}$$

where $IK_{it}=\frac{I_{it}}{K_{i,t-1}}$, and $CFK_{it}=\frac{CF_{it}}{K_{i,t-1}}$. Here we don't restrict the specific structure on the FEs and they can be either additive or interactive.

The data are taken from the China Stock Market & Accounting Research (CSMAR) Database. We use quarterly data for 195 manufacturing firms in China from 2003 to 2020. Based on the model (6.1), we define corporate investment as $I_{it}=LI_{it}-LI_{i,t-1}$, where LI_{it} is the total value of long-term corporate investment as the sum of long-term equity investment, long-term bound investment, fixed assets and immaterial assets. The investment measures the change of firm's total investment compared to the last period. All these four variables can be easily obtained from the balance sheet. We directly use Tobin's q from the CSMAR database, where by definition $q=\frac{MV}{K}$ and MV is the market value of the firm. We obtain a balanced panel dataset with 195 firms and 72 time periods. The units of corporate investment, capital and cash flow are measured by billions of Chinese RMB.

By using the SVT approach, we obtain the estimates of the ranks of Θ_1 and Θ_2 : $\hat{r}_1=\hat{r}_2=1$ for each $\tau=\{0.25,0.5,0.75\}$. Consequently, we can consider the test that whether $\Theta_{j,it}$ is constant over i or constant over t for both $j=1,2$. Specifically, we want to test whether the effect of

Table 5: Empirical size and power of testing slope homogeneity across either i or t (nominal level: 0.05)

DGP	N	T	$\tau=0.25$				$\tau=0.5$				$\tau=0.75$			
			u_1	v_1	u_2	v_2	u_1	v_1	u_2	v_2	u_1	v_1	u_2	v_2
DGP 1	75	35	0.040	0.051	0.049	0.032	0.024	0.054	0.034	0.054	0.036	0.047	0.036	0.048
		70	0.040	0.055	0.050	0.044	0.020	0.056	0.017	0.068	0.025	0.037	0.029	0.029
	150	35	0.028	0.036	0.058	0.048	0.065	0.054	0.052	0.055	0.074	0.030	0.076	0.024
		70	0.034	0.025	0.030	0.023	0.035	0.048	0.028	0.040	0.035	0.025	0.039	0.025
DGP 2	75	35	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	150	35	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DGP 3	75	35	0.045	0.057	0.050	0.041	0.022	0.050	0.038	0.089	0.054	0.047	0.049	0.047
		70	0.048	0.031	0.031	0.033	0.028	0.086	0.023	0.069	0.041	0.046	0.032	0.038
	150	35	0.065	0.054	0.058	0.034	0.064	0.051	0.068	0.045	0.084	0.018	0.089	0.023
		70	0.046	0.030	0.044	0.025	0.022	0.037	0.037	0.030	0.046	0.022	0.048	0.015
DGP 4	75	35	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	150	35	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
		70	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

cash flow and Tobin's q on the firm's investment is homogeneous over i or across t with market imperfection. That is, for $j \in \{1,2\}$, we shall test

- H_0^a: $\Theta_{j,it}$ is a constant over i,
- H_0^b: $\Theta_{j,it}$ is a constant over t.

Figure 1 shows the estimation results for the factor and factor loadings of two slope coefficient matrices under different quantiles. In each sub-figure, the first and second rows report the results for Θ_1 and Θ_2, respectively. Specifically, the first row of Figure 1(a) gives the plot of $\{\hat{u}_{i,1}\}_{i\in[N]}$ as a catenation of $\{\hat{u}_{i,1}^{(1,2)}, \hat{u}_{i,1}^{(2,3)}, \hat{u}_{i,1}^{(3,1)}\}$ at the left and as a cantenation of $\{\hat{u}_{i,1}^{(1,3)}, \hat{u}_{i,1}^{(2,1)}, \hat{u}_{i,1}^{(3,2)}\}$ at the right in the first row, and similarly the plot of $\{\hat{u}_{i,2}\}_{i\in[N]}$ in the second row. Similarly, the first row of Figure 1(d) shows $\{\hat{v}_{t,1}^{(a,b)}\}_{t\in[T]}$ for $a \in [3]$, $b \in [3]\setminus\{a\}$ in the first row and $\{\hat{v}_{t,3}^{(a,b)}\}_{t\in[T]}$ for $a \in [3]$, $b \in [3]\setminus\{a\}$ in the second row.

Table 7 reports the test statistics, critical values, and p-values. Tobin's q can measure a firm's investment demand. After controlling the Tobin's q and the intercept FEs, the coefficient of cash flow captures a firm's potential for external investment with the variation of internal finance. It is clear that we can reject the homogeneous hypotheses for both i and t at the 1% significance level for each $\tau \in \{0.25, 0.5, 0.75\}$. This indicates that with high probability, the slope coefficient of both CFK and Tobin's q follow the factor structure with one factor.

Table 6: Empirical size and power for testing additive slopes (nomial level: 0.05)

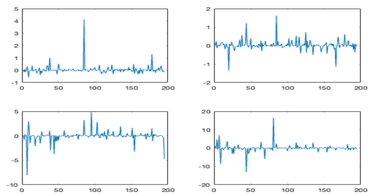
DGP	N	T	$\tau = 0.25$ size	$\tau = 0.25$ power	$\tau = 0.50$ size	$\tau = 0.50$ power	$\tau = 0.75$ size	$\tau = 0.75$ power
DGP 5	75	35	0.027	0.807	0.034	1.00	0.026	0.979
		70	0.065	1.00	0.061	1.00	0.034	1.00
	150	35	0.018	1.00	0.026	1.00	0.011	1.00
		70	0.012	1.00	0.029	1.00	0.010	1.00
DGP 6	75	35	0.039	0.796	0.058	1.00	0.045	0.979
		70	0.024	1.00	0.06	1.00	0.032	1.00
	150	35	0.019	1.00	0.025	1.00	0.022	1.00
		70	0.01	1.00	0.022	1.00	0.014	1.00

The above study shows strong evidence that under imperfect market, the sensitivity of corporate investment to cash flow exhibits both individual heterogeneity and time heterogeneity across quantiles. It implies that neither the usual homogenous panel QR model nor the panel QR model with either cross-section or time heterogeneity alone in the slope coefficients fails to fully capture the unobserved heterogeneity in the investment equation.

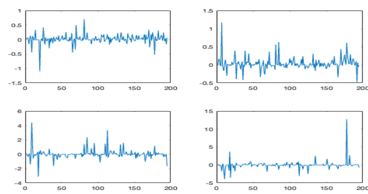
Table 7: Test results under different quantiles for the investment equation

τ	Test	S	$cv_{\alpha=0.01}$	$cv_{\alpha=0.05}$	$cv_{\alpha=0.1}$	p-value
0.25	u_{CFK}	1.28×10^3	16.94	13.68	12.24	0.00
	u_q	4.16×10^4				0.00
	v_{CFK}	13.85	5.70	4.07	3.35	0.00
	v_q	870.85				0.00
0.50	u_{CFK}	148.28	16.94	13.68	12.24	0.00
	u_q	1.24×10^5				0.00
	v_{CFK}	49.57	5.70	4.07	3.35	0.00
	v_q	138.83				0.00
0.75	u_{CFK}	313.21	16.94	13.68	12.24	0.00
	u_q	2.03×10^4				0.00
	v_{CFK}	31.50	5.70	4.07	3.35	0.00
	v_q	58.29				0.00

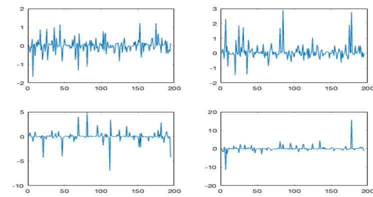
Notes: S is the test statistics for the factor or factor loadings under different quantiles, $H_0^a(CFK)$ and $H_0^a(q)$ refer to the hypotheses that the slope of CFK and Tobin'q is homogeneous across i, respectively. $H_0^b(CFK)$ and $H_0^b(q)$ refer to the the hypotheses that the slope of CFK and Tobin'q is homogeneous across t, respectively. $\text{cv}_{\alpha=a}$ is the critical value under the significance level a where a=0.1, 0.05, and 0.01.



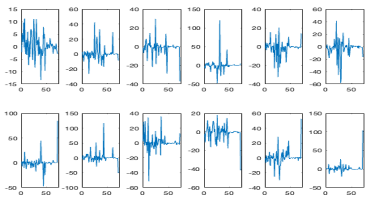
(a) factor loading estimates under $\tau = 0.25$



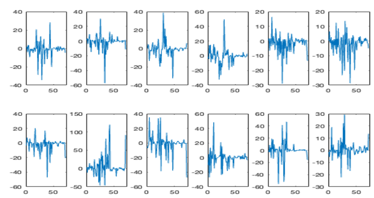
(b) factor loading estimates under $\tau = 0.5$



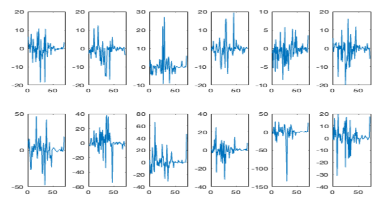
(c) factor loading estimates under $\tau = 0.75$



(d) factor estimates under $\tau = 0.25$



(e) factor estimates under $\tau = 0.5$



(f) factor estimates under $\tau = 0.75$

Figure 1: Factor loading and factor estimates under different quantiles

6.2 Foreign Direct Investment and Unemployment

Investment is one of the major driving forces for economic growth and employment. Among the investment, foreign direct investment (FDI) is an important contributor to the employment. See Craigwell (2006), Aktar et al. (2009), Karlsson et al. (2009), Mucuk and Demirsel (2013), and Strat et al. (2015), among others. Controversially, Mucuk and Demirsel (2013) argue that FDI may have both positive and negative effects on employment. On the one hand, FDI adds to the net capital and creates jobs through forward and backward linkages and multiplier effects in local economy. On the other hand, acquisitions may rely on imports or displacement of existing firms which may result in job loss.

To study the relationship of FDI, economic growth rate and unemployment at the country level, we consider the following panel quantile regression model,

$$\mathscr{Q}_\tau \left(U_{it} \middle| \{G_{i,t-1}, FDI_{it}\}_{t\in[T]}, \{\Theta_{j,it}\}_{t\in[T],j\in\{0,1,2\}} \right) = \Theta_{0,it}(\tau) + \Theta_{1,it}(\tau)G_{i,t-1} + \Theta_{2,it}(\tau)FDI_{it},$$

where U_{it} is the unemployment rate of country i at year t, $G_{i,t-1}$ is the economic growth measured by the growth of real GDP. $\Theta_{0,it}$ is the FEs of country i and year t, $\Theta_{1,it}$ is the elasticity of the economic growth in the previous year to the unemployment this year, and $\Theta_{2,it}$ is the elasticity of FDI to the unemployment.

We draw the data for 126 countries from 1992-2019. The data for the unemployment rate are taken from International Labor Organization (ILO) and GDP growth and FDI are from the World Bank Development Indicators (WDI) historical database. The rank estimation procedure shows that $\hat{r}_1 = 2$ and $\hat{r}_2 = 1$. Consequently, we can test whether the elasticity of FDI to the unemployment rate is homogeneous across individual countries and over years 1992-2009, and whether the elasticity of growth rate to unemployment follows the additive structure, i.e.,

- H_0^c: $\Theta_{1,it} = \Theta_{1,i} + \Theta_{1,t}$,
- H_0^d: $\Theta_{2,it}$ is a constant over i,
- H_0^e: $\Theta_{2,it}$ is a constant over t.

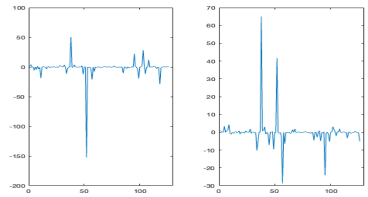
Table 8 reports the test results under quantiles 0.25, 0.5 and 0.75 for the above three null hypotheses. Figure 2 gives the estimation results for the factor and factor loading estimates of the slope coefficient Θ_2. As Table 8 suggests, we can reject all the above three null hypotheses safely at the conventional 5% significance level. This means that the effect of FDI on the unemployment rate is different across both countries and time even though the estimated rank of Θ_2 is one, and the effect of economic growth rate on the unemployment is heterogeneous across both countries and time and it does not exhibit an additive structure.

Table 8: Test results under different quantiles

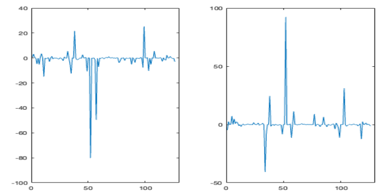
Test	τ	S	$cv_{\alpha=0.01}$	$cv_{\alpha=0.05}$	$cv_{\alpha=0.10}$	$p-$value
H_0^c	0.25	38.92	35.55	32.29	30.85	0.00
	0.50	80.84	22.29	19.03	17.59	0.00
	0.75	66.24	35.55	32.29	30.85	0.00
H_0^d	0.25	1.41×10^6	16.15	12.89	11.45	0.00
	0.50	6.39×10^6				0.00
	0.75	3.07×10^7				0.00
H_0^e	0.25	36.36	5.70	4.07	3.35	0.00
	0.50	164.03				0.00
	0.75	5.44				0.013

7 Conclusion

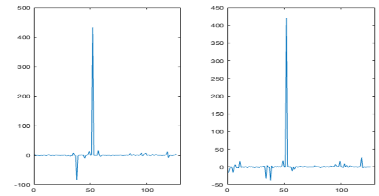
This paper considers panel QR model with heterogeneous slopes over both i and t. Compared to Chernozhukov et al. (2019), to remove the bias from the nuclear norm regularization, we split the full sample into three subsamples. We then use the first subsample to compute initial estimators via NNR, the second sample to refine the convergence rate of the initial estimator, and the last subsample to debias the refined estimator. Our asymptotic theory shows that the factor estimates, factor loading estimates and the slope estimates all follow the normal distributions asymptotically. By constructing the consistent estimator for the asymptotic variance, we also conduct two specification tests: (1) the slope coefficient is constant over time or individuals under the case that true rank of slope matrix equals one and (2) the slope coefficient exhibits the additive structure under the case that true rank of the slope coefficient matrix equals two. Our test statistics



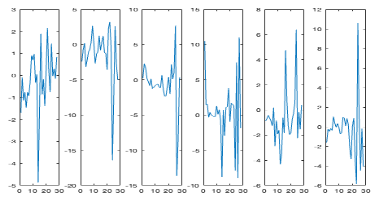
(a) factor loading estimates under $\tau = 0.25$



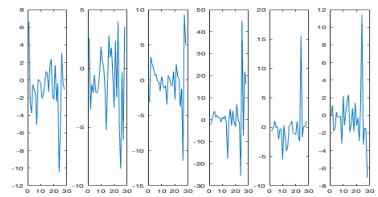
(b) factor loading estimates under $\tau = 0.5$



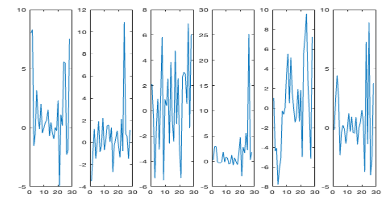
(c) factor loading estimates under $\tau = 0.75$



(d) factor estimates under $\tau = 0.25$



(e) factor estimates under $\tau = 0.5$



(f) factor estimates under $\tau = 0.75$

Figure 2: Factor loading and factor estimates of Θ_2 under different quantiles

are shown to follow the Gumbel distribution asymptotically under the null, consistent under the global alternative and have non-trivial power against local alternatives. Monte Carlo simulation and empirical studies illustrate the finite sample performance of our algorithm and test statistics.

References

Aktar, İ., Demirci, N., and Öztürk, L. (2009). Can unemployment be cured by economic growth and foreign direct investment? *Sosyal Ekonomik Araştırmalar Dergisi*, 9(17):452–467.

Alidaee, H., Auerbach, E., and Leung, M. P. (2020). Recovering network structure from aggregated relational data using penalized regression. *arXiv preprint arXiv:2001.06052*.

Almeida, H. and Campello, M. (2007). Financial constraints, asset tangibility, and corporate investment. *The Review of Financial Studies*, 20(5):1429–1460.

Ando, T. and Bai, J. (2016). Panel data models with grouped factor structure under unknown group membership. *Journal of Applied Econometrics*, 31(1):163–191.

Athey, S., Bayati, M., Doudchenko, N., Imbens, G., and Khosravi, K. (2021). Matrix completion methods for causal panel data models. *Journal of the American Statistical Association*, pages 1–15.

Bai, J. (2003). Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171.

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

Bai, J. and Ng, S. (2019). Rank regularized estimation of approximate factor models. *Journal of Econometrics*, 212(1):78–96.

Bai, J. and Ng, S. (2020). Simpler proofs for approximate factor models of large dimensions. *arXiv preprint arXiv:2008.00254*.

Baltagi, B. H., Kao, C., and Wang, F. (2021). Estimating and testing high dimensional factor models with multiple structural changes. *Journal of Econometrics*, 220(2):349–365.

Belloni, A., Chen, M., Padilla, O. H. M., and Wang, Z. (2022). High dimensional latent panel quantile regression with an application to asset pricing. *Annals of statistics*, forthcoming.

Bonhomme, S. and Manresa, E. (2015). Grouped patterns of heterogeneity in panel data. *Econometrica*, 83(3):1147–1184.

Castagnetti, C., Rossi, E., and Trapani, L. (2015). Inference on factor structures in heterogeneous panels. *Journal of Econometrics*, 184(1):145–157.

Chen, L. (2015). Estimating the common break date in large factor models. *Economics Letters*, 131:70–74.

Chen, L. (2022). Two-step estimation of quantile panel data models with interactive fixed effects. *Econometric Theory*, forthcoming.

Chen, L., Dolado, J. J., and Gonzalo, J. (2021). Quantile factor models. *Econometrica*, 89(2):875–910.

Chen, Y., Chi, Y., Fan, J., Ma, C., and Yan, Y. (2020). Noisy matrix completion: Understanding statistical guarantees for convex relaxation via nonconvex optimization. *SIAM journal on optimization*, 30(4):3098–3121.

Cheng, X., Liao, Z., and Schorfheide, F. (2016). Shrinkage estimation of high-dimensional factor models with structural instabilities. *The Review of Economic Studies*, 83(4):1511–1543.

Chernozhukov, V., Hansen, C. B., Liao, Y., and Zhu, Y. (2019). Inference for heterogeneous effects using low-rank estimations. Technical report, CEMMAP working paper.

Cleary, S. (1999). The relationship between firm investment and financial status. *The Journal of Finance*, 54(2):673–692.

Craigwell, R. (2006). Foreign direct investment and employment in the english and dutch-speaking caribbean. *ILO Subregional Office for the Caribbean.*

Devereux, M. and Schiantarelli, F. (1990). Investment, financial factors, and cash row: Evidence from uk panel data. *Asymmetric Information, Corporate Finance, and Investment*, pages 279–306.

Fazzari, S., Hubbard, R. G., and Petersen, B. C. (1988). Financing constraints and corporate investment. *Brookings Papers on Economic Activity*, (1):141–206.

Feng, J. (2019). Regularized quantile regression with interactive fixed effects. *arXiv preprint arXiv:1911.00166*.

Galvao, A. F., Gu, J., and Volgushev, S. (2020). On the unbiased asymptotic normality of quantile regression with fixed effects. *Journal of Econometrics*, 218(1):178–215.

Galvao, A. F., Juhl, T., Montes-Rojas, G., and Olmo, J. (2018). Testing slope homogeneity in quantile regression panel data with an application to the cross-section of stock returns. *Journal of Financial Econometrics*, 16(2):211–243.

Galvao, A. F. and Kato, K. (2016). Smoothed quantile regression for panel data. *Journal of Econometrics*, 193(1):92–112.

Galvao, A. F. and Montes-Rojas, G. V. (2010). Penalized quantile regression for dynamic panel data. *Journal of Statistical Planning and Inference*, 140(11):3476–3497.

Galvao, A. F. and Wang, L. (2015). Efficient minimum distance estimator for quantile regression fixed effects panel data. *Journal of Multivariate Analysis*, 133:1–26.

Gilchrist, S. and Himmelberg, C. P. (1995). Evidence on the role of cash flow for investment. *Journal of Monetary Economics*, 36(3):541–572.

Golub, G. H. and Van Loan, C. F. (1996). *Matrix Computations*. JHU Press.

Hong, S., Su, L., and Jiang, T. (2022). Profile gmm estimation of panel data models with interactive fixed effects. *Journal of Econometrics*, forthcoming.

Kaplan, S. N. and Zingales, L. (1995). Do financing constraints explain why investment is correlated with cash flow? *NBER Working paper*, (No. 5267).

Karlsson, S., Lundin, N., Sjöholm, F., and He, P. (2009). Foreign firms and chinese employment. *World Economy*, 32(1):178–201.

Kato, K., Galvao Jr, A. F., and Montes-Rojas, G. V. (2012). Asymptotics for panel quantile regression models with individual effects. *Journal of Econometrics*, 170(1):76–91.

Knight, K. (1998). Limiting distributions for l_1 regression estimators under general conditions. *Annals of Statistics*, pages 755–770.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91(1):74–89.

Leadbetter, M. and Rootzen, H. (1988). Extremal theory for stochastic processes. *The Annals of Probability*, pages 431–478.

Ledoux, M. and Talagrand, M. (1989). Comparison theorems, random geometry and some limit theorems for empirical processes. *The Annals of Probability*, pages 596–631.

Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces: isoperimetry and processes*, volume 23. Springer Science & Business Media.

Lin, C.-C. and Ng, S. (2012). Estimation of panel data models with parameter heterogeneity when group membership is unknown. *Journal of Econometric Methods*, 1(1):42–55.

Lu, X. and Su, L. (2021). Uniform inference in linear panel data models with two-dimensional heterogeneity. *Journal of Econometrics*, forthcoming.

Lumsdaine, R. L., Okui, R., and Wang, W. (2021). Estimation of panel group structure models with structural breaks in group memberships and coefficients. *Journal of Econometrics*, forthcoming.

Ma, S. and Su, L. (2018). Estimation of large dimensional factor models with an unknown number of breaks. *Journal of Econometrics*, 207(1):1–29.

Ma, S., Su, L., and Zhang, Y. (2020). Detecting latent communities in network formation models. *arXiv preprint arXiv:2005.03226*.

Machado, J. A. and Silva, J. S. (2019). Quantiles via moments. *Journal of Econometrics*, 213(1):145–173.

Merlevède, F., Peligrad, M., and Rio, E. (2009). Bernstein inequality and moderate deviations under strong mixing conditions. In *High dimensional probability V: the Luminy volume*, pages 273–292. Institute of Mathematical Statistics.

Miao, K., Phillips, P. C., and Su, L. (2022). High-dimensional vars with common factors. *Journal of Econometrics*, forthcoming.

Moon, H. R. and Weidner, M. (2018). Nuclear norm regularized estimation of panel regression models. *arXiv preprint arXiv:1810.10987*.

Mucuk, M. and Demirsel, M. T. (2013). The effect of foreign direct investments on unemployment: Evidence from panel data for seven developing countries. *Journal of Business Economics and Finance*, 2(3):53–66.

Okui, R. and Wang, W. (2021). Heterogeneous structural breaks in panel data models. *Journal of Econometrics*, 220(2):447–473.

Prakasa Rao, B. L. (2009). Conditional independence, conditional mixing and conditional association. *Annals of the Institute of Statistical Mathematics*, 61(2):441–460.

Rakhlin, A., Sridharan, K., and Tewari, A. (2015). Sequential complexities and uniform martingale laws of large numbers. *Probability Theory and Related Fields*, 161(1-2):111–153.

Rauh, J. D. (2006). Investment and financing constraints: Evidence from the funding of corporate pension plans. *The Journal of Finance*, 61(1):33–71.

Strat, V. A., Davidescu, A., and Paul, A. M. (2015). Fdi and the unemployment-a causality analysis for the latest eu members. *Procedia Economics and Finance*, 23:635–643.

Su, L. and Chen, Q. (2013). Testing homogeneity in panel data models with interactive fixed effects. *Econometric Theory*, 29(6):1079–1135.

Su, L. and Ju, G. (2018). Identifying latent grouped patterns in panel data models with interactive fixed effects. *Journal of Econometrics*, 206(2):554–573.

Su, L., Shi, Z., and Phillips, P. C. (2016). Identifying latent structures in panel data. *Econometrica*, 84(6):2215–2264.

Su, L., Wang, W., and Zhang, Y. (2020). Strong consistency of spectral clustering for stochastic block models. *IEEE Transactions on Information Theory*, 66(1):324–338.

Su, L., Wang, X., and Jin, S. (2019). Sieve estimation of time-varying panel data models with latent structures. *Journal of Business & Economic Statistics*, 37(2):334–349.

Tropp, J. A. (2011). User-friendly tail bounds for matrix martingales. Technical report, California Institute of Technology.

Vershynin, R. (2010). Introduction to the non-asymptotic analysis of random matrices. *arXiv preprint arXiv:1011.3027*.

Wainwright, M. J. (2019). *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press.

Wang, W., Phillips, P. C., and Su, L. (2018). Homogeneity pursuit in panel data models: Theory and application. *Journal of Applied Econometrics*, 33(6):797–815.

Wang, W., Phillips, P. C., and Su, L. (2019). The heterogeneous effects of the minimum wage on employment across states. *Economics Letters*, 174:179–185.

Wang, W. and Su, L. (2021). Identifying latent group structures in nonlinear panels. *Journal of Econometrics*, 220(2):272–295.

Yu, Y., Wang, T., and Samworth, R. J. (2015). A useful variant of the davis–kahan theorem for statisticians. *Biometrika*, 102(2):315–323.

Zhang, Y., Wang, H. J., and Zhu, Z. (2019). Quantile-regression-based clustering for panel data. *Journal of Econometrics*, 213(1):54–67.

Online Supplement for
"Low-rank Panel Quantile Regression: Estimation and Inference"

Yiren Wang[a], Liangjun Su[b] and Yichong Zhang[a]
[a]School of Economics, Singapore Management University, Singapore
[b]School of Economics and Management, Tsinghua University, China

This supplement contains three sections. Section A contains the proofs of the main results by calling upon some technical lemmas in Section B. Section B states and proves the technical lemmas used in Section A. Section C provides detail algorithm for the nuclear norm regularized panel quantile regression.

A Proofs of the Main Results

A.1 Proof of Theorem 3.1

We focus on the full sample estimators $\tilde{\Delta}_{\Theta_j}$, $\tilde{\sigma}_{k,j}$, and $\tilde{V}_j$ in the proof. The results for their subsample counterparts can be established in the same manner, and we omit the detail for brevity.

A.1.1 Proof of Statement (i)

Recall that

$$\mathcal{R}(C_1, C_2) := \left\{ (\{\Delta_{\Theta_j}\}_{j=0}^p) : \sum_{j=0}^p \left\|\mathcal{P}_j^{\perp}(\Delta_{\Theta_j})\right\|_* \leq C_1 \sum_{j=0}^p \left\|\mathcal{P}_j(\Delta_{\Theta_j})\right\|_*, \ \sum_{j=0}^p \left\|\Delta_{\Theta_j}\right\|_F^2 \geq C_2\sqrt{NT} \right\}.$$

Define $\mathcal{R}(C_1) := \left\{ \{\Delta_{\Theta_j}\}_{j=0}^p : \sum_{j=0}^p \left\|\mathcal{P}_j^{\perp}(\Delta_{\Theta_j})\right\|_* \leq C_1 \sum_{j=0}^p \left\|\mathcal{P}_j(\Delta_{\Theta_j})\right\|_* \right\}$. By Lemma B.4, $\mathbb{P}\{\{\tilde{\Delta}_{\Theta_j}(\tau)\}_{j=0}^p \in \mathcal{R}(3)\} \to 1$. When $\{\tilde{\Delta}_{\Theta_j}\}_{j=0}^p \in \mathcal{R}(C_1)$ and $\{\tilde{\Delta}_{\Theta_j}\}_{j=0}^p \notin \mathcal{R}(3, C_2)$, we have $\sum_{j=0}^p \left\|\tilde{\Delta}_{\Theta_j}\right\|_F^2 < C_2\sqrt{NT}$, which implies $\frac{1}{\sqrt{NT}}\left\|\tilde{\Delta}_{\Theta_j}\right\|_F = O_p\left((N \wedge T)^{-1/2}\right), \quad \forall j \in [p] \cup \{0\}$. It suffices to consider the case that $\{\tilde{\Delta}_{\Theta_j}\}_{j=0}^p \in \mathcal{R}(3, C_2)$.

Define

$$\begin{aligned}\mathbb{Q}_\tau\left(\{\Theta_j\}_{j=0}^p\right) &= \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \rho_\tau\left(Y_{it} - \Theta_{0,it} - \sum_{j=1}^p X_{j,it}\Theta_{j,it}\right), \text{ and} \\ Q_\tau(\{\Theta_j\}_{j=0}^p) &= \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \mathbb{E}\left[\rho_\tau\left(Y_{it} - \Theta_{0,it} - \sum_{j=1}^p X_{j,it}\Theta_{j,it}\right)\middle|\mathscr{G}_{i,t-1}\right],\end{aligned}$$

where $\mathscr{G}_{i,t-1}$ is defined in Assumption 1. Then we have

$$0 \geq \mathbb{Q}_\tau\left(\left\{\Theta_j^0 + \tilde{\Delta}_{\Theta_j}\right\}_{j=0}^p\right) - \mathbb{Q}_\tau\left(\{\Theta_j^0\}_{j=0}^p\right) + \sum_{j=0}^p \nu_j\left(\left\|\Theta_j^0 + \tilde{\Delta}_{\Theta_j}\right\|_* - \|\Theta_j^0\|_*\right)$$

$$= \left\{ \mathbb{Q}_\tau \left(\left\{ \Theta_j^0 + \tilde{\Delta}_{\Theta_j} \right\}_{j=0}^p \right) - \mathbb{Q}_\tau \left(\left\{ \Theta_j^0 \right\}_{j=0}^p \right) - \left[Q_\tau \left(\left\{ \Theta_j^0 + \tilde{\Delta}_{\Theta_j} \right\}_{j=0}^p \right) - Q_\tau \left(\left\{ \Theta_j^0 \right\}_{j=0}^p \right) \right] \right\}$$
$$+ \left[Q_\tau \left(\left\{ \Theta_j^0 + \tilde{\Delta}_{\Theta_j} \right\}_{j=0}^p \right) - Q_\tau \left(\left\{ \Theta_j^0 \right\}_{j=0}^p \right) \right] + \sum_{j=0}^p \nu_j \left(\left\| \Theta_j^0 + \tilde{\Delta}_{\Theta_j} \right\|_* - \left\| \Theta_j^0 \right\|_* \right), \tag{A.1}$$

where the first inequality holds by the definition of the estimator. Noted that

$$\nu_j \left| \left\| \Theta_j^0 + \tilde{\Delta}_{\Theta_j} \right\|_* - \left\| \Theta_j^0 \right\|_* \right| \leq \nu_j \left\| \tilde{\Delta}_{\Theta_j} \right\|_* \leq c_8 \nu_j \sum_{j=0}^p \left\| \tilde{\Delta}_{\Theta_j} \right\|_F \tag{A.2}$$

where the first inequality is due to triangle inequality and the second inequality holds by Lemma B.7 with positive constant c_8 defined in the lemma.

Define

$$\rho_{it} \left(\left\{ \Delta_{\Theta_j,it}, X_{j,it} \right\}_{j=0}^p, \epsilon_{it} \right) = \rho_\tau \left(\epsilon_{it} - \Delta_{0,it} - \sum_{j=1}^p X_{j,it} \Delta_{\Theta_j,it} \right) - \rho_\tau \left(\epsilon_{it} \right), \tag{A.3}$$
$$\bar{\rho}_{it} \left(\left\{ \Delta_{\Theta_j,it}, X_{j,it} \right\}_{j=0}^p, \epsilon_{it} \right) = \mathbb{E} \left[\rho_\tau \left(\epsilon_{it} - \Delta_{0,it} - \sum_{j=1}^p X_{j,it} \Delta_{\Theta_j,it} \right) - \rho_\tau \left(\epsilon_{it} \right) \middle| \mathscr{G}_{i,t-1} \right],$$
$$\tilde{\rho}_{it} \left(\left\{ \Delta_{\Theta_j,it}, X_{j,it} \right\}_{j=0}^p, \epsilon_{it} \right) = \rho_{it} \left(\left\{ \Delta_{\Theta_j,it}, X_{j,it} \right\}_{j=0}^p, \epsilon_{it} \right) - \bar{\rho}_{it} \left(\left\{ \Delta_{\Theta_j,it}, X_{j,it} \right\}_{j=0}^p, \epsilon_{it} \right), \tag{A.4}$$
$$\mathscr{A}_1 = \left\{ \sup_{\left\{ \Delta_{\Theta_j} \right\}_{j=0}^p \in \mathcal{R}(3, C_2)} \frac{\left| \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \tilde{\rho}_{it} \left(\left\{ \Delta_{\Theta_j,it}, X_{j,it} \right\}_{j=0}^p, \epsilon_{it} \right) \right|}{\sum_{j=0}^p \left\| \Delta_{\Theta_j} \right\|_F} \leq C_5 a_{NT} \right\},$$

with $a_{NT} = \frac{\sqrt{(N \vee T) \log(N \vee T)}}{NT}$ for some positive constant C_5, and $\mathscr{A}_1^c$ as the complement of $\mathscr{A}_1$.

On $\mathscr{A}_1$, following (A.1), we have w.p.a.1

$$0 \geq \left[Q_\tau \left(\left\{ \Theta_j^0 + \tilde{\Delta}_{\Theta_j} \right\}_{j=0}^p \right) - Q_\tau \left(\left\{ \Theta_j^0 \right\}_{j=0}^p \right) \right]$$
$$- \left| \mathbb{Q}_\tau \left(\left\{ \Theta_j^0 + \tilde{\Delta}_{\Theta_j} \right\}_{j=0}^p \right) - \mathbb{Q}_\tau \left(\left\{ \Theta_j^0 \right\}_{j=0}^p \right) - \left[Q_\tau \left(\left\{ \Theta_j^0 + \tilde{\Delta}_{\Theta_j} \right\}_{j=0}^p \right) - Q_\tau \left(\left\{ \Theta_j^0 \right\}_{j=0}^p \right) \right] \right|$$
$$- \sum_{j=0}^p \nu_j \left| \left\| \Theta_j^0 + \tilde{\Delta}_{\Theta_j} \right\|_* - \left\| \Theta_j^0 \right\|_* \right|$$
$$\geq \frac{c_7 C_3}{NT \xi_N^2} \sum_{j=0}^p \left\| \tilde{\Delta}_{\Theta_j} \right\|_F^2 - \frac{c_7 C_4}{NT \xi_N^2} (N + T) - \frac{C_5 \sum_{j=0}^p \left\| \tilde{\Delta}_{\Theta_j} \right\|_F \sqrt{(N \vee T) \log (N \vee T)}}{NT}$$
$$- c_8 \sum_{j=0}^p \nu_j \sum_{j=0}^p \left\| \tilde{\Delta}_{\Theta_j} \right\|_F,$$

where the first inequality is by triangle inequality, the second inequality holds by (A.2) and Lemmas B.6 and B.11. It follows that

$$\frac{c_7 C_3}{NT \xi_N^2} \sum_{j=0}^p \left\| \tilde{\Delta}_{\Theta_j} \right\|_F^2 - \frac{c_8 (p+1) c_0 \sqrt{N \vee (T \log T)} + C_5 \sqrt{(N \vee T) \log (N \vee T)}}{NT} \sum_{j=0}^p \left\| \tilde{\Delta}_{\Theta_j} \right\|_F$$
$$- \frac{c_7 C_4}{NT \xi_N^2} (N + T) \leq 0,$$

which implies

$$\frac{1}{\sqrt{NT}}\left\|\tilde{\Delta}_{\Theta_j}\right\|_F = O\left(\frac{\sqrt{\log(N\vee T)}\xi_N^2}{\sqrt{N\wedge T}}\right)$$

under the event $\mathscr{A}_1$.

By Lemma B.11, for any $\delta > 0$, we can choose a sufficiently large C_5 such that $\mathbb{P}\{\mathscr{A}_1^c\} \leq \delta$. This implies

$$\frac{1}{\sqrt{NT}}\left\|\tilde{\Theta}_j - \Theta_j^0\right\|_F = O_p\left(\frac{\sqrt{\log(N\vee T)}\xi_N^2}{\sqrt{N\wedge T}}\right), \quad \forall j \in \{0,\cdots,p\}. \quad \blacksquare$$

A.1.2 Proof of Statement (ii)

With the statement (i), the second statement holds by the Weyl's inequality. $\blacksquare$

A.1.3 Proof of Statement (iii)

For $\forall j \in \{0,\cdots,p\}$, let $\tilde{D}_j = \frac{1}{NT}\tilde{\Theta}_j'\tilde{\Theta}_j = \hat{\tilde{\mathcal{V}}}_j\hat{\tilde{\Sigma}}_j\hat{\tilde{\mathcal{V}}}_j'$, and recall that $D_j^0 = \frac{1}{NT}\Theta_j^{0\prime}\Theta_j^0 = \mathcal{V}_j^0\Sigma_j^0\mathcal{V}_j^{0\prime}$. Define the event $\mathscr{A}_2(M) = \left\{\frac{1}{\sqrt{NT}}\left\|\tilde{\Theta}_j - \Theta_j^0\right\|_F \leq M\eta_N, \forall j \in \{0,\cdots,p\}\right\}$ with $\eta_N = \frac{\sqrt{\log(N\vee T)}\xi_N^2}{\sqrt{N\wedge T}}$. On event $\mathscr{A}_2(M)$, for some positive constant C_6,

$$\left\|\tilde{D}_j - D_j^0\right\|_F^2 \leq C_6\eta_N.$$

By Lemma C.1 of Su et al. (2020) and Davis-Kahan sinΘ theorem in Yu et al. (2015), there exists an orthogonal rotation matrix O_j such that

$$\left\|\mathcal{V}_j^0 - \hat{\tilde{\mathcal{V}}}_jO_j\right\|_F \leq \sqrt{K_j}\left\|\mathcal{V}_j^0 - \hat{\tilde{\mathcal{V}}}_jO_j\right\|_{op} \leq \sqrt{K_j}\frac{\sqrt{2}C_6\eta_N}{\Sigma_{K_j,1}^2 - C_6\eta_N} \leq \sqrt{K_j}\frac{\sqrt{2}C_6\eta_N}{c_\sigma^2 - C_6\eta_N} \leq \sqrt{K_j}\frac{\sqrt{2}C_6\eta_N}{C_7c_\sigma^2} \leq C_8\eta_N,$$

for $C_8 = \frac{\sqrt{2}C_6\sqrt{\bar{K}}}{C_7c_\sigma^2}$. The second last inequality holds with some positive constant C_7 and the fact that η_N is sufficiently small.

Then $\left\|V_j^0 - \tilde{V}_jO_j\right\|_F \leq C_8\sqrt{T}\eta_N$ by the definition of $\tilde{V}_j$ and V_j. Let $\mathscr{A}_2^c(M)$ be the complement of event $\mathscr{A}_2(M)$. Combining the fact that $\mathbb{P}\{\mathscr{A}_2^c(M)\} \to 0$, it implies $\left\|V_j^0 - \tilde{V}_jO_j\right\|_F = O_p\left(\sqrt{T}\eta_N\right)$. $\blacksquare$

A.2 Proof of Theorem 3.2

A.2.1 Proof of Statement (i)

We prove that $\max_{i\in I_2}\left\|O_j^{(1)\prime}\dot{u}_{i,j}^{(1)} - u_{i,j}^0\right\|_2 = O_p(\eta_N)$ and the $\max_{i\in I_3}\left\|O_j^{(1)\prime}\dot{u}_{i,j}^{(1)} - u_{i,j}^0\right\|_2 = O_p(\eta_N)$ can be derived in the same manner once statement (ii) is satisfied. Define

$$\tilde{\mathbb{Q}}_{\tau i}\left(\{u_{i,j}\}_{j\in[p]\cup\{0\}}\right) = \frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_{it} - u_{i,0}'\tilde{v}_{t,0}^{(1)} - \sum_{j=1}^{p}u_{i,j}'\tilde{v}_{t,j}^{(1)}X_{j,it}\right),$$

$$u_i^0 = \left[u_{i,0}^{0\prime},\cdots,u_{i,p}^{0\prime}\right]', \quad \dot{\Delta}_{i,j} = O_j^{(1)\prime}\dot{u}_{i,j}^{(1)} - u_{i,j}^0, \quad \dot{\Delta}_{i,u} = \left(\dot{\Delta}_{i,0}',\cdots,\dot{\Delta}_{i,p}'\right)',$$

$$\tilde{\Phi}_{it}^{(1)} = \left[\left(O_0^{(1)\prime}\tilde{v}_{t,0}^{(1)}\right)',\left(O_1^{(1)\prime}\tilde{v}_{t,1}^{(1)}X_{1,it}\right)',\cdots,\left(O_p^{(1)\prime}\tilde{v}_{t,p}^{(1)}X_{p,it}\right)'\right]', \quad \tilde{\Phi}_i^{(1)} = \frac{1}{T}\sum_{t=1}^{T}\tilde{\Phi}_{it}^{(1)}\tilde{\Phi}_{it}^{(1)\prime},$$

$$w_{1,it} = Y_{it} - \left(O_0^{(1)} u_{i0}^0\right)' \tilde{v}_{t,0}^{(1)} - \sum_{j=1}^{p} \left(O_j^{(1)} u_{i,j}^0\right)' \tilde{v}_{t,j}^{(1)} X_{j,it} = Y_{it} - u_i^{0\prime} \tilde{\Phi}_{it}^{(1)} = \epsilon_{it} - u_i^{0\prime} (\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^0),$$

and for $i \in I_2$, recall that $\mathscr{D}_{e_i}^{I_1}$ is the σ-field generated by

$$\{\epsilon_{i^*t}, e_{i^*t}\}_{i^* \in I_1, t \in [T]} \bigcup \{e_{it}\}_{t \in [T]} \bigcup \left\{V_j^0\right\}_{j \in [p] \cup \{0\}} \bigcup \left\{W_j^0\right\}_{j \in [p]}.$$

By construction, we have

$$\begin{aligned}
0 \geq\ & \dot{Q}_{\tau i,u}\left(\left\{\dot{u}_{i,j}^{(1)}\right\}_{j \in [p] \cup \{0\}}\right) - \dot{Q}_{\tau i,u}\left(\left\{O_j^{(1)} u_{i,j}^0\right\}_{j \in [p] \cup \{0\}}\right) \\
=\ & \frac{1}{T} \sum_{t=1}^{T} \rho_\tau \left(Y_{it} - \dot{u}_{i,0}^{(1)\prime} O_0^{(1)} O_0^{(1)\prime} \tilde{v}_{t,0}^{(1)} - \sum_{j=1}^{p} \dot{u}_{i,j}^{(1)\prime} O_j^{(1)} O_j^{(1)\prime} \tilde{v}_{t,j}^{(1)} X_{j,it}\right) \\
& - \frac{1}{T} \sum_{t=1}^{T} \rho_\tau \left(Y_{it} - u_{i,0}^{0\prime} O_0^{(1)\prime} \tilde{v}_{t,0}^{(1)} - \sum_{j=1}^{p} u_{i,j}^{0\prime} O_j^{(1)\prime} \tilde{v}_{t,j}^{(1)} X_{j,it}\right) \\
=\ & \frac{1}{T} \sum_{t=1}^{T} \left[\tilde{\Phi}_{it}^{(1)\prime} \dot{\Delta}_{i,u} \left(\tau - \mathbf{1}\{w_{1,it} \leq 0\}\right)\right] + \frac{1}{T} \sum_{t=1}^{T} \int_0^{\tilde{\Phi}_{it}^{(1)\prime} \dot{\Delta}_{i,u}} \left(\mathbf{1}\{w_{1,it} \leq s\} - \mathbf{1}\{w_{1,it} \leq 0\}\right) ds \\
=\ & \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[\tilde{\Phi}_{it}^{(1)\prime} \left(\tau - \mathbf{1}\{w_{1,it} \leq 0\}\right) \middle| \mathscr{D}_{e_i}^{I_1}\right] \dot{\Delta}_{i,u} \\
& + \left\{\frac{1}{T} \sum_{t=1}^{T} \left[\tilde{\Phi}_{it}^{(1)\prime} \left(\tau - \mathbf{1}\{w_{1,it} \leq 0\}\right) - \mathbb{E}\left(\tilde{\Phi}_{it}^{(1)\prime} \left(\tau - \mathbf{1}\{w_{1,it} \leq 0\}\right) \middle| \mathscr{D}_{e_i}^{I_1}\right)\right]\right\} \dot{\Delta}_{i,u} \\
& + \frac{1}{T} \sum_{t=1}^{T} \int_0^{\tilde{\Phi}_{it}^{(1)\prime} \dot{\Delta}_{i,u}} \left(\mathbf{1}\{\epsilon_{it} \leq s\} - \mathbf{1}\{\epsilon_{it} \leq 0\}\right) - \mathbb{E}\left(\mathbf{1}\{\epsilon_{it} \leq s\} - \mathbf{1}\{\epsilon_{it} \leq 0\} \middle| \mathscr{D}_{e_i}^{I_1}\right) ds \\
& + \frac{1}{T} \sum_{t=1}^{T} \int_0^{\tilde{\Phi}_{it}^{(1)\prime} \dot{\Delta}_{i,u}} \mathbb{E}\left(\mathbf{1}\{\epsilon_{it} \leq s\} - \mathbf{1}\{\epsilon_{it} \leq 0\} \middle| \mathscr{D}_{e_i}^{I_1}\right) ds \\
& + \frac{1}{T} \sum_{t=1}^{T} \int_0^{\tilde{\Phi}_{it}^{(1)\prime} \dot{\Delta}_{i,u}} \left(\mathbf{1}\{w_{1,it} \leq s\} - \mathbf{1}\{\epsilon_{it} \leq s\}\right) - \mathbb{E}\left(\mathbf{1}\{w_{1,it} \leq s\} - \mathbf{1}\{\epsilon_{it} \leq s\} \middle| \mathscr{D}_{e_i}^{I_1}\right) ds \\
& + \frac{1}{T} \sum_{t=1}^{T} \int_0^{\tilde{\Phi}_{it}^{(1)\prime} \dot{\Delta}_{i,u}} \mathbb{E}\left(\mathbf{1}\{w_{1,it} \leq s\} - \mathbf{1}\{\epsilon_{it} \leq s\} \middle| \mathscr{D}_{e_i}^{I_1}\right) ds \\
& + \frac{1}{T} \sum_{t \in [T]} \left[\tilde{\Phi}_{it}^{(1)\prime} \dot{\Delta}_{i,u} \left(\mathbf{1}\{\epsilon_{it} \leq 0\} - \mathbf{1}\{w_{1,it} \leq 0\}\right)\right] \\
:=\ & \sum_{m=1}^{7} A_{m,i},
\end{aligned} \tag{A.5}$$

where the first inequality holds by the definition of the estimator and the second equality holds by Knight's identity in Knight (1998) which states that

$$\rho_\tau(u - v) - \rho_\tau(u) = v(\tau - \mathbf{1}\{u \leq 0\}) + \int_0^v \left(\mathbf{1}\{u \leq s\} - \mathbf{1}\{u \leq 0\}\right) ds.$$

After simple manipulation, we have

$$|A_{4,i}| = A_{4,i} \leq -A_{1,i} - A_{2,i} - A_{3,i} - A_{5,i} - A_{6,i} - A_{7,i}$$

$$\leq |A_{1,i}| + |A_{2,i}| + |A_{3,i}| + |A_{5,i}| + |A_{6,i}| + |A_{7,i}| .$$

Define, for some constant M, an event set

$$\mathscr{A}_3(M) = \left\{ \max_{i \in I_2} \left(|A_{m,i}| / \left\| \dot{\Delta}_{i,u} \right\|_2 \right) \leq M\eta_N, \quad m = 1,2,3,5,6,7 \right\}$$

and

$$q_i^I = \inf_{\Delta} \frac{\left[\frac{1}{T} \sum_{t \in [T]} \left(\tilde{\Phi}_{it}^{(1)\prime} \Delta \right)^2 \right]^{\frac{3}{2}}}{\frac{1}{T} \sum_{t \in [T]} \left| \tilde{\Phi}_{it}^{(1)\prime} \Delta \right|^3}. \tag{A.6}$$

Then, under $\mathscr{A}_3(M)$, we have

$$\begin{aligned} M\eta_N \left\| \dot{\Delta}_{i,u} \right\|_2 &\geq |A_{4,i}| \\ &\geq \min \left(\frac{\left(3c_{11}^2 \underline{\mathfrak{f}} - c_{11}^3 \bar{\mathfrak{f}}'\right) c_\phi \max_{i \in I_2} \left\| \dot{\Delta}_{i,u} \right\|_2^2}{12}, \frac{\left(3c_{11}^2 \underline{\mathfrak{f}} - c_{11}^3 \bar{\mathfrak{f}}'\right) \sqrt{c_\phi} q_i^I \max_{i \in I_2} \left\| \dot{\Delta}_{i,u} \right\|_2}{6\sqrt{2}} \right), \end{aligned} \tag{A.7}$$

where $c_{11} < \min(\frac{3\underline{\mathfrak{f}}}{\bar{\mathfrak{f}}'}, 1)$ and the second inequality holds by Lemma B.15. In addition, note that

$$\begin{aligned} \max_{i \in I_2} \frac{1}{T} \sum_{t \in [T]} \left\| \tilde{\Phi}_{it}^{(1)} \right\|_2^3 &= \max_{i \in I_2} \frac{1}{T} \sum_{t \in [T]} \left(\left\| \tilde{v}_{t,0}^{(1)} \right\|_2^2 + \sum_{j \in [p]} \left\| \tilde{v}_{t,j}^{(1)} X_{j,it} \right\|_2^2 \right)^{3/2} \\ &\leq \max_{i \in I_2} \frac{1}{T} \sum_{t \in [T]} \left[\left(\frac{2M}{c_\sigma} \right)^2 \left(1 + \sum_{j \in [p]} X_{j,it}^2 \right) \right]^{3/2} \leq C_9, \quad \text{a.s.} \end{aligned}$$

with C_9 being a positive constant, where the first inequality holds by Lemma B.13(ii) and the second inequality is by Assumption 1(iv). Then we have by Lemma B.14

$$q_i^I \geq \inf_{\Delta} \frac{\|\Delta\|_2^3 \left[\lambda_{\min} \left(\tilde{\Phi}_i^{(1)} \right) \right]^{2/3}}{\|\Delta\|_2^3 \frac{1}{T} \sum_{t \in [T]} \left\| \tilde{\Phi}_{it}^{(1)} \right\|_2^3} \geq \frac{\min_{i \in I_2} \left[\lambda_{\min} \left(\tilde{\Phi}_i^{(1)} \right) \right]^{2/3}}{\max_{i \in I_2} \frac{1}{T} \sum_{t \in [T]} \left\| \tilde{\Phi}_{it}^{(1)} \right\|_2^3} > \frac{(c_\phi/2)^{2/3}}{C_9},$$

which implies

$$\frac{\left(3c_{11}^2 \underline{\mathfrak{f}} - c_{11}^3 \bar{\mathfrak{f}}'\right) \sqrt{c_\phi} q_i^I \max_{i \in I_2} \left\| \dot{\Delta}_{i,u} \right\|_2}{6\sqrt{2}} > C_{10} \max_{i \in I_2} \left\| \dot{\Delta}_{i,u} \right\|_2 \eta_N,$$

as C_{10} is defined to be the positive constant and $\eta_N = o(1)$. Combining this with (A.7), we have

$$\max_{i \in I_2} \left\| O_j^{(1)\prime} \dot{u}_{i,j}^{(1)} - u_{i,j}^0 \right\|_2 \leq \left\| \dot{\Delta}_{i,u} \right\|_2 \leq M' \eta_N$$

for some constant M' which may depends on M. In addition, for an arbitrary constant $e > 0$, we can find a sufficiently large constant M such that $\mathbb{P}(\mathscr{A}_2^c(M)) \leq e$, which implies $\max_{i \in I_2} \left\| O_j^{(1)\prime} \dot{u}_{i,j}^{(1)} - u_{i,j}^0 \right\|_2 = O_p(\eta_N)$. ■

A.2.2 Proof of Statement (ii)

Differently from the proof in the previous subsection, owing to the dependence of $\dot{u}_{i,j}^{(1)}$ and ϵ_{it}, we can not directly use conditional exponential inequality. In this subsection, we will show how to handle this dependence in detail. Recall that $\mathscr{D}_{e_{it}}$ is the σ-field generated by $\{e_{j,it}\}_{j\in[p]} \cup \left\{V_j^0\right\}_{j\in[p]\cup\{0\}} \cup \left\{W_j^0\right\}_{j\in[p]}$ and define

$$\dot{Q}_{\tau t,v}\left(\{v_{t,j}\}_{j\in[p]\cup\{0\}}\right) = \frac{1}{N_2}\sum_{i\in I_2}\rho_\tau\left(Y_{it} - v'_{t,0}\dot{u}_{i,0}^{(1)} - \sum_{j=1}^{p} v'_{t,j}\dot{u}_{i,j}^{(1)}X_{j,it}\right),$$

$$v_t^0 = \left(v_{t,0}^{0\prime},\cdots,v_{t,p}^{0\prime}\right)', \quad \dot{\Delta}_{t,j} = O_j^{(1)\prime}\dot{v}_{t,j}^{(1)} - v_{t,j}^0, \quad \dot{\Delta}_{t,v} = \left(\dot{\Delta}'_{t,0},\cdots,\dot{\Delta}'_{t,p}\right)',$$

$$\dot{\Psi}_{it}^{(1)} = \left[\left(O_0^{(1)\prime}\dot{u}_{i,0}^{(1)}\right)', \left(O_1^{(1)\prime}\dot{u}_{i,1}^{(1)}X_{1,it}\right)', ..., \left(O_p^{(1)\prime}\dot{u}_{i,p}^{(1)}X_{p,it}\right)'\right]',$$

$$\dot{\Psi}_t^{(1)} = \frac{1}{N_2}\sum_{i\in I_2}\dot{\Psi}_{it}^{(1)}\dot{\Psi}_{it}^{(1)\prime}.$$

As in (A.5), we have

$$\begin{aligned}
0 &\geq \dot{Q}_{\tau t,v}\left(\left\{\dot{v}_{t,j}^{(1)}\right\}_{j\in[p]\cup\{0\}}\right) - \dot{Q}_{\tau t,v}\left(\left\{O_j^{(1)}v_{t,j}^0\right\}_{j\in[p]\cup\{0\}}\right) \\
&= \frac{1}{N_2}\sum_{i\in I_2}\rho_\tau\left(Y_{it} - \dot{v}_{t,0}^{(1)\prime}O_0^{(1)}O_0^{(1)\prime}\dot{u}_{i,0}^{(1)} - \sum_{j=1}^{p}\dot{v}_{i,j}^{(1)\prime}O_j^{(1)}O_j^{(1)\prime}\dot{u}_{t,j}^{(1)}X_{j,it}\right) \\
&\quad - \frac{1}{N_2}\sum_{i\in I_2}\rho_\tau\left(Y_{it} - v_{t,0}^{0\prime}O_0^{(1)\prime}\dot{u}_{i,0}^{(1)} - \sum_{j=1}^{p}v_{t,j}^{0\prime}O_j^{(1)\prime}\dot{u}_{t,j}^{(1)}X_{j,it}\right) \\
&= \frac{1}{N_2}\sum_{i\in I_2}\left[\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}\left(\tau - \mathbf{1}\{w_{3,it}\leq 0\}\right)\right] + \frac{1}{N_2}\sum_{i\in I_2}\int_0^{\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}}\left(\mathbf{1}\{w_{3,it}\leq s\} - \mathbf{1}\{w_{3,it}\leq 0\}\right)ds \\
&= \frac{1}{N_2}\sum_{i\in I_2}\left[\Psi_{it}^{0\prime}\dot{\Delta}_{t,v}\left(\tau - \mathbf{1}\{\epsilon_{it}\leq 0\}\right)\right] + \frac{1}{N_2}\sum_{i\in I_2}\left[\left(\dot{\Psi}_{it}^{(1)} - \Psi_{it}^0\right)'\dot{\Delta}_{t,v}\left(\tau - \mathbf{1}\{\epsilon_{it}\leq 0\}\right)\right] \\
&\quad + \frac{1}{N_2}\sum_{i\in I_2}\left[\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}\left(\mathbf{1}\{\epsilon_{it}\leq 0\} - \mathbf{1}\{w_{3,it}\leq 0\}\right)\right] \\
&\quad + \frac{1}{N_2}\sum_{i\in I_2}\int_0^{\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}}\mathbb{E}\left[\left(\mathbf{1}\{\epsilon_{it}\leq s\} - \mathbf{1}\{\epsilon_{it}\leq 0\}\right)ds\middle|\mathscr{D}_{e_{it}}\right] \\
&\quad + \frac{1}{N_2}\sum_{i\in I_2}\left\{\int_0^{\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}}\left[\left(\mathbf{1}\{\epsilon_{it}\leq s\} - \mathbf{1}\{\epsilon_{it}\leq 0\}\right) - \mathbb{E}\left(\left(\mathbf{1}\{\epsilon_{it}\leq s\} - \mathbf{1}\{\epsilon_{it}\leq 0\}\right)ds\middle|\mathscr{D}_{e_{it}}\right)\right]\right\} \\
&\quad + \frac{1}{N_2}\sum_{i\in I_2}\int_0^{\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}}\left(\mathbf{1}\{w_{3,it}\leq s\} - \mathbf{1}\{\epsilon_{it}\leq s\}\right)ds \\
&\quad + \frac{1}{N_2}\sum_{i\in I_2}\int_0^{\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}}\left(\mathbf{1}\{\epsilon_{it}\leq 0\} - \mathbf{1}\{w_{3,it}\leq 0\}\right)ds \\
&= \frac{1}{N_2}\sum_{i\in I_2}\left[\Psi_{it}^{0\prime}\dot{\Delta}_{t,v}\left(\tau - \mathbf{1}\{\epsilon_{it}\leq 0\}\right)\right] + \frac{1}{N_2}\sum_{i\in I_2}\left[\left(\dot{\Psi}_{it}^{(1)} - \Psi_{it}^0\right)'\dot{\Delta}_{t,v}\left(\tau - \mathbf{1}\{\epsilon_{it}\leq 0\}\right)\right]
\end{aligned}$$

$$
\begin{aligned}
&+\frac{2}{N_2}\sum_{i\in I_2}\left[\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}\left(\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{w_{3,it}\leq 0\right\}\right)\right]\\
&+\frac{1}{N_2}\sum_{i\in I_2}\int_0^{\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}}\mathbb{E}\left[\left(\mathbf{1}\left\{\epsilon_{it}\leq s\right\}-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\middle|\mathscr{D}_{e_{it}}\right]ds\\
&+\frac{1}{N_2}\sum_{i\in I_2}\left\{\int_0^{\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}}\left[\left(\mathbf{1}\left\{\epsilon_{it}\leq s\right\}-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)-\mathbb{E}\left(\left(\mathbf{1}\left\{\epsilon_{it}\leq s\right\}-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\middle|\mathscr{D}_{e_{it}}\right)\right]ds\right\}\\
&+\frac{1}{N_2}\sum_{i\in I_2}\int_0^{\dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}}\left(\mathbf{1}\left\{w_{3,it}\leq s\right\}-\mathbf{1}\left\{\epsilon_{it}\leq s\right\}\right)ds\\
&:=\sum_{m=1}^{6}B_{m,t} \qquad\qquad (A.8)
\end{aligned}
$$

where $w_{3,it}=Y_{it}-v_{t,0}^{0\prime}O_0^{(1)\prime}\dot{u}_{i,0}^{(1)}-\sum_{j=1}^{p}v_{t,j}^{0\prime}O_j^{(1)\prime}\dot{u}_{i,j}^{(1)}X_{j,it}=Y_{it}-v_t^{0\prime}\dot{\Psi}_{it}^{(1)}=\epsilon_{it}-v_t^{0\prime}\left(\dot{\Psi}_{it}^{(1)}-\Psi_{it}^{0}\right)$, the last equality is by the fact that the third and the last terms after the second equality are identical. Then we obtain $|B_{4,t}|\leq\sum_{m\neq 4}|B_{m,t}|$. By Lemma B.16 and similar arguments for Theorem 3.2(i), we obtain that

$$
\max_{t\in[T]}\left\|O_j^{(1)\prime}\dot{v}_{t,j}^{(1)}-v_{t,j}^{0}\right\|_2\leq\left\|\dot{\Delta}_{t,v}\right\|_2=O_p(\eta_N).
$$

■

A.2.3 Proof of Statement (iii)

In this proof we derive the linear expansion of $\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}$ for each $i\in I_3$. Recall that $\forall i\in I_3$,

$$
\left\{\dot{u}_{i,j}^{(1)}\right\}_{j\in[p]\cup\{0\}}=\underset{\{u_{i,j}\}_{j\in[p]\cup\{0\}}}{\arg\min}\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(Y_{it}-u_{i,0}^{\prime}\dot{v}_{t,0}^{(1)}-\sum_{j\in[p]}u_{i,j}^{\prime}\dot{v}_{t,j}^{(1)}X_{j,it}\right).
$$

Define $\dot{\varpi}_{it}=\left(\dot{v}_{t,0}^{(1)\prime},X_{1,it}\dot{v}_{t,1}^{(1)\prime},\cdots,X_{p,it}\dot{v}_{t,p}^{(1)\prime}\right)^{\prime}$ and

$$
\begin{aligned}
\dot{\mathbb{H}}_i\left(\{u_{i,j}\}_{j\in[p]\cup\{0\}}\right)&=\frac{1}{T}\sum_{t=1}^{T}\left[\tau-\mathbf{1}\left\{\epsilon_{it}\leq g_{it}(\{u_{i,j}\}_{j\in[p]\cup\{0\}})\right\}\right]\dot{\varpi}_{it},\\
\dot{\mathcal{H}}_i\left(\{u_{i,j}\}_{j\in[p]\cup\{0\}}\right)&=\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left\{\left[\tau-\mathbf{1}\left\{\epsilon_{it}\leq g_{it}(\{u_{i,j}\}_{j\in[p]\cup\{0\}})\right\}\right]\dot{\varpi}_{it}\middle|\mathscr{D}_{e_i}^{I_1\cup I_2}\right\},\\
&=\frac{1}{T}\sum_{t=1}^{T}\left\{\left[\tau-\mathfrak{F}_{it}\left(g_{it}(\{u_{i,j}\}_{j\in[p]\cup\{0\}})\right)\right]\dot{\varpi}_{it}\right\},\\
\dot{\mathbb{W}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right)&=\dot{\mathbb{H}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right)-\dot{\mathbb{H}}_i\left(\{O_j^{(1)}u_{i,j}^{0}\}_{j\in[p]\cup\{0\}}\right)\\
&\quad-\left\{\dot{\mathcal{H}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right)-\dot{\mathcal{H}}_i\left(\{O_j^{(1)}u_{i,j}^{0}\}_{j\in[p]\cup\{0\}}\right)\right\},
\end{aligned}
$$

where $\mathscr{D}_{e_i}^{I_1\cup I_2}$ being the σ-field generated by $\{\epsilon_{i^*t},e_{i^*t}\}_{i^*\in I_1\cup I_2,t\in[T]}\bigcup\{e_{it}\}_{t\in[T]}\bigcup\left\{V_j^0\right\}_{j\in[p]\cup\{0\}}\bigcup\left\{W_j^0\right\}_{j\in[p]}$, and

$$
g_{it}(\{u_{i,j}\}_{j\in[p]\cup\{0\}})=u_{i,0}^{\prime}\dot{v}_{t,0}^{(1)}+\sum_{j\in[p]}u_{i,j}^{\prime}\dot{v}_{t,j}^{(1)}X_{j,it}-u_{i,0}^{0\prime}v_{t,0}^{0}-\sum_{j\in[p]}u_{i,j}^{0\prime}v_{t,j}^{0}X_{j,it}.
$$

Then we have

$$\dot{\mathbb{H}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right)=\dot{\mathbb{H}}_i\left(\{O_j^{(1)}u_{i,j}^0\}_{j\in[p]\cup\{0\}}\right)+\dot{\mathcal{H}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right)-\dot{\mathcal{H}}_i\left(\{O_j^{(1)}u_{i,j}^0\}_{j\in[p]\cup\{0\}}\right)$$
$$+\dot{\mathbb{W}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right). \tag{A.9}$$

By Assumptions 1(v) and 2 and Theorem 3.2(ii), we have $\max_{i\in I_3,t\in[T]}\|\dot{\varpi}_{it}\|_2\le C_{11}\xi_N$ a.s.. By the first order condition of the quantile regression, we have

$$\max_{i\in I_3}\left\|\dot{\mathbb{H}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right)\right\|_2=O_p\left(\frac{1}{T}\max_{i\in I_3,t\in[T]}\|\dot{\varpi}_{it}\|_2\right)=O_p\left(\frac{\xi_N}{T}\right). \tag{A.10}$$

Next, we show that $\max_{i\in I_3}\left\|\dot{\mathbb{W}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right)\right\|_2=o_p\left((N\vee T)^{-1/2}\right)$. Notice that

$$\begin{aligned}\dot{\mathbb{W}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right)&=\frac{1}{T}\sum_{t=1}^T\dot{\varpi}_{it}\left(\mathbf{1}\left\{\epsilon_{it}\le g_{it}(\{O_j^{(1)}u_{i,j}^0\}_{j\in[p]\cup\{0\}})\right\}-\mathbf{1}\left\{\epsilon_{it}\le g_{it}(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}})\right\}\right)\\&\quad-\frac{1}{T}\sum_{t\in[T]}\dot{\varpi}_{it}\left[\mathfrak{F}_{it}\left(g(\{O_j^{(1)}u_{i,j}^0\}_{j\in[p]\cup\{0\}})\right)-\mathfrak{F}_{it}\left(g_{it}(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}})\right)\right]\\&=\dot{\mathbb{W}}_i^I-\dot{\mathbb{W}}_i^{II}\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right),\end{aligned} \tag{A.11}$$

where $\dot{\mathbb{W}}_i^I=\frac{1}{T}\sum_{t=1}^T\dot{\mathbb{W}}_{it}^I$ with

$$\dot{\mathbb{W}}_{it}^I=\dot{\varpi}_{it}\left\{\left(\mathbf{1}\left\{\epsilon_{it}\le g_{it}(\{O_j^{(1)}u_{i,j}^0\}_{j\in[p]\cup\{0\}})\right\}-\mathbf{1}\{\epsilon_{it}\le 0\}\right)-\left[\mathfrak{F}_{it}\left(g_{it}(\{O_j^{(1)}u_{i,j}^0\}_{j\in[p]\cup\{0\}})\right)-\mathfrak{F}_{it}(0)\right]\right\}$$

and

$$\begin{aligned}\dot{\mathbb{W}}_i^{II}\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}}\right)&=\frac{1}{T}\sum_{t=1}^T\dot{\varpi}_{it}\left\{\left(\mathbf{1}\left\{\epsilon_{it}\le g_{it}(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}})\right\}-\mathbf{1}\{\epsilon_{it}\le 0\}\right)\right.\\&\quad\left.-\left[\mathfrak{F}_{it}\left(g_{it}(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}})\right)-\mathfrak{F}_{it}(0)\right]\right\}.\end{aligned}$$

Noting that

$$g_{it}(\{O_j^{(1)}u_{i,j}^0\}_{j\in[p]\cup\{0\}})=\left(O_0^{(1)}u_{i,0}\right)'\left(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right)+\sum_{j\in[p]}\left(O_j^{(1)}u_{i,j}\right)'\left(\dot{v}_{t,j}^{(1)}-O_j^{(1)}v_{t,j}^0\right)X_{j,it},$$

and $\max_{i\in I_2,t\in[T]}\left|g_{it}(\{O_j^{(1)}u_{i,j}^0\}_{j\in[p]\cup\{0\}})\right|=O_p(\xi_N\eta_N)$, we have, $\max_{i\in I_3,t\in[T]}\left\|\dot{\mathbb{W}}_{it}^I\right\|_2=O_p(\xi_N)$, $\mathbb{E}\left(\mathbb{W}_{it}^I\middle|\mathscr{D}_{e_i}^{I_1\cup I_2}\right)=0$, and for some positive constants C_{11},

$$\begin{aligned}&\max_{i\in I_3,t\in[T]}\left\|Var\left(\dot{\mathbb{W}}_{it}^I\middle|\mathscr{D}_{e_i}^{I_1\cup I_2}\right)\right\|_F\\&\le\max_{i\in I_3,t\in[T]}\left\|\mathbb{E}\left\{\dot{\varpi}_{it}\dot{\varpi}_{it}'\left(\mathbf{1}\left\{\epsilon_{it}\le g(\{O_j^{(1)}u_{i,j}^0\}_{j\in[p]\cup\{0\}})\right\}-\mathbf{1}\{\epsilon_{it}\le 0\}\right)^2\middle|\mathscr{D}_{e_i}^{I_1\cup I_2}\right\}\right\|_F\\&\le\max_{i\in I_3,t\in[T]}\|\dot{\varpi}_{it}\|_2^2\,\mathbb{E}\left(\mathbf{1}\left\{0\le|\epsilon_{it}|\le\left|g(\{O_j^{(1)}u_{i,j}^0\}_{j\in[p]\cup\{0\}})\right|\right\}\middle|\mathscr{D}_{e_i}^{I_1\cup I_2}\right)=O_p(\xi_N^3\eta_N),\end{aligned}$$

and

$$\max_{i\in I_3,t\in[T]}\sum_{s=t+1}^T\left\|Cov\left(\dot{\mathbb{W}}_{it}^I,\dot{\mathbb{W}}_{is}^I\middle|\mathscr{D}_{e_i}^{I_1\cup I_2}\right)\right\|_F$$

$$
\begin{aligned}
&\lesssim \max_{i\in I_3,t\in[T]} \sum_{s=t+1}^{T} \left[\mathbb{E}\left(\left\|\dot{\mathbb{W}}_{it}^{I}\right\|_F^{2+\vartheta} \middle| \mathscr{D}_{e_i}^{I_1\cup I_2}\right)\right]^{\frac{1}{2+\vartheta}} \left[\mathbb{E}\left(\left\|\dot{\mathbb{W}}_{is}^{I}\right\|_F^{2+\vartheta} \middle| \mathscr{D}_{e_i}^{I_1\cup I_2}\right)\right]^{\frac{1}{2+\vartheta}} [\alpha(t-s)]^{1-\frac{2}{2+\vartheta}} \\
&\leq \max_{i\in I_3,t\in[T]} \left[\mathbb{E}\left(\left\|\dot{\mathbb{W}}_{it}^{I}\right\|_F^{2+\vartheta} \middle| \mathscr{D}_{e_i}^{I_1\cup I_2}\right)\right]^{\frac{2}{2+\vartheta}} \max_{t\in[T]} \sum_{s=t+1}^{T} [\alpha(t-s)]^{1-\frac{2}{2+\vartheta}} \text{ a.s.} \\
&= O_p\left(\xi_N^{\frac{6+2\vartheta}{2+\vartheta}} \eta_N^{\frac{2}{2+\vartheta}}\right),
\end{aligned}
$$

for any $\vartheta > 0$, where the first inequality holds by Davydov's inequality for conditional strong mixing processes, and the last equality holds by the fact that $\mathbb{E}\left(\left\|\dot{\mathbb{W}}_{it}^{I}\right\|^{2+\vartheta} \middle| \mathscr{D}_{e_i}^{I_1\cup I_2}\right) = O_p\left(\xi_N^{3+\vartheta}\eta_N\right)$ following the similar argument as in (A.10). Combining (A.10) and (A.11) yields

$$
\max_{i\in I_3,t\in[T]} \left\{\left\|Var\left(\dot{\mathbb{W}}_{it}^{I}\right)\right\|_F + 2\sum_{s>t}\left\|Cov\left(\dot{\mathbb{W}}_{it}^{I}, \dot{\mathbb{W}}_{is}^{I}\right)\right\|_F\right\} = O_p\left(\xi_N^{\frac{6+2\vartheta}{2+\vartheta}} \eta_N^{\frac{2}{2+\vartheta}}\right).
$$

For any constant $C_{12} > 0$, define

$$
\begin{aligned}
\mathscr{A}_{5,i}(C_{12}) &= \left\{\max_{t\in[T]} \left\{\left\|Var\left(\dot{\mathbb{W}}_{it}^{I}\right)\right\|_F + 2\sum_{s>t}\left\|Cov\left(\dot{\mathbb{W}}_{it}^{I}, \dot{\mathbb{W}}_{is}^{I}\right)\right\|_F\right\} \leq C_{12}\xi_N^{\frac{6+2\vartheta}{2+\vartheta}} \eta_N^{\frac{2}{2+\vartheta}}\right\}, \\
\mathscr{A}_{5}(C_{12}) &= \left\{\max_{i\in I_3,t\in[T]} \left\{\left\|Var\left(\dot{\mathbb{W}}_{it}^{I}\right)\right\|_F + 2\sum_{s>t}\left\|Cov\left(\dot{\mathbb{W}}_{it}^{I}, \dot{\mathbb{W}}_{is}^{I}\right)\right\|_F\right\} \leq C_{12}\xi_N^{\frac{6+2\vartheta}{2+\vartheta}} \eta_N^{\frac{2}{2+\vartheta}}\right\}.
\end{aligned}
$$

For any $e > 0$, we can find a sufficiently large constants C_{12} such that $\mathbb{P}(\mathscr{A}_5^c(C_{12})) \leq e$. Therefore, we have

$$
\begin{aligned}
&\mathbb{P}\left\{\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\dot{\mathbb{W}}_{it}^{I}\right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}\right\} \\
&\leq \mathbb{P}\left\{\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\dot{\mathbb{W}}_{it}^{I}\right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}, \mathscr{A}_5(C_{12})\right\} + e \\
&\leq \sum_{i\in I_3}\mathbb{P}\left\{\left\|\frac{1}{T}\sum_{t=1}^{T}\dot{\mathbb{W}}_{it}^{I}\right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}, \mathscr{A}_5(C_{12})\right\} + e \\
&\leq \sum_{i\in I_3}\mathbb{P}\left\{\left\|\frac{1}{T}\sum_{t=1}^{T}\dot{\mathbb{W}}_{it}^{I}\right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}, \mathscr{A}_{5,i}(C_{12})\right\} + e \\
&= \sum_{i\in I_3}\mathbb{E}\mathbb{P}\left\{\left\|\frac{1}{T}\sum_{t=1}^{T}\dot{\mathbb{W}}_{it}^{I}\right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}\middle|\mathscr{D}_{e_i}^{I_1\cup I_2}\right\} 1\{\mathscr{A}_{5,i}(C_{12})\} + e,
\end{aligned}
\tag{A.12}
$$

where the second inequality is by the union bound, the third inequality is by $\mathscr{A}_5(C_{12}) \subset \mathscr{A}_{5,i}(C_{12})$, and the last equality is owing to the fact that $\mathscr{A}_{5,i}(C_{12})$ is $\mathscr{D}_{e_i}^{I_1\cup I_2}$ measurable. Given $\mathscr{D}_{e_i}^{I_1\cup I_2}$, the randomness in $\dot{\mathbb{W}}_{i}^{I}$ only comes from $\{\epsilon_{it}\}_{t\in[T]}$, which are strong mixing given $\mathscr{D}_{e_i}^{I_1\cup I_2}$. Therefore, on $\mathscr{A}_{5,i}(C_{12})$, Lemma B.12(ii) implies

$$
\mathbb{P}\left\{\left\|\frac{1}{T}\sum_{t=1}^{T}\dot{\mathbb{W}}_{it}^{I}\right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}\middle|\mathscr{D}_{e_i}^{I_1\cup I_2}\right\}
$$

$$= \mathbb{P}\left\{ \left\| \sum_{t=1}^{T} \dot{\mathbb{W}}_{it}^{I} \right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N \vee T)}{N \wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{T \log(N \vee T)} \middle| \mathscr{D}_{e_i}^{I_1 \cup I_2} \right\}$$

$$\leq \exp\left\{ - \frac{c_{12} C_{13}^2 T \left(\frac{\log(N \vee T)}{N \wedge T} \right)^{\frac{1}{2+\vartheta}} \xi_N^{\frac{10+2\vartheta}{2+\vartheta}} \log(N \vee T)}{C_{12} T \left(\frac{\log(N \vee T)}{N \wedge T} \right)^{\frac{1}{2+\vartheta}} \xi_N^{\frac{10+2\vartheta}{2+\vartheta}} + C_{11}^2 \xi_N^2 + C_{11} C_{13} \sqrt{T \log(N \vee T)} \xi_N^{\frac{7+2\vartheta}{2+\vartheta}} \left(\frac{\log(N \vee T)}{N \wedge T} \right)^{\frac{1}{4+2\vartheta}} (\log T)^2} \right\}, \tag{A.13}$$

which further implies

$$\mathbb{P}\left\{ \max_{i \in I_3} \left\| \frac{1}{T} \sum_{t=1}^{T} \dot{\mathbb{W}}_{it}^{I} \right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N \vee T)}{N \wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N \vee T)}{T}} \right\} = o(1) + e.$$

As e is arbitrary, by Assumption 1(ix), we obtain that

$$\max_{i \in I_3} \left\| \frac{1}{T} \sum_{t=1}^{T} \dot{\mathbb{W}}_{it}^{I} \right\|_2 = O_p\left(C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N \vee T)}{N \wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N \vee T)}{T}} \right) = o_p\left((N \vee T)^{-1/2} \right).$$

For $\dot{\mathbb{W}}_i^{II}\left(\{\dot{u}_{i,j}^{(1)}\}_{j \in [p] \cup \{0\}} \right)$, we observe that

$$\begin{aligned}
g_{it}(\{\dot{u}_{i,j}^{(1)}\}_{j \in [p] \cup \{0\}}) &= \dot{u}_{i,0}^{(1)\prime} \dot{v}_{t,0}^{(1)} + \sum_{j \in [p]} \dot{u}_{i,j}^{(1)\prime} \dot{v}_{t,j}^{(1)} X_{j,it} - u_{i,0}^{0\prime} v_{t,0}^{0} - \sum_{j \in [p]} u_{i,j}^{0\prime} v_{t,j}^{0} X_{j,it} \\
&= \left(\dot{u}_{i,0}^{(1)} - O_0^{(1)} u_{i,0}^{0} \right)' \dot{v}_{t,0}^{(1)} + \left(O_0^{(1)} u_{i,0}^{0} \right)' \left(\dot{v}_{t,0}^{(1)} - O_0^{(1)} v_{t,0}^{0} \right) \\
&\quad + \sum_{j \in [p]} \left(\dot{u}_{i,j}^{(1)} - O_j^{(1)} u_{i,j}^{0} \right)' \dot{v}_{t,j}^{(1)} X_{j,it} + \sum_{j \in [p]} \left(O_j^{(1)} u_{i,j}^{0} \right)' \left(\dot{v}_{t,j}^{(1)} - O_j^{(1)} v_{t,j}^{0} \right) X_{j,it} \\
&= \dot{\Delta}_{i,u}' \dot{\Phi}_{it}^{(1)} + \dot{\Delta}_{t,v}' \Psi_{it}^{0},
\end{aligned}$$

where $\dot{\Delta}_{i,u} = \left(\dot{\Delta}_{i,0}', \cdots, \dot{\Delta}_{i,p}' \right)'$ with $\dot{\Delta}_{i,j} = O_j^{(1)\prime} \dot{u}_{i,j}^{(1)} - u_{i,j}^{0}$, $\dot{\Delta}_{t,v} = \left(\dot{\Delta}_{t,0}', \cdots, \dot{\Delta}_{t,p}' \right)'$ with $\dot{\Delta}_{t,j} = O_j^{(1)\prime} \dot{v}_{t,j}^{(1)} - v_{t,j}^{0}$,

$$\begin{aligned}
\dot{\Phi}_{it}^{(1)} &= \left[\left(O_0^{(1)\prime} \dot{v}_{t,0}^{(1)} \right)', \left(O_1^{(1)\prime} \dot{v}_{t,1}^{(1)} X_{1,it} \right)', \cdots, \left(O_p^{(1)\prime} \dot{v}_{t,p}^{(1)} X_{p,it} \right)' \right]', \quad \text{and} \\
\Psi_{it}^{0} &= (u_{i,0}^{0\prime}, u_{i,1}^{0\prime} X_{1,it}, ..., u_{i,p}^{0\prime} X_{p,it})'.
\end{aligned}$$

Unlike the analysis for $\dot{\mathbb{W}}_{it}^{I}$, to handle the dependence between ϵ_{it} and $\dot{\Delta}_{i,u}$, for any constant $C_{14} > 0$, we first define an event set $\mathscr{A}_6(C_{14}) = \left\{ \max_{i \in I_3} \left\| \dot{\Delta}_{i,u} \right\|_2 \leq C_{14}\eta_N \right\}$ with $\mathbb{P}(\mathscr{A}_6^c(C_{14})) \leq e$ for any $e > 0$ by Theorem 3.2(i), then we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{i \in I_3} \left\| \dot{\mathbb{W}}_i^{II}\left(\{\dot{u}_{i,j}^{(1)}\}_{j \in [p] \cup \{0\}} \right) \right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N \vee T)}{N \wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N \vee T)}{T}} \right) \\
&\leq \mathbb{P}\left(\max_{i \in I_3} \left\| \dot{\mathbb{W}}_i^{II}\left(\{\dot{u}_{i,j}^{(1)}\}_{j \in [p] \cup \{0\}} \right) \right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N \vee T)}{N \wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N \vee T)}{T}}, \mathscr{A}_6(C_{14}) \right) + e \\
&\leq \mathbb{P}\left(\sup_{s \in \mathbb{S}} \max_{i \in I_3} \left\| \overline{\mathbb{W}}_i^{II}(s) \right\|_2 > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N \vee T)}{N \wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N \vee T)}{T}} \right) + e
\end{aligned} \tag{A.14}$$

with $\mathbb{S} = \left\{ s \in \mathbb{R}^{\left(\sum_{j\in[p]\cup\{0\}} K_j\right)\times 1} : \|s\|_2 \le C_{14}\eta_N \right\}$ and

$$\overline{\mathbb{W}}_i^{II}(s) = \frac{1}{T}\sum_{t=1}^{T} \dot{\varpi}_{it} \left\{ \left(\mathbf{1}\left\{ \epsilon_{it} \le s'\dot{\Phi}_{it}^{(1)} + \dot{\Delta}_{t,v}'\Psi_{it}^{0} \right\} - \mathbf{1}\left\{ \epsilon_{it} \le 0 \right\} \right) - \left[\mathfrak{F}_{it}\left(s'\dot{\Phi}_{it}^{(1)} + \dot{\Delta}_{t,v}'\Psi_{it}^{0} \right) - \mathfrak{F}_{it}(0) \right] \right\}.$$

Similarly as in (B.37), we sketch the proof. Divide $\mathbb{S}$ into $\mathbb{S}_m$ with center s_m for $m = 1, \cdots, n_{\mathbb{S}}$ if $s \in \mathbb{S}_m$, then $\|s - s_m\|_2 < \frac{\varepsilon}{T}$ and $n_{\mathbb{S}} \lesssim T^{\sum_{j\in[p]\cup\{0\}} K_j}$. Then, $\forall s \in \mathbb{S}_m$, we have

$$\left\|\overline{\mathbb{W}}_i^{II}(s)\right\|_2 \le \left\|\overline{\mathbb{W}}_i^{II}(s_m)\right\|_2 + \left\|\overline{\mathbb{W}}_i^{II}(s) - \overline{\mathbb{W}}_i^{II}(s_m)\right\|_2, \tag{A.15}$$

with

$$\begin{aligned}
&\max_{i\in I_3, m\in[n_{\mathbb{S}}]} \sup_{s\in\mathbb{S}_m} \left\|\overline{\mathbb{W}}_i^{II}(s) - \overline{\mathbb{W}}_i^{II}(s_m)\right\|_2 \\
&\le \max_{i\in I_3, m\in[n_{\mathbb{S}}]} \sup_{s\in\mathbb{S}_m} \left\| \frac{1}{T}\sum_{t=1}^{T} \dot{\varpi}_{it} \left(\mathbf{1}\left\{ \epsilon_{it} \le s'\dot{\Phi}_{it}^{(1)} + \dot{\Delta}_{t,v}'\Psi_{it}^{0} \right\} - \mathbf{1}\left\{ \epsilon_{it} \le s_m'\dot{\Phi}_{it}^{(1)} + \dot{\Delta}_{t,v}'\Psi_{it}^{0} \right\} \right) \right\|_2 \\
&+ \max_{i\in I_3, m\in[n_{\mathbb{S}}]} \sup_{s\in\mathbb{S}_m} \left\| \frac{1}{T}\sum_{t=1}^{T} \dot{\varpi}_{it} \left[\mathfrak{F}_{it}\left(s'\dot{\Phi}_{it}^{(1)} + \dot{\Delta}_{t,v}'\Psi_{it}^{0} \right) - \mathfrak{F}_{it}\left(s_m'\dot{\Phi}_{it}^{(1)} + \dot{\Delta}_{t,v}'\Psi_{it}^{0} \right) \right] \right\|_2 \\
&\le \max_{i\in I_3, m\in[n_{\mathbb{S}}]} \frac{1}{T}\sum_{t\in[T]} \|\dot{\varpi}_{it}\|_2 \mathbb{E}\left(\mathbf{1}\left\{ \left| \epsilon_{it} - \dot{\Delta}_{t,v}'\Psi_{it}^{0} - s_m'\dot{\Phi}_{it}^{(1)} \right| \le \frac{\varepsilon||\dot{\Phi}_{it}^{(1)}||_2}{T} \right\} \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right) + \max_{i\in I_2, m\in[n_{\mathbb{S}}]} \left| \overline{\mathbb{W}}_i^{III}(m) \right| \\
&+ \max_{i\in I_3, m\in[n_{\mathbb{S}}]} \sup_{s\in\mathbb{S}_m} \frac{1}{T}\sum_{t\in[T]} \|\dot{\varpi}_{it}\|_2 ||\dot{\Phi}_{it}^{(1)}||_2 \|s - s_m\|_2
\end{aligned} \tag{A.16}$$

such that $\overline{\mathbb{W}}_i^{III}(m) := \frac{1}{T}\sum_{t\in[T]} \overline{\mathbb{W}}_{it}^{III}(m)$ and

$$\begin{aligned}
\overline{\mathbb{W}}_{it}^{III}(m) := \|\dot{\varpi}_{it}\|_2 &\left[\left(\mathbf{1}\left\{ \left| \epsilon_{it} - \dot{\Delta}_{t,v}'\Psi_{it}^{0} - s_m'\dot{\Phi}_{it}^{(1)} \right| \le \frac{\varepsilon||\dot{\Phi}_{it}^{(1)}||_2}{T} \right\} \right) \right. \\
&\left. - \mathbb{E}\left(\mathbf{1}\left\{ \left| \epsilon_{it} - \dot{\Delta}_{t,v}'\Psi_{it}^{0} - s_m'\dot{\Phi}_{it}^{(1)} \right| \le \frac{\varepsilon||\dot{\Phi}_{it}^{(1)}||_2}{T} \right\} \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right) \right].
\end{aligned}$$

Like (B.39), we can show that $\max_{i\in I_3, m\in[n_{\mathbb{S}}]} \frac{1}{T}\sum_{t\in[T]} \|\dot{\varpi}_{it}\|_2 \mathbb{E}\left(\mathbf{1}\left\{ \left| \epsilon_{it} - \dot{\Delta}_{t,v}'\Psi_{it}^{0} - s_m'\dot{\Phi}_{it} \right| \le \frac{\varepsilon||\dot{\Phi}_{it}^{(1)}||_2}{T} \right\} \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right)$ $= O_p(\frac{\varepsilon}{T})$ because $\dot{\Delta}_{t,v}'\Psi_{it}^{0} + s_m'\dot{\Phi}_{it}^{(1)}$ and $||\dot{\Phi}_{it}||$ are measurable in $\mathscr{D}_{e_i}^{I_1\cup I_2}$ and the conditional density of ϵ_{it} given $\mathscr{D}_{e_i}^{I_1\cup I_2}$ is bounded. Also we have

$$\max_{i\in I_3, m\in[n_{\mathbb{S}}]} \sup_{s\in\mathbb{S}_m} \frac{1}{T}\sum_{t\in[T]} \|\dot{\varpi}_{it}\|_2 ||\dot{\Phi}_{it}^{(1)}||_2 \|s - s_m\|_2 = \frac{\varepsilon}{T} \max_{i\in I_3} \frac{1}{T}\sum_{t\in[T]} \|\dot{\varpi}_{it}\|_2 ||\dot{\Phi}_{it}^{(1)}||_2 = O_p\left(\frac{\varepsilon}{T}\right).$$

In addition, we note that

$$\begin{aligned}
&\max_{i\in I_3, m\in[n_{\mathbb{S}}], t\in[T]} \|\dot{\varpi}_{it}\|_2 = O_p(\xi_N), \quad \max_{i\in I_3, m\in[n_{\mathbb{S}}], t\in[T]} Var\left(\overline{\mathbb{W}}_{it}^{III}(m) \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right) = O_p\left(\frac{\xi_N^3 \varepsilon}{T} \right), \\
&\max_{i\in I_3, m\in[n_{\mathbb{S}}], t\in[T]} \sum_{s=t+1}^{T} \left| Cov\left(\overline{\mathbb{W}}_{it}^{III}(m), \overline{\mathbb{W}}_{is}^{III}(m) \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right) \right| = O_p\left(\xi_N^{8/3} \left(\frac{\varepsilon}{T} \right)^{2/3} \right),
\end{aligned}$$

and for any positive constant C_{15} and C_{16}, define event set

$$\mathscr{A}_{8,N} = \left(\begin{array}{c} \max_{i\in I_3, m\in[n_{\mathbb{S}}], t\in[T]} \left\{ Var\left(\overline{\mathbb{W}}_{it}^{III}(m) \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right) + 2\sum_{s=t+1}^{T} \left| Cov\left(\overline{\mathbb{W}}_{it}^{III}(m), \overline{\mathbb{W}}_{is}^{III}(m) \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right) \right| \right\} \leq C_{15}\xi_N^{8/3} \left(\frac{\varepsilon}{T}\right)^{2/3} \\ \text{and} \max_{i\in I_3, t\in[T]} \|\dot{\varpi}_{it}\|_2 \leq C_{16}\xi_N \end{array} \right),$$

$$\mathscr{A}_{8,N,i} = \left(\begin{array}{c} \max_{m\in[n_{\mathbb{S}}], t\in[T]} \left\{ Var\left(\overline{\mathbb{W}}_{it}^{III}(m) \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right) + 2\sum_{s=t+1}^{T} \left| Cov\left(\overline{\mathbb{W}}_{it}^{III}(m), \overline{\mathbb{W}}_{is}^{III}(m) \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right) \right| \right\} \leq C_{15}\xi_N^{8/3} \left(\frac{\varepsilon}{T}\right)^{2/3} \\ \text{and} \max_{t\in[T]} \|\dot{\varpi}_{it}\|_2 \leq C_{16}\xi_N \end{array} \right),$$

with $\mathbb{P}\left(\mathscr{A}_{8,N}^c\right) \leq e$ for any positive e. Then we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_3, m\in[n_{\mathbb{S}}]} \left| \overline{\mathbb{W}}_i^{III}(m) \right| > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N\vee T)}{N\wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N\vee T)}{T}} \right) \\
&\leq \mathbb{P}\left(\max_{i\in I_3, m\in[n_{\mathbb{S}}]} \left| \overline{\mathbb{W}}_i^{III}(m) \right| > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N\vee T)}{N\wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N\vee T)}{T}}, \mathscr{A}_{8,N} \right) + e \\
&\leq \sum_{i\in[I_3], m\in[n_{\mathbb{S}}]} \mathbb{P}\left(\left| \overline{\mathbb{W}}_i^{III}(m) \right| > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N\vee T)}{N\wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N\vee T)}{T}}, \mathscr{A}_{8,N,i} \right) + e \\
&= \sum_{i\in[I_3], m\in[n_{\mathbb{S}}]} \mathbb{E}\mathbb{P}\left(\left| \overline{\mathbb{W}}_i^{III}(m) \right| > C_{13}\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N\vee T)}{N\wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N\vee T)}{T}} \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right) \mathbf{1}\{\mathscr{A}_{8,N,i}\} + e \\
&= o(1) + e, \qquad \text{(A.17)}
\end{aligned}$$

where the last line is by Bernstein's inequality similar to (A.13). As e is arbitrary, we have

$$\max_{i\in I_3, m\in[n_{\mathbb{S}}]} \left\| \overline{\mathbb{W}}_i^{III}(m) \right\|_2 = O_p\left(\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N\vee T)}{N\wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N\vee T)}{T}} \right),$$

which implies that $\max_{i\in I_3, m\in[n_{\mathbb{S}}]} \sup_{s\in\mathbb{S}_m} \left\| \overline{\mathbb{W}}_i^{II}(s) - \overline{\mathbb{W}}_i^{II}(s_m) \right\|_2 = O_p\left(\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N\vee T)}{N\wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N\vee T)}{T}} \right)$. Following the same argument in (A.17), we can show that

$$\max_{i\in I_3, m\in[n_{\mathbb{S}}]} \left\| \overline{\mathbb{W}}_i^{II}(s_m) \right\|_2 = O_p\left(\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N\vee T)}{N\wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N\vee T)}{T}} \right),$$

which, combined with (A.14) and (A.15), implies that

$$\begin{aligned}
&\max_{i\in I_3} \left\| \dot{\mathbb{W}}_i^{II}\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}} \right) \right\|_2 = O_p\left(\xi_N^{\frac{5+\vartheta}{2+\vartheta}} \left(\frac{\log(N\vee T)}{N\wedge T} \right)^{\frac{1}{4+2\vartheta}} \sqrt{\frac{\log(N\vee T)}{T}} \right) = o_p\left((N\vee T)^{-1/2} \right), \quad \text{and thus,} \\
&\max_{i\in I_3} \left\| \dot{\mathbb{W}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j\in[p]\cup\{0\}} \right) \right\|_2 = o_p\left((N\vee T)^{-1/2} \right). \qquad \text{(A.18)}
\end{aligned}$$

Next, we observe that

$$\dot{\mathcal{H}}_i\left(\{O_j^{(1)} u_{i,j}^0\}_{j=0}^p \right) - \dot{\mathcal{H}}_i\left(\{\dot{u}_{i,j}^{(1)}\}_{j=0}^p \right)$$

$$= \frac{1}{T}\sum_{t=1}^{T}\mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^{0}\right)\dot{\varpi}_{it}\dot{\varpi}_{it}'\dot{\Delta}_{i,u} + O_p\left(\max_{i\in i_3}\left\|\dot{\Delta}_{i,u}\right\|_2^2\right)$$

$$= \frac{1}{T}\sum_{t=1}^{T}\mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^{0}\right)\dot{\varpi}_{it}\dot{\varpi}_{it}'\dot{\Delta}_{i,u} + o_p\left((N\vee T)^{-1/2}\right) \quad \text{uniformly over } i\in I_3, \tag{A.19}$$

where the first equality holds by Taylor expansion and Lemma B.17 and the second equality is by Assumption 1(ix). Combining (A.9), (A.10), (A.18) and (A.19), we have shown that

$$\dot{\Delta}_{i,u} = \left[\dot{D}_i^I\right]^{-1}\dot{D}_i^{II} + o_p\left((N\vee T)^{-1/2}\right) \quad \text{uniformly over } i\in I_3, \tag{A.20}$$

where $\dot{D}_i^I := \frac{1}{T}\sum_{t=1}^{T}\mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^{0}\right)\dot{\varpi}_{it}\dot{\varpi}_{it}'$ and $\dot{D}_i^{II} := \frac{1}{T}\sum_{t=1}^{T}\left[\tau - \mathbf{1}\left\{\epsilon_{it}\le \dot{\Delta}_{t,v}'\Psi_{it}^{0}\right\}\right]\dot{\varpi}_{it}$. Recall that $\varpi_{it}^0 = \left(\left(O_0^{(1)}v_{t,0}^0\right)', \left(O_1^{(1)}v_{t,1}^0\right)'X_{1,it}, \cdots, \left(O_1^{(1)}v_{t,p}^0\right)'X_{p,it}\right)'$, $D_i^I = \frac{1}{T}\sum_{t=1}^{T}\mathfrak{f}_{it}(0)\varpi_{it}^0\varpi_{it}^{0\prime}$, and $D_i^{II} = \frac{1}{T}\sum_{t=1}^{T}\left[\tau - \mathbf{1}\left\{\epsilon_{it}\le 0\right\}\right]\varpi_{it}^0$. Noting that

$$\begin{aligned}\left[\dot{D}_i^I\right]^{-1}\dot{D}_i^{II} &= \left[D_i^I\right]^{-1}D_i^{II} + \left[D_i^I\right]^{-1}\left(\dot{D}_i^{II} - D_i^{II}\right) + \left[\left(\dot{D}_i^I\right)^{-1} - \left(D_i^I\right)^{-1}\right]D_i^{II} \\ &\quad + \left[\left(\dot{D}_i^I\right)^{-1} - \left(D_i^I\right)^{-1}\right]\left(\dot{D}_i^{II} - D_i^{II}\right),\end{aligned}$$

we have by (A.20) and Lemma B.18, uniformly

$$\dot{\Delta}_{i,u} = \left[D_i^I\right]^{-1}D_i^{II} + \left[D_i^I\right]^{-1}\frac{1}{T}\sum_{t=1}^{T}\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\} - \mathbf{1}\left\{\epsilon_{it}\le \dot{\Delta}_{t,v}'\Psi_{it}^{0}\right\}\right]\varpi_{it}^0 + o_p\left((N\vee T)^{-1/2}\right)$$

uniformly over $i\in I_3$, where we use the fact that $\sqrt{\frac{\log(N\vee T)}{T}}\xi_N\eta_N = o\left((N\vee T)^{-1/2}\right)$ by Assumption 1(ix). ■

A.3 Proof of Theorem 3.3

In this section, we extend the distribution theory of the least squares framework in Chernozhukov et al. (2019) to the quantile regression framework and obtain the uniform error bound. We assume the model has only one regressor in this section for notation simplicity.

A.3.1 Proof of Statement (i)

For $\forall i\in I_3$, recall from (2.7) that

$$\left\{\hat{u}_{i,j}^{(3,1)}\right\}_{j\in[p]} = \underset{\{u_{i,j}\}_{j\in[p]}}{\arg\min}\frac{1}{T}\sum_{t=1}^{T}\rho_\tau\left(\tilde{Y}_{it} - u_{i,0}'\dot{v}_{t,0}^{(1)} - u_{i,1}'\dot{v}_{t,1}^{(1)}\hat{e}_{1,it}\right). \tag{A.21}$$

where $\tilde{Y}_{it} = Y_{it} - \hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}$. Let $\hat{\Delta}_{i,u} = \begin{bmatrix}\hat{u}_{i,0}^{(3,1)} - O_0^{(1)}u_{i,0}^0\\ \hat{u}_{i,1}^{(3,1)} - O_1^{(1)}u_{i,1}^0\end{bmatrix}$ and $\dot{\omega}_{it} = \begin{bmatrix}\dot{v}_{t,0}^{(1)}\\ \dot{v}_{t,1}^{(1)}\hat{e}_{1,it}\end{bmatrix}$. With generic $(u_{i,0}, u_{i,1}, \mathrm{u}_{i,1})$, define

$$\hat{\mathbb{H}}_i\left(u_{i,0}, u_{i,1}, \mathrm{u}_{i,1}\right) = \frac{1}{T}\sum_{t=1}^{T}\left[\tau - \mathbf{1}\left\{\epsilon_{it}\le \iota_{it}\left(u_{i,0}, u_{i,1}, \mathrm{u}_{i,1}\right),\right\}\right]\dot{\omega}_{it} \text{ and}$$

$$\iota_{it}\left(u_{i,0}, u_{i,1}, \mathrm{u}_{i,1}\right) = u_{i,0}'\dot{v}_{t,0}^{(1)} - u_{i,0}^{0\prime}v_{t,0}^{0} + \hat{\mu}_{1,it}\mathrm{u}_{i,1}'\dot{v}_{t,1}^{(1)} - \mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0} + \hat{e}_{1,it}u_{i,1}'\dot{v}_{t,1}^{(1)} - e_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}. \tag{A.22}$$

We can see that

$$\hat{\mathbb{H}}_i\left(u_{i,0}, u_{i,1}, \dot{u}_{i,1}^{(1)}\right) = \frac{1}{T}\sum_{t=1}^{T}\left[\tau - \mathbf{1}\left\{\tilde{Y}_{it} - u_{i,0}'\dot{v}_{t,0}^{(1)} - u_{i,1}'\dot{v}_{t,1}^{(1)}\hat{e}_{1,it} \leq 0\right\}\right]\dot{\omega}_{it}$$

is the first order subgradient of (A.21). In addition, we define

$$\begin{aligned}\hat{\mathcal{H}}_i\left(u_{i,0}, u_{i,1}, \mathrm{u}_{i,1}\right) &= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left\{\left[\tau - \mathbf{1}\left\{\epsilon_{it} \leq \left(\iota_{it}\left(u_{i,0}, u_{i,1}, \mathrm{u}_{i,1}\right)\right)\right\}\right]\dot{\omega}_{it}\middle|\mathscr{D}_e^{I_1\cup I_2}\right\} \\ &= \frac{1}{T}\sum_{t=1}^{T}\left[\tau - F_{it}\left(\iota_{it}\left(u_{i,0}, u_{i,1}, \mathrm{u}_{i,1}\right)\right)\right]\dot{\omega}_{it},\end{aligned}$$

and

$$\begin{aligned}\hat{\mathbb{W}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right) &= \hat{\mathbb{H}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right) - \hat{\mathbb{H}}_i\left(O_0^{(1)}u_{i,0}^{0}, O_1^{(1)}u_{i,1}^{0}, \dot{u}_{i,1}^{(1)}\right) \\ &\quad - \left\{\hat{\mathcal{H}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right) - \hat{\mathcal{H}}_i\left(O_0^{(1)}u_{i,0}^{0}, O_1^{(1)}u_{i,1}^{0}, \dot{u}_{i,1}^{(1)}\right)\right\},\end{aligned}$$

where $\mathscr{D}_e^{I_1\cup I_2}$ is the σ-field generated by $\{\epsilon_{it}\}_{i\in I_1\cup I_2, t\in[T]}\bigcup\{e_{it}\}_{i\in[N], t\in[T]}\bigcup\left\{V_j^0\right\}_{j\in[p]\cup\{0\}}\bigcup\left\{W_j^0\right\}_{j\in[p]}$. Then it is clear that

$$\begin{aligned}\hat{\mathbb{H}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right) = \hat{\mathbb{H}}_i\left(O_0^{(1)}u_{i,0}^{0}, O_1^{(1)}u_{i,1}^{0}, \dot{u}_{i,1}^{(1)}\right)) + \hat{\mathcal{H}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right) - \hat{\mathcal{H}}_i\left(O_0^{(1)}u_{i,0}^{0}, O_1^{(1)}u_{i,1}^{0}, \dot{u}_{i,1}^{(1)}\right) \\ + \hat{\mathbb{W}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right).\end{aligned} \tag{A.23}$$

For specific $u_{i,0}$ and $u_{i,1}$, let $u_i = \left(u_{i,0}', u_{i,1}'\right)'$. Following similar arguments as used in the proof of Lemma B.17, the second order partial derivative of the function $\hat{\mathcal{H}}_i(\cdot)$ with respect to u_i at the true value can be shown to be bounded in probability. By Taylor expansion, it yields

$$\hat{\mathcal{H}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right) - \hat{\mathcal{H}}_i\left(O_0^{(1)}u_{i,0}^{0}, O_1^{(1)}u_{i,1}^{0}, \dot{u}_{i,1}^{(1)}\right) = \frac{\partial\hat{\mathcal{H}}_i\left(O_0^{(1)}u_{i,0}^{0}, O_1^{(1)}u_{i,1}^{0}, \dot{u}_{i,1}^{(1)}\right)}{\partial u_i'}\hat{\Delta}_{i,u} + R_i, \tag{A.24}$$

where $\max_{i\in I_3}|R_i| \lesssim \max_{i\in I_3}\left\|\hat{\Delta}_{i,u}\right\|_2^2$ and

$$\frac{\partial\hat{\mathcal{H}}_i\left(O_0^{(1)}u_{i,0}^{0}, O_1^{(1)}u_{i,1}^{0}, \dot{u}_{i,1}^{(1)}\right)}{\partial u_i'} = -\frac{1}{T}\sum_{t=1}^{T}f_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^{0}, O_1^{(1)}u_{i,1}^{0}, \dot{u}_{i,1}^{(1)}\right)\right]\begin{bmatrix}\dot{v}_{t,0}^{(1)}\dot{v}_{t,0}^{(1)\prime} & \hat{e}_{1,it}\dot{v}_{t,0}^{(1)}\dot{v}_{t,1}^{(1)\prime} \\ \hat{e}_{1,it}\dot{v}_{t,1}^{(1)}\dot{v}_{t,0}^{(1)\prime} & \hat{e}_{1,it}^2\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime}\end{bmatrix} = -\dot{D}_i^F.$$

Combing (A.23) and (A.24), we have

$$\begin{aligned}\hat{\Delta}_{i,u} &= \left(\dot{D}_i^F\right)^{-1}\left\{\hat{\mathbb{H}}_i\left(O_0^{(1)}u_{i,0}^{0}, O_1^{(1)}u_{i,1}^{0}, \dot{u}_{i,1}^{(1)}\right) + \hat{\mathbb{W}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right) - \hat{\mathbb{H}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right) + R_i\right\} \\ &= \left(\dot{D}_i^F\right)^{-1}\left\{\hat{\mathbb{H}}_i\left(O_0^{(1)}u_{i,0}^{0}, O_1^{(1)}u_{i,1}^{0}, \dot{u}_{i,1}^{(1)}\right) + \hat{\mathbb{W}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right) + o_p\left((N\vee T)^{-1/2}\right)\right\},\end{aligned} \tag{A.25}$$

uniformly over $i \in I_3$, where the second line is due to the fact that

$$\max_{i\in I_3}\left\|\hat{\mathbb{H}}_i\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)}\right)\right\|_{\max} = O_p\left(\frac{\xi_N}{T}\right) \quad \text{and} \quad \max_{i\in I_3}\left\|\hat{\Delta}_{i,u}\right\|_2^2 = O_p\left(\eta_N^2\right) = o_p\left((N\vee T)^{-1/2}\right),$$

following similar arguments as in (A.9) and the proof of Theorem 3.2.

Next, we analyze the term $\hat{\mathbb{H}}_i\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)$ in (A.24), which can be written as

$$\begin{aligned}
&\hat{\mathbb{H}}_i\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\\
&=\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left(\tau-\mathbf{1}\left\{\epsilon_{it}\le\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right\}\right)\\
&=\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left(\tau-\mathbf{1}\left\{\epsilon_{it}\le 0\right\}\right)+\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left(\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right\}\right)\\
&=\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left[\tau-\mathbf{1}\left\{\epsilon_{it}\le 0\right\}\right)+\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left\{F_{it}(0)-F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]\right\}\\
&+\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left(\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-F_{it}(0)\right]\\
&-\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left\{\mathbf{1}\left\{\epsilon_{it}\le\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right\}-F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]\right\}. \qquad \text{(A.26)}
\end{aligned}$$

For the second term after the last equality, we notice that

$$\begin{aligned}
&\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left\{F_{it}(0)-F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]\right\}\\
&=-\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}f_{it}\left(\tilde{\iota}_{it}\right)\left[(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0)' \quad (\hat{e}_{1,it}\dot{v}_{t,1}^{(1)}-e_{1,it}O_1^{(1)}v_{t,1}^0)'\right]\begin{bmatrix}O_0^{(1)}u_{i,0}^0\\ O_1^{(1)}u_{i,1}^0\end{bmatrix}\\
&-\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}f_{it}\left(\tilde{\iota}_{it}\right)\left(\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}-\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right)\\
&:=\dot{D}_i^J\begin{bmatrix}O_0^{(1)}u_{i,0}^0\\ O_1^{(1)}u_{i,1}^0\end{bmatrix}+\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}f_{it}\left(\tilde{\iota}_{it}\right)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right), \qquad \text{(A.27)}
\end{aligned}$$

where $|\tilde{\iota}_{it}|$ lies between 0 and $\left|\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right|$ and

$$\dot{D}_i^J=\frac{1}{T}\sum_{t=1}^{T}f_{it}(\tilde{\iota}_{it})\begin{bmatrix}\dot{v}_{t,0}^{(1)}\left(O_0^{(1)}v_{t,0}^0-\dot{v}_{t,0}^{(1)}\right)' & \dot{v}_{t,0}^{(1)}\left(e_{1,it}O_1^{(1)}v_{t,1}^0-\hat{e}_{it}\dot{v}_{t,1}^{(1)}\right)'\\ \hat{e}_{1,it}\dot{v}_{t,1}^{(1)}\left(O_0^{(1)}v_{t,0}^0-\dot{v}_{t,0}^{(1)}\right)' & \hat{e}_{1,it}\dot{v}_{t,1}^{(1)}\left(e_{1,it}O_1^{(1)}v_{t,1}^0-\hat{e}_{1,it}\dot{v}_{t,1}^{(1)}\right)'\end{bmatrix}.$$

The first equality above is due to the mean-value theorem and the definition for $\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)$ in (A.22). Inserting (A.27) into (A.26), we obtain that

$$\begin{aligned}
\hat{\mathbb{H}}_i\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)=&\dot{D}_i^J\begin{bmatrix}O_0^{(1)}u_{i,0}^0\\ O_1^{(1)}u_{i,1}^0\end{bmatrix}+\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left[\tau-\mathbf{1}\left\{\epsilon_{it}\le 0\right\}\right)\\
&+\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}f_{it}\left(\tilde{\iota}_{it}\right)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\\
&+\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right\}\right]
\end{aligned}$$

$$- \left(F_{it}(0) - F_{it}\left[\iota_{it}\left(O_0^{(1)} u_{i,0}^0, O_1^{(1)} u_{i,1}^0, \dot{u}_{i,1}^{(1)} \right) \right] \right) \Bigg\}$$

$$:= \dot{D}_i^J \begin{bmatrix} O_0^{(1)} u_{i,0}^0 \\ O_1^{(1)} u_{i,1}^0 \end{bmatrix} + \hat{\mathbb{I}}_{1,i} + \hat{\mathbb{I}}_{2,i} + \hat{\mathbb{I}}_{3,i}. \tag{A.28}$$

Combining (A.25) and (A.28), we obtain that

$$\hat{\Delta}_{i,u} = \left(\dot{D}_i^F \right)^{-1} \left\{ \dot{D}_i^J \begin{bmatrix} O_0^{(1)} u_{i,0}^0 \\ O_1^{(1)} u_{i,1}^0 \end{bmatrix} + \hat{\mathbb{I}}_{1,i} + \hat{\mathbb{I}}_{2,i} + \hat{\mathbb{I}}_{3,i} + \hat{\mathbb{W}}_i \left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \dot{u}_{i,1}^{(1)} \right) + o_p\left((N \vee T)^{-1/2} \right) \right\}, \tag{A.29}$$

where the $o_p\left((N \vee T)^{-1/2} \right)$ term holds uniformly over $i \in I_3$. To prove Theorem 3.3(i), we analyze each term in (A.29) step by step.

Step 1: Uniform Convergence for $\dot{D}_i^F$ and $\dot{D}_i^J$.

Define

$$D_i^F = \frac{1}{T} \sum_{t=1}^{T} f_{it}(0) \begin{bmatrix} O_0^{(1)} v_{t,0}^0 v_{t,0}^{0\prime} O_0^{(1)\prime} & 0 \\ 0 & e_{1,it}^2 O_1^{(1)} v_{t,1}^0 v_{t,1}^{0\prime} O_1^{(1)\prime} \end{bmatrix},$$

$$D_i^J = \frac{1}{T} \sum_{t=1}^{T} f_{it}(0) \begin{bmatrix} \dot{v}_{t,0}^{(1)} \left(O_0^{(1)} v_{t,0}^0 - \dot{v}_{t,0}^{(1)} \right)' & 0 \\ 0 & e_{1,it}^2 O_1^{(1)} v_{t,1}^0 \left(O_1^{(1)} v_{t,1}^0 - \dot{v}_{t,1}^{(1)} \right)' \end{bmatrix}.$$

Lemmas B.22 and B.23 show that

$$\max_{i \in I_3} ||\dot{D}_i^F - D_i^F||_F = O_p(\eta_N) \quad \text{and} \quad \max_{i \in I_3} ||\dot{D}_i^J - D_i^J||_F = \left\| \begin{bmatrix} O_p(\eta_N^2) & O_p(\eta_N) \\ O_p(\eta_N^2) & O_p(\eta_N^2) \end{bmatrix} \right\|_F,$$

with $\eta_N = \frac{\sqrt{\log(N \vee T)}\xi_N^2}{\sqrt{N \wedge T}}$.

Step 2: Uniform Convergence for $\hat{\mathbb{I}}_{1,i}$.

Let $\omega_{it}^0 = \begin{bmatrix} O_0^{(1)} v_{t,0}^0 \\ O_1^{(1)} v_{t,1}^0 e_{1,it} \end{bmatrix}$. Then we can see that

$$\hat{\mathbb{I}}_{1,i} = \frac{1}{T} \sum_{t=1}^{T} \omega_{it}^0 \left(\tau - \mathbf{1}\left\{ \epsilon_{it} \le 0 \right\} \right) + \frac{1}{T} \sum_{t=1}^{T} \left(\dot{\omega}_{it} - \omega_{it}^0 \right) \left(\tau - \mathbf{1}\left\{ \epsilon_{it} \le 0 \right\} \right).$$

Noting that

$$\dot{\omega}_{it} - \omega_{it}^0 = \begin{bmatrix} \dot{v}_{t,0}^{(1)} - O_0^{(1)} v_{t,0}^0 \\ \dot{v}_{t,1}^{(1)} \hat{e}_{1,it} - O_1^{(1)} v_{t,1}^0 e_{1,it}^0 \end{bmatrix}$$

$$= \begin{bmatrix} \dot{v}_{t,0}^{(1)} - O_0^{(1)} v_{t,0}^0 \\ \left(\dot{v}_{t,1}^{(1)} - O_0^{(1)} v_{t,1}^0 \right) \left(\hat{e}_{1,it} - e_{1,it} \right) + e_{1,it} \left(\dot{v}_{t,1}^{(1)} - O_0^{(1)} v_{t,1}^0 \right) + O_1^{(1)} v_{t,1}^{(1)} \left(\hat{e}_{1,it} - e_{1,it} \right) \end{bmatrix}, \tag{A.30}$$

we have

$$\max_{i \in I_3} \frac{1}{T} \sum_{t \in [T]} \left\| \dot{\omega}_{it} - \omega_{it}^0 \right\|_2 = O_p(\eta_N). \tag{A.31}$$

In addition, $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^{0}\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_2=O_p\left(\eta_N^2\right)$ by Lemma B.24. It follows that

$$\hat{\mathbb{I}}_{1,i}=\frac{1}{T}\sum_{t=1}^{T}\omega_{it}^{0}\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)+O_p\left(\eta_N^2\right), \tag{A.32}$$

uniformly over $i\in I_3$.

Step 3: Uniform Convergence for $\hat{\mathbb{I}}_{2,i}$.

Note that

$$\begin{aligned}
\hat{\mathbb{I}}_{2,i}&=\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}f_{it}\left(\tilde{\iota}_{it}\right)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\\
&=\frac{1}{T}\sum_{t=1}^{T}\omega_{it}^{0}f_{it}(0)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\\
&\quad+\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^{0}\right)\left[f_{it}\left(\tilde{\iota}_{it}\right)-f_{it}(0)\right]\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\\
&\quad+\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^{0}\right)f_{it}(0)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\\
&\quad+\frac{1}{T}\sum_{t=1}^{T}\omega_{it}^{0}\left[f_{it}\left(\tilde{\iota}_{it}\right)-f_{it}(0)\right]\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)
\end{aligned}\tag{A.33}$$

where

$$\begin{aligned}
&\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^{0}\right)\left[f_{it}\left(\tilde{\iota}_{it}\right)-f_{it}(0)\right]\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\right\|_2=O_p(\eta_N^3),\\
&\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^{0}\right)f_{it}(0)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\right\|_2=O_p(\eta_N^2),\\
&\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\omega_{it}^{0}\left[f_{it}\left(\tilde{\iota}_{it}\right)-f_{it}(0)\right]\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\right\|_2=O_p(\eta_N^2).
\end{aligned}$$

To see why these three equalities hold, we focus on the third one. By Cauchy's inequality, Theorem 3.2, and Lemma B.21, we have

$$\begin{aligned}
&\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\omega_{it}^{0}\left[f_{it}\left(\tilde{\iota}_{it}\right)-f_{it}(0)\right]\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\right\|_2\\
&\lesssim\sqrt{\max_{i\in I_3}\frac{1}{T}\sum_{t\in[T]}\left\|\omega_{it}^{0}\right\|_2^2\left|\tilde{\iota}_{it}\right|^2}\sqrt{\max_{i\in I_3}\frac{1}{T}\sum_{t\in[T]}\left|\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right|^2}=O_p\left(\eta_N^2\right).
\end{aligned}$$

For the first term on the right hand side (RHS) of the second equality of (A.33), we have by Lemma B.25

$$\begin{aligned}
&\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}O_0^{(1)}v_{t,0}^{0}f_{it}(0)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\right\|_2=O_p(\eta_N),\\
&\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}e_{1,it}O_1^{(1)}v_{t,1}^{0}f_{it}(0)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)\right\|_2=O_p\left(\eta_N^2\right),
\end{aligned}$$

and thus $\max_{i\in I_3}\left\|\hat{\mathbb{I}}_{2,i}\right\|_2 = \left\|\begin{bmatrix} O_p(\eta_N) \\ O_p(\eta_N^2)\end{bmatrix}\right\|_2$.

Step 4: Uniform Convergence for $\hat{\mathbb{I}}_{3,i}$.

Note that

$$\begin{aligned}
\hat{\mathbb{I}}_{3,i} &= \frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le \iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right\}\right] \\
&\quad -\left(F_{it}(0)-F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]\right)\Bigg\} \\
&= \frac{1}{T}\sum_{t=1}^{T}\omega_{it}^0\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le \iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right\}\right] \\
&\quad -\left(F_{it}(0)-F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]\right)\Bigg\} \\
&\quad +\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^0\right)\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le \iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right\}\right] \\
&\quad -\left(F_{it}(0)-F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]\right)\Bigg\}. \qquad \text{(A.34)}
\end{aligned}$$

By (B.73), we can show that

$$\begin{aligned}
&\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\le R_{\iota,it}^1\left(|\mu_{1,it}|+|e_{1,it}|\right)+R_{\iota,it}^2 \quad \text{with} \\
&\max_{i\in I_3,t\in[T]}\left|R_{\iota,it}^1\right| = O_p(\eta_N),\ \max_{i\in I_3,t\in[T]}\left|R_{\iota,it}^2\right| = O_p(\eta_N). \qquad \text{(A.35)}
\end{aligned}$$

For the second term on the RHS of the second equality in (A.34), we notice that

$$\begin{aligned}
&\max_{i\in I_3}\Bigg\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^0\right)\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le \iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right\}\right] \\
&-\left(F_{it}(0)-F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]\right)\Bigg\}\Bigg\|_2 \\
&\le \max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|\dot{\omega}_{it}-\omega_{it}^0\right\|_2\mathbf{1}\left\{0\le|\epsilon_{it}|\le\left|\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right|\right\} \\
&+\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|\dot{\omega}_{it}-\omega_{it}^0\right\|_2\left|F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]-F_{it}(0)\right| \\
&\le \max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|\dot{\omega}_{it}-\omega_{it}^0\right\|_2\mathbf{1}\left\{0\le|\epsilon_{it}|\le\left|\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right|\right\}+O_p(\eta_N^2), \qquad \text{(A.36)}
\end{aligned}$$

where the last line is by (A.30), (A.35) and Assumption 1(iv).

Define the event $\mathscr{A}_9(M) := \left\{\max_{i\in I_3,t\in[T]}\left|\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right|\le M\eta_N\left(|\mu_{1,it}|+|e_{1,it}|+1\right)\right\}$ with $\mathbb{P}\left\{\mathscr{A}_9^c(M)\right\}\le e$ for any $e>0$. Then for a large enough constant C_{17}, we have

$$\mathbb{P}\left(\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|\dot{\omega}_{it}-\omega_{it}^0\right\|_2\mathbf{1}\left\{0\le|\epsilon_{it}|\le\left|\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right|\right\}>C_{17}\eta_N^2\right)$$

$$
\begin{aligned}
&\le \mathbb{P}\left(\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|\dot{\omega}_{it}-\omega_{it}^0\right\|_2 \mathbf{1}\left\{0\le|\epsilon_{it}|\le\left|\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0,\dot{u}_{i,1}^{(1)}\right)\right|\right\}>C_{17}\eta_N^2\middle|\mathscr{A}_9(M)\right)+e\\
&\le \mathbb{P}\left(\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|\dot{\omega}_{it}-\omega_{it}^0\right\|_2 \mathbf{1}\left\{0\le|\epsilon_{it}|\le M\eta_N\left(|\mu_{1,it}|+|e_{1,it}|+1\right)\right\}>C_{17}\eta_N^2\right)+e\\
&\le \mathbb{E}\left\{\mathbb{P}\left(\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|\dot{\omega}_{it}-\omega_{it}^0\right\|_2\left[\mathbf{1}_{it}-\bar{\mathbf{1}}_{it}\right]>\frac{C_{17}\eta_N^2}{2}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\right\}\\
&+\mathbb{E}\left\{\mathbb{P}\left(\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left\{\left\|\dot{\omega}_{it}-\omega_{it}^0\right\|_2 \mathbf{1}\left\{0\le|\epsilon_{it}|\le M\eta_N\left(|\mu_{1,it}|+|e_{1,it}|+1\right)\right\}\middle|\mathscr{D}_e^{I_1\cup I_2}\right\}>\frac{C_{17}\eta_N^2}{2}\right)\right\}+e\\
&=o(1)+e
\end{aligned}
\tag{A.37}
$$

where $\mathbf{1}_{it}-\bar{\mathbf{1}}_{it}:=\mathbf{1}\left\{0\le|\epsilon_{it}|\le M\eta_N\left(|\mu_{1,it}|+|e_{1,it}|+1\right)\right\}-\mathbb{E}\left(\mathbf{1}\left\{0\le|\epsilon_{it}|\le M\eta_N\left(|\mu_{1,it}|+|e_{1,it}|+1\right)\right\}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)$, the last line holds by the fact that $\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left\{\left\|\dot{\omega}_{it}-\omega_{it}^0\right\|_2\mathbf{1}\left\{0\le|\epsilon_{it}|\le M\eta_N\left(|\mu_{1,it}|+|e_{1,it}|+1\right)\right\}\middle|\mathscr{D}_e^{I_1\cup I_2}\right\}$ $=O_p(\eta_N^2)$ and

$$
\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|\dot{\omega}_{it}-\omega_{it}^0\right\|_2\left[\mathbf{1}_{it}-\bar{\mathbf{1}}_{it}\right]>C_{17}\eta_N^2\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\\
&\lesssim \exp\left(-\frac{T^2\eta_N^4}{T\xi_N^2\eta_N^2+T\eta_N^3\xi_N\log T\log\log T}\right)=o(1)
\end{aligned}
$$

by Bernstein's inequality in Lemma B.12(i). Combining (A.36) and (A.37), we have shown the second term on the RHS of the second equality in (A.34) is $O_p(\eta_N^2)$ uniformly over $i\in I_3$. This result, in conjunction with Lemma B.25, (A.34) and Assumption 1(ix), implies that $\max_{i\in I_3}\left\|\hat{\mathbb{I}}_{3,i}\right\|=o_p\left((N\vee T)^{-\frac{1}{2}}\right)$.

Step 5: Uniform Convergence for $\hat{\mathbb{W}}_i\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)$.

Note that

$$
\begin{aligned}
&\hat{\mathbb{W}}_i\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\\
&=\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left(\mathbf{1}\left\{\epsilon_{it}\le\iota_{it}\left(O_0^{(1)}u_{i,0}^0,O_1^{(1)}u_{i,1}^0,\dot{u}_{i,1}^{(1)}\right)\right\}-\mathbf{1}\left\{\epsilon_{it}\le\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right\}\right)\\
&-\frac{1}{T}\sum_{t=1}^{T}\dot{\omega}_{it}\left(F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0,O_1^{(1)}u_{i,1}^0,\dot{u}_{i,1}^{(1)}\right)\right]-F_{it}\left[\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right]\right)\\
&=\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^0\right)\left(\mathbf{1}\left\{\epsilon_{it}\le\iota_{it}\left(O_0^{(1)}u_{i,0}^0,O_1^{(1)}u_{i,1}^0,\dot{u}_{i,1}^{(1)}\right)\right\}-\mathbf{1}\left\{\epsilon_{it}\le\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right\}\right)\\
&-\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^0\right)\left(F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0,O_1^{(1)}u_{i,1}^0,\dot{u}_{i,1}^{(1)}\right)\right]-F_{it}\left[\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right]\right)\\
&+\frac{1}{T}\sum_{t=1}^{T}\omega_{it}^0\Bigg\{\left(\mathbf{1}\left\{\epsilon_{it}\le\iota_{it}\left(O_0^{(1)}u_{i,0}^0,O_1^{(1)}u_{i,1}^0,\dot{u}_{i,1}^{(1)}\right)\right\}-\mathbf{1}\left\{\epsilon_{it}\le\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right\}\right)\\
&-\left(F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0,O_1^{(1)}u_{i,1}^0,\dot{u}_{i,1}^{(1)}\right)\right]-F_{it}\left[\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right]\right)\Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
&:= \frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^{0}\right)\left(\mathbf{1}\left\{\epsilon_{it}\leq \iota_{it}\left(O_0^{(1)}u_{i,0}^{0},O_1^{(1)}u_{i,1}^{0},\dot{u}_{i,1}^{(1)}\right)\right\}-\mathbf{1}\left\{\epsilon_{it}\leq \iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right\}\right)\\
&-\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^{0}\right)\left(F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^{0},O_1^{(1)}u_{i,1}^{0},\dot{u}_{i,1}^{(1)}\right)\right]-F_{it}\left[\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right]\right)\\
&+\hat{\mathbb{W}}_i^{I}-\hat{\mathbb{W}}_i^{II},
\end{aligned}
$$

where we define

$$
\begin{aligned}
\hat{\mathbb{W}}_i^{I} &= \frac{1}{T}\sum_{t=1}^{T}\omega_{it}^{0}\Bigg\{\mathbf{1}\left\{\epsilon_{it}\leq \iota_{it}\left(O_0^{(1)}u_{i,0}^{0},O_1^{(1)}u_{i,1}^{0},\dot{u}_{i,1}^{(1)}\right)\right\}-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\\
&\quad -\left\{F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^{0},O_1^{(1)}u_{i,1}^{0},\dot{u}_{i,1}^{(1)}\right)\right]-F_{it}(0)\right\}\Bigg\} \quad \text{and}\\
\hat{\mathbb{W}}_i^{II} &= \frac{1}{T}\sum_{t=1}^{T}\omega_{it}^{0}\Bigg\{\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{\epsilon_{it}\leq \iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right\}\\
&\quad -\left\{F_{it}(0)-F_{it}\left[\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)},\right)\right]\right\}\Bigg\}.
\end{aligned}
$$

We first observe that

$$
\begin{aligned}
&\iota_{it}\left(O_0^{(1)}u_{i,0}^{0},O_1^{(1)}u_{i,1}^{0},\dot{u}_{i,1}^{(1)}\right)-\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\\
&=\left(\hat{u}_{i,0}^{(3,1)}-O_0^{(1)}u_{i,0}^{0}\right)'\dot{v}_{t,0}^{(1)}+\left(\hat{u}_{i,1}^{(3,1)}-O_1^{(1)}u_{i,1}^{0}\right)'\dot{v}_{t,1}^{(1)}\hat{e}_{1,it}\\
&=\left(O_0^{(1)\prime}\hat{u}_{i,0}^{(3,1)}-u_{i,0}^{0}\right)'v_{t,0}^{0}+\left(O_1^{(1)\prime}\hat{u}_{i,1}^{(3,1)}-u_{i,1}^{0}\right)'v_{t,1}^{(1)}e_{1,it}+O_p(\eta_N^2)\\
&=R_{\iota,it}^{3}e_{1,it}+R_{\iota,it}^{4} \qquad \text{(A.38)}
\end{aligned}
$$

such that $\max_{i\in I_3,t\in[T]}\left|R_{\iota,it}^{3}\right|=O_p(\eta_N)$ and $\max_{i\in I_3,t\in[T]}\left|R_{\iota,it}^{4}\right|=O_p(\eta_N)$. As in Step 4, we can show that

$$
\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^{0}\right)\left(\mathbf{1}\left\{\epsilon_{it}\leq \iota_{it}\left(O_0^{(1)}u_{i,0}^{0},O_1^{(1)}u_{i,1}^{0},\dot{u}_{i,1}^{(1)}\right)\right\}-\mathbf{1}\left\{\epsilon_{it}\leq \iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right\}\right)\right\|=O_p(\eta_N^2)
$$

and

$$
\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{\omega}_{it}-\omega_{it}^{0}\right)\left(F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^{0},O_1^{(1)}u_{i,1}^{0},\dot{u}_{i,1}^{(1)}\right)\right]-F_{it}\left[\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right]\right)\right\|=O_p(\eta_N^2).
$$

Then by Lemma B.26, Lemma B.27 and Assumption 1(ix), we obtain that

$$
\max_{i\in I_3}\left\|\hat{\mathbb{W}}_i\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right\|_2=o_p\left((N\vee T)^{-\frac{1}{2}}\right).
$$

Step 6: Distribution Theory for $\hat{\Delta}_{i,u}$

Combining the above results, we have that uniformly over $i\in I_3$,

$$
\begin{bmatrix}\hat{u}_{i,0}^{(3,1)}\\ \hat{u}_{i,1}^{(3,1)}\end{bmatrix}=\begin{bmatrix}O_0^{(1)}u_{i,0}^{0}\\ O_1^{(1)}u_{i,1}^{0}\end{bmatrix}+\left(\dot{D}_i^{F}\right)^{-1}\left\{\dot{D}_i^{J}\begin{bmatrix}O_0^{(1)}u_{i,0}^{0}\\ O_1^{(1)}u_{i,1}^{0}\end{bmatrix}+\hat{\mathbb{I}}_{1,i}+\hat{\mathbb{I}}_{2,i}+\mathbb{I}_{3,i}+\hat{\mathbb{W}}_i\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)+o_p\left((N\vee T)^{-\frac{1}{2}}\right)\right\}
$$

$$= \left[I_{K_0+K_1} + \left(D_i^F\right)^{-1} D_i^J\right] \begin{bmatrix} O_0^{(1)} u_{i,0}^0 \\ O_1^{(1)} u_{i,1}^0 \end{bmatrix} + \left[\left(\dot{D}_i^F\right)^{-1} \dot{D}_i^J - \left(D_i^F\right)^{-1} D_i^J\right] \begin{bmatrix} O_0^{(1)} u_{i,0}^0 \\ O_1^{(1)} u_{i,1}^0 \end{bmatrix}$$

$$+ \left(D_i^F\right)^{-1} \frac{1}{T} \sum_{t=1}^{T} \omega_{it}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it} \le 0\right\}\right) + \left[\left(\dot{D}_i^F\right)^{-1} - \left(D_i^F\right)^{-1}\right] \frac{1}{T} \sum_{t=1}^{T} \omega_{it}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it} \le 0\right\}\right)$$

$$+ \left(D_i^F\right)^{-1} \hat{\mathbb{I}}_{2,i} + o_p\left((N \vee T)^{-\frac{1}{2}}\right). \tag{A.39}$$

Owing to the fact that

$$\max_{i \in I_3} \left\| \frac{1}{T} \sum_{t=1}^{T} \omega_{it}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it} \le 0\right\}\right) \right\|_2 = O_p\left(\sqrt{\frac{\log(N \vee T)}{T}}\right) \tag{A.40}$$

by similar arguments as (B.57) using Bernstein's inequality in Lemma B.12(ii) and Lemma B.22, we notice that

$$\max_{i \in I_3} \left\| \left[\left(\dot{D}_i^F\right)^{-1} - \left(D_i^F\right)^{-1}\right] \frac{1}{T} \sum_{t=1}^{T} \omega_{it}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it} \le 0\right\}\right) \right\|_2 = o_p\left((N \vee T)^{-\frac{1}{2}}\right).$$

Next, we define

$$D^F = O^{(1)} \begin{bmatrix} \frac{1}{T}\sum_{t=1}^T \mathbb{E}\left[f_{it}(0)\middle|\mathscr{D}\right] v_{t,0}^0 v_{t,0}^{0\prime} & 0 \\ 0 & \frac{1}{T}\sum_{t=1}^T \mathbb{E}\left[e_{1,it}^2 f_{it}(0)\middle|\mathscr{D}\right] v_{t,1}^0 v_{t,1}^{0\prime} \end{bmatrix} O^{(1)\prime},$$

$$D^J = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^T \mathbb{E}\left[f_{it}(0)\middle|\mathscr{D}\right] \dot{v}_{t,0}^{(1)} \left(O_0^{(1)} v_{t,0}^0 - \dot{v}_{t,0}^{(1)}\right)' & 0 \\ 0 & O_1^{(1)} \frac{1}{T}\sum_{t=1}^T \mathbb{E}\left[e_{1,it}^2 f_{it}(0)\middle|\mathscr{D}\right] v_{t,1}^0 \left(O_1^{(1)} v_{t,1}^0 - \dot{v}_{t,1}^{(1)}\right)' \end{bmatrix},$$

where $O^{(1)} = diag\left(O_0^{(1)}, O_1^{(1)}\right)$. Here D^F and D^J do not depend on i owing to stationary assumption of sequence $\left\{f_{it}, f_{it}(0) e_{j,it}\right\}_{j \in [p]}$ conditional on all factors in Assumption 9(iii). By Bernstein's inequality conditional on all factors similarly as in (B.57), we can show that $\max_{i \in I_3} \left\|D_i^F - D^F\right\|_F = O_p\left(\sqrt{\frac{\log(N \vee T)}{T}} \xi_N\right)$. Analogously, by Bernstein's inequality conditional on $\mathscr{D}^{I_1 \cup I_2}$, we can show that $\max_{i \in I_3} \left\|D_i^J - D_j\right\|_F = O_p\left(\sqrt{\frac{\log(N \vee T)}{T}} \eta_N \xi_N\right)$. Then it follows that

$$\max_{i \in I_3} \left\| \left(D_i^F\right)^{-1} D_i^J - \left(D^F\right)^{-1} D^J \right\|_F = O_p(\eta_N^2).$$

In addition, uniformly over $i \in I_3$,

$$\left(\dot{D}_i^F\right)^{-1} \dot{D}_i^J - \left(D_i^F\right)^{-1} D_i^J$$

$$= \left[\left(\dot{D}_i^F\right)^{-1} - \left(D_i^F\right)^{-1}\right] \left[\dot{D}_i^J - D_i^J\right] + D_i^J \left[\left(\dot{D}_i^F\right)^{-1} - \left(D_i^F\right)^{-1}\right] + \left(D_i^F\right)^{-1} \left[\dot{D}_i^J - D_i^J\right]$$

$$= \left(D_i^F\right)^{-1} \left[\dot{D}_i^J - D_i^J\right] + O_p(\eta_N^2)$$

$$= \begin{bmatrix} O_p\left(\eta_N^2\right) & O_p(\eta_N) \\ O_p\left(\eta_N^2\right) & O_p\left(\eta_N^2\right) \end{bmatrix},$$

where the upper right block is dominated by $\frac{1}{T}\sum_{t=1}^{T} f_{it}(0) O_0^{(1)} v_{t,0}^0 v_{t,1}^{0\prime} O_1^{(1)\prime}\left(e_{1,it}-\hat{e}_{1,it}\right)$ in the analysis of $J_{1,i}^2$ in (B.62).

Let $I_{K_0+K_1}+\left(D^F\right)^{-1} D^J = O_{u,1}^{(1)} := \begin{pmatrix} \bar{O}_{u,0} & 0 \\ 0 & \bar{O}_{u,1} \end{pmatrix}$, $O_{u,0}^{(1)} = \bar{O}_{u,0} O_0^{(1)}$, and $O_{u,1}^{(1)} = \bar{O}_{u,1} O_1^{(1)}$. Combining the above arguments, we obtain that

$$\hat{u}_{i,1}^{(3,1)} - O_{u,1}^{(1)} u_{i,1}^0 = O_1^{(1)} \hat{V}_{u_1}^{-1} \frac{1}{T}\sum_{t=1}^{T} e_{1,it} v_{t,1}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it} \leq 0\right\}\right) + \mathcal{R}_{i,u}^1, \tag{A.41}$$

$$\begin{aligned} \hat{u}_{i,0}^{(3,1)} - O_{u,0}^{(1)} u_{i,0}^0 &= O_0^{(1)} \hat{V}_{u_0,i}^{-1} \frac{1}{T}\sum_{t=1}^{T} v_{t,0}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it} \leq 0\right\}\right) \\ &\quad + O_0^{(1)} \hat{V}_{u_0,i}^{-1} \frac{1}{T}\sum_{t=1}^{T} f_{it}(0) v_{t,0}^0 v_{t,1}^{0\prime} u_{i,1}^0 \left(e_{1,it} - \hat{e}_{1,it}\right) \\ &\quad + O_0^{(1)} \hat{V}_{u_0,i}^{-1} \frac{1}{T}\sum_{t=1}^{T} f_{it}(0) v_{t,0}^0 \left(\mu_{1,it} u_{i,1}^{0\prime} v_{t,1}^0 - \hat{\mu}_{1,it} \dot{u}_{i,1}^{(1)\prime} \dot{v}_{t,1}^{(1)}\right) + \mathcal{R}_{i,u}^0 \end{aligned} \tag{A.42}$$

such that $\max_{i\in I_3}\left|\mathcal{R}_{i,u}^0\right| = o_p\left((N\vee T)^{-\frac{1}{2}}\right)$, $\max_{i\in I_3}\left|\mathcal{R}_{i,u}^1\right| = o_p\left(\frac{1}{\sqrt{T}}\right)$, $\hat{V}_{u_0} = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[f_{it}(0)\middle|\mathscr{D}\right] v_{t,0}^0 v_{t,0}^{0\prime}$ and $\hat{V}_{u_1} = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[f_{it}(0) e_{1,it}^2\middle|\mathscr{D}\right] v_{t,1}^0 v_{t,1}^{0\prime}$. From (A.41), owing to the fact that $\hat{V}_{u_1}$ is bounded a.s. and $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T} e_{1,it} v_{t,1}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_2 = O_p\left(\sqrt{\frac{\log(N\vee T)}{T}}\right)$ by Bernstein's inequality in Lemma B.12(ii), we obtain that

$$\max_{i\in I_3}\left\|\hat{u}_{i,1}^{(3,1)} - O_{u,1}^{(1)} u_{i,1}^0\right\|_2 = O_p\left(\sqrt{\frac{\log(N\vee T)}{T}}\right) \quad \text{and} \quad \max_{i\in I_3}\left\|\hat{u}_{i,0}^{(3,1)} - O_{u,0}^{(1)} u_{i,0}^0\right\|_2 = O_p(\eta_N). \tag{A.43}$$

Last, noting that $O_1^{(1)}$ is a rotation matrix and the normal distribution is invariant to rotation, for each $i \in I_3$, we have

$$\sqrt{T}\left(\hat{u}_{i,1}^{(3,1)} - O_{u,1}^{(1)} u_{i,1}^0\right) \rightsquigarrow \mathcal{N}(0, \Sigma_{u_1,i}),$$

where $\Sigma_{u_1,i} = O_1^{(1)} V_{u_1}^{-1} \Omega_{u_1} V_{u_1}^{-1} O_1^{(1)\prime}$, $V_{u_1,i} = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left(f_{it}(0) e_{1,it}^2 v_{t,1}^0 v_{t,1}^{0\prime}\right)$, and

$$\Omega_{u_1,i} = Var\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_{1,it} v_{t,1}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right]. \quad \blacksquare$$

A.3.2 Proof of Statement (ii)

Steps for the proof for statement (ii) are the same as those in the proof of statement (i). Hence, we only sketch the proof. Recall from (2.8) that $\forall t \in [T]$,

$$\{\hat{v}_{t,j}^{(3,1)}\}_{j\in[p]} = \underset{\{v_{t,j}\}_{j\in[p]}}{\arg\min}\ \frac{1}{N_3}\sum_{i\in I_3}\rho_\tau\left(\hat{Y}_{it} - v_{t,0}'\hat{u}_{i,0}^{(3,1)} - v_{t,1}'\hat{u}_{i,1}^{(3,1)}\hat{e}_{1,it}\right),$$

where $\hat{Y}_{it} = Y_{it} - \hat{\mu}_{1,it}\hat{u}_{i,1}^{(3,1)\prime}\dot{v}_{t,1}^{(1)}$. Let

$$\hat{\Delta}_{t,v} = \begin{bmatrix} \hat{v}_{t,0}^{(3,1)} - \left(O_{u,0}^{(1)}\right)^{\prime -1} v_{t,0}^0 \\ \hat{v}_{t,1}^{(3,1)} - \left(O_{u,1}^{(1)}\right)^{\prime -1} v_{t,1}^0 \end{bmatrix} \text{ and } \hat{\omega}_{it} = \begin{bmatrix} \hat{u}_{i,0}^{(3,1)} \\ \hat{u}_{i,1}^{(3,1)}\hat{e}_{1,it} \end{bmatrix}.$$

For generic $(u_{i,0}, u_{i,1}, v_{i,0}, v_{i,1})$, define

$$\hat{\mathbb{S}}_t\left(u_{i,0}, u_{i,1}, v_{i,0}, v_{i,1}\right) = \frac{1}{N_3}\sum_{i\in I_3}\left[\tau - \mathbf{1}\left\{\epsilon_{it} \leq \varrho_{it}\left(u_{i,0}, u_{i,1}, v_{i,0}, v_{i,1}\right)\right\}\right]\hat{\omega}_{it},$$

with

$$\varrho_{it}\left(u_{i,0}, u_{i,1}, v_{i,0}, v_{i,1}\right) = u_{i,0}'v_{t,0} - u_{i,0}^{0\prime}v_{t,0}^0 + \hat{\mu}_{1,it}u_{i,1}'\dot{v}_{t,1}^{(1)} - \mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0 + \hat{e}_{1,it}u_{i,1}'v_{t,1} - e_{1,it}u_{i,1}^{0\prime}v_{t,1}^0.$$

We also define

$$\begin{aligned}\hat{\mathcal{S}}_t\left(u_{i,0}, u_{i,1}, v_{i,0}, v_{i,1}\right) &= \frac{1}{N_3}\sum_{i\in I_3}\mathbb{E}\left\{\left[\tau - \mathbf{1}\left\{\epsilon_{it} \leq \left(\varrho_{it}\left(u_{i,0}, u_{i,1}, v_{i,0}, v_{i,1}\right)\right)\right\}\right]\dot{\omega}_{it}\middle|\mathscr{D}_e^{I_1\cup I_2}\right\}\\ &= \frac{1}{N_3}\sum_{i\in I_3}\left[\tau - F_{it}\left(\varrho_{it}\left(u_{i,0}, u_{i,1}, v_{i,0}, v_{i,1}\right)\right)\right]\dot{\omega}_{it}\end{aligned}$$

and

$$\begin{aligned}\hat{\mathbb{M}}_t\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \hat{v}_{t,0}^{(3,1)}, \hat{v}_{t,1}^{(3,1)}\right) &= \hat{\mathbb{S}}_t\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \hat{v}_{t,0}^{(3,1)}, \hat{v}_{t,1}^{(3,1)}\right) - \hat{\mathbb{S}}_t\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)^{\prime-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)^{\prime-1}v_{t,1}^0\right)\\ &- \left\{\hat{\mathcal{S}}_t\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \hat{v}_{t,0}^{(3,1)}, \hat{v}_{t,1}^{(3,1)}\right) - \hat{\mathcal{S}}_t\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)^{\prime-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)^{\prime-1}v_{t,1}^0\right)\right\}.\end{aligned}$$

Then a similar result to that in (A.29) holds:

$$\hat{\Delta}_{t,v} = \left(\hat{D}_t^F\right)^{-1}\left\{\hat{D}_t^J\begin{bmatrix}\left(O_{u,0}^{(1)}\right)^{\prime-1}v_{t,0}^0\\ \left(O_{u,1}^{(1)}\right)^{\prime-1}v_{t,1}^0\end{bmatrix} + \hat{\mathbb{I}}_{4,t} + \hat{\mathbb{I}}_{5,t} + \hat{\mathbb{I}}_{6,t} + \hat{\mathbb{M}}_t\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \hat{v}_{t,0}^{(1)}, \hat{v}_{t,1}^{(3,1)}\right) + o_p\left((N\vee T)^{-\frac{1}{2}}\right)\right\}, \tag{A.44}$$

where

$$\begin{aligned}\hat{D}_t^F &= \frac{1}{N_3}\sum_{i\in I_3}f_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)^{\prime-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)^{\prime-1}v_{t,1}^0\right)\right]\begin{bmatrix}\hat{u}_{i,0}^{(3,1)}\hat{u}_{i,0}^{(3,1)\prime} & \hat{e}_{1,it}\hat{u}_{i,0}^{(3,1)}\hat{u}_{i,1}^{(3,1)\prime}\\ \hat{e}_{1,it}\hat{u}_{i,1}^{(3,1)}\hat{u}_{i,0}^{(3,1)\prime} & \hat{e}_{1,it}^2\hat{u}_{i,1}^{(3,1)}\hat{u}_{i,1}^{(3,1)\prime}\end{bmatrix},\\ \hat{D}_t^J &= \frac{1}{N_3}\sum_{i\in I_3}f_{it}(\tilde{\varrho}_{it})\begin{bmatrix}\hat{u}_{i,0}^{(3,1)}\left(O_{u,0}^{(1)}u_{i,0}^0 - \hat{u}_{i,0}^{(3,1)}\right)' & \hat{u}_{i,0}^{(3,1)}\left(e_{1,it}O_{u,1}^{(1)}u_{i,1}^0 - \hat{e}_{it}\hat{u}_{i,1}^{(3,1)}\right)'\\ \hat{e}_{1,it}\hat{u}_{i,1}^{(3,1)}\left(O_{u,0}^{(1)}u_{i,0}^0 - \hat{u}_{i,0}^{(3,1)}\right)' & \hat{e}_{1,it}\hat{u}_{i,1}^{(3,1)}\left(e_{1,it}O_{u,1}^{(1)}u_{i,1}^0 - \hat{e}_{1,it}\hat{u}_{i,1}^{(3,1)}\right)'\end{bmatrix},\\ \hat{\mathbb{I}}_{4,t} &= \frac{1}{N_3}\sum_{i\in I_3}\hat{\omega}_{it}\left[\tau - \mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right),\\ \hat{\mathbb{I}}_{5,t} &= \frac{1}{N_3}\sum_{i\in I_3}\hat{\omega}_{it}f_{it}\left(\tilde{\varrho}_{it}\right)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0 - \hat{\mu}_{1,it}\hat{u}_{i,1}^{(3,1)\prime}\dot{v}_{t,1}^{(1)}\right),\\ \hat{\mathbb{I}}_{6,t} &= \frac{1}{N_3}\sum_{i\in I_3}\hat{\omega}_{it}\left\{\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\} - \mathbf{1}\left\{\epsilon_{it}\leq\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)^{\prime-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)^{\prime-1}v_{t,1}^0\right)\right\}\right]\right.\\ &\left.- \left(f_{it}(0) - f_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)^{\prime-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)^{\prime-1}v_{t,1}^0\right)\right]\right)\right\},\end{aligned}$$

and $|\tilde{\varrho}_{it}|$ lies between 0 and $\left|\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)^{\prime-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)^{\prime-1}v_{t,1}^0\right)\right|$.

Then, we derive the linear expansion for $\hat{\Delta}_{t,v}$ by analyzing each term in (A.44). Define

$$D_t^F = \frac{1}{N_3}\sum_{i\in I_3} f_{it}(0)\begin{bmatrix} O_{u,0}^{(1)}u_{i,0}^0 u_{i,0}^{0\prime}O_{u,0}^{(1)\prime} & 0 \\ 0 & e_{1,it}^2 O_{u,1}^{(1)}u_{i,1}^0 u_{i,1}^{0\prime}O_{u,1}^{(1)\prime}\end{bmatrix},$$

$$D_t^J = \frac{1}{N_3}\sum_{i\in I_3} f_{it}(0)\begin{bmatrix} \hat{u}_{i,0}^{(3,1)}\left(O_{u,0}^{(1)}u_{i,0}^0 - \hat{u}_{i,0}^{(3,1)}\right)' & 0 \\ 0 & e_{1,it}^2 O_{u,1}^{(1)}u_{i,1}^0\left(O_{u,1}^{(1)}u_{i,1}^0 - \hat{u}_{i,1}^{(3,1)}\right)'\end{bmatrix},$$

such that

$$\max_{t\in[T]}\left\|\hat{D}_t^F - D_t^F\right\|_F = O_p(\eta_N) \text{ and } \max_{t\in[T]}\left\|\hat{D}_t^J - D_t^J\right\|_F = \left\|\begin{bmatrix} O_p\left(\eta_N^2\right) & O_p(\eta_N) \\ O_p\left(\eta_N^2\right) & O_p\left(\eta_N^2\right)\end{bmatrix}\right\|_F$$

by Lemma B.28. Let $\omega_{it}^* = \begin{bmatrix} O_{u,0}^{(1)}u_{i,0}^0 \\ e_{1,it}O_{u,1}^{(1)}u_{i,1}^0\end{bmatrix}$. We can show that

$$\hat{\mathbb{I}}_{4,t} = \frac{1}{N_3}\sum_{i\in I_3}\omega_{it}^*\left[\tau - \mathbf{1}\left\{\epsilon_{it}\le 0\right\}\right] + O_p\left(\eta_N^2\right)$$

uniformly over $t\in[T]$ by analogous analysis as in Step 2 in the previous subsection. With Lemma B.29 and by similar arguments as in Step 3 in the previous subsection, we obtain

$$\max_{t\in[T]}\left\|\hat{\mathbb{I}}_{5,t}\right\|_2 = \left\|\begin{bmatrix} O_p(\eta_N) \\ o_p\left((N\vee T)^{-\frac{1}{2}}\right)\end{bmatrix}\right\|_2$$

uniformly over $t\in[T]$.

Next, for $\hat{\mathbb{I}}_{6,t}$, From (B.87), we first note that

$$\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right) \le R_{\varrho,it}^1\left(|\mu_{1,it}| + |e_{1,it}|\right) + R_{\varrho,it}^2 \quad \text{with}$$
$$\max_{i\in I_3,t\in[T]}\left|R_{\varrho,it}^1\right| = O_p(\eta_N), \max_{i\in I_3,t\in[T]}\left|R_{\varrho,it}^2\right| = O_p(\eta_N). \tag{A.45}$$

We then observe that, uniformly over $t\in[T]$,

$$\begin{aligned}
\hat{\mathbb{I}}_{6,t} &= \frac{1}{N_3}\sum_{i\in I_3}\hat{\omega}_{it}\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\} - \mathbf{1}\left\{\epsilon_{it}\le \varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right\}\right] \\
&- \left(F_{it}(0) - F_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right]\right)\Bigg\} \\
&= \frac{1}{N_3}\sum_{i\in I_3}\omega_{it}^*\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\} - \mathbf{1}\left\{\epsilon_{it}\le \varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right\}\right] \\
&- \left(F_{it}(0) - F_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right]\right)\Bigg\} \\
&+ \frac{1}{N_3}\sum_{i\in I_3}\left(\hat{\omega}_{it} - \omega_{it}^*\right)\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\} - \mathbf{1}\left\{\epsilon_{it}\le \varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right\}\right] \\
&- \left(F_{it}(0) - F_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0, \left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right]\right)\Bigg\}
\end{aligned}$$

$$= \frac{1}{N_3} \sum_{i \in I_3} \omega_{it}^* \Bigg\{ \left[\mathbf{1}\left\{\epsilon_{it} \le 0\right\} - \mathbf{1}\left\{ \epsilon_{it} \le \varrho_{it} \left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)^{\prime -1} v_{t,0}^0, \left(O_{u,1}^{(1)}\right)^{\prime -1} v_{t,1}^0 \right) \right\} \right]$$

$$- \left(F_{it}(0) - F_{it} \left[\varrho_{it} \left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)^{\prime -1} v_{t,0}^0, \left(O_{u,1}^{(1)}\right)^{\prime -1} v_{t,1}^0 \right) \right] \right) \Bigg\} + o_p\left((N \vee T)^{-\frac{1}{2}} \right)$$

$$= \begin{bmatrix} O_p(\eta_N) \\ o_p\left((N \vee T)^{-\frac{1}{2}} \right) \end{bmatrix},$$

where the third equality holds by (A.45) and similar arguments as used in (A.36) and (A.37), and the last equality holds by Lemma B.30.

Similarly, combining Lemma B.30 and Lemma B.31, it yields

$$\max_{t \in [T]} \left\| \hat{\mathbb{M}}_t \left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \hat{v}_{t,0}^{(3,1)}, \hat{v}_{t,1}^{(3,1)} \right) \right\|_2 = \left\| \begin{bmatrix} O_p(\eta_N) \\ o_p\left((N \vee T)^{-\frac{1}{2}} \right) \end{bmatrix} \right\|_2.$$

Let $\hat{V}_{v_1,t}^3 = \frac{1}{N_3} \sum_{i \in I_3} f_{it}(0) e_{1,it}^2 u_{i,1}^0 u_{i,1}^{0\prime}$, and

$$O_{v_1,t}^{(1)} = \left\{ I_{K_1} + \left(O_{u,1}^{(1)\prime}\right)^{-1} \left[\hat{V}_{v_1,t}^3\right]^{-1} \left[\frac{1}{N_3} \sum_{i \in I_3} f_{it}(0) e_{1,it}^2 u_{i,1}^0 \left(O_{u,1}^{(1)} u_{i,1}^0 - \hat{u}_{i,1}^{(3,1)} \right)^{\prime} \right] \right\} \left(O_{u,1}^{(1)\prime}\right)^{-1}.$$

Combining arguments above from (A.44), we have

$$\hat{v}_{t,1}^{(3,1)} - O_{v_1,t}^{(1)} v_{t,1}^0 = \left(O_{u,1}^{(1)\prime}\right)^{-1} \left(\hat{V}_{v_1,t}^3\right)^{-1} \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it} \le 0\right\} \right) + \mathcal{R}_{t,v}^1,$$

such that $\max_{t \in [T]} \left| \mathcal{R}_{t,v}^1 \right| = o_p\left((N \vee T)^{-\frac{1}{2}} \right)$. This, in conjunction with the result in Lemma B.32, i.e., $\max_{t \in [T]} \left\| O_{v_1,t}^{(1)} - \left(O_{u,1}^{(1)\prime}\right)^{-1} \right\|_F = o_p\left((N \vee T)^{-\frac{1}{2}} \right)$, implies that

$$\begin{aligned} \hat{v}_{t,1}^{(3,1)} - \left(O_{u,1}^{(1)}\right)^{\prime -1} v_{t,1}^0 &= \left(O_{u,1}^{(1)\prime}\right)^{-1} \left(\hat{V}_{v_1,t}^3\right)^{-1} \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it} \le 0\right\} \right) + \mathcal{R}_{t,v}^1, \\ &= O_1^{(1)} \left(\hat{V}_{v_1,t}^3\right)^{-1} \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it} \le 0\right\} \right) + \mathcal{R}_{t,v}^1. \end{aligned} \tag{A.46}$$

where the second line holds by the fact that $\left\| O_{u,1}^{(1)} - O_1^{(1)} \right\|_F = O_p(\eta_N)$, $\hat{V}_{v_1,t}^I$ is uniformly bounded and that

$$\max_{t \in [T]} \left\| \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \left(\tau - \mathbf{1}\left\{\epsilon_{it} \le 0\right\} \right) \right\|_2 = O_p\left(\sqrt{\frac{\log(N \vee T)}{N}} \right)$$

by similar arguments as in (A.40). Further define

$$\Sigma_{v_1} = O_1^{(1)} V_{v_1}^{-1} \Omega_{u_1} V_{u_1}^{-1} O_1^{(1)\prime}, \quad V_{v_1} = \frac{1}{N_3} \sum_{i \in I_3} \mathbb{E}\left(f_{it}(0) e_{1,it}^2 u_{i,1}^0 u_{i,1}^{0\prime} \right), \text{ and}$$

$$\Omega_{v_1} = \tau\left(1 - \tau\right) \frac{1}{N_3} \sum_{i \in I_3} \mathbb{E}\left(e_{1,it}^2 u_{i,1}^0 u_{i,1}^{0\prime} \right).$$

Then we have

$$\sqrt{N_3} \left(\hat{v}_{t,1}^{(3,1)} - \left(O_{u,1}^{(1)}\right)^{\prime -1} v_{t,1}^0 \right) \rightsquigarrow \mathcal{N}(0, \Sigma_{v_1}),$$

$$\max_{t\in[T]}\left\|\hat{v}_{t,1}^{(3,1)}-O_{u,1}^{(1)}v_{t,1}^{0}\right\|_2=O_p\left(\sqrt{\frac{\log(N\vee T)}{N}}\right),\quad \text{and } \max_{t\in[T]}\left\|\hat{v}_{t,0}^{(3,1)}-O_{u,0}^{(1)}v_{t,0}^{0}\right\|_2=O_p(\eta_N), \tag{A.47}$$

where the second line holds by Bernstein's inequality with independent data and is similar to (A.43). ■

A.4 Proof of Proposition 3.4

A.4.1 Proof of Statement (i)

Focusing on the slope estimators for $i \in I_3$, we notice that $\hat{\Theta}_{j,it} = \frac{1}{2}\left\{\hat{u}_{t,j}^{(3,1)\prime}\hat{v}_{t,j}^{(3,1)} + \hat{u}_{t,j}^{(3,2)\prime}\hat{v}_{t,j}^{(3,2)}\right\}$. It follows that

$$\begin{aligned}
&\hat{\Theta}_{j,it}-\Theta_{j,it}^{0}\\
&=\frac{1}{2}\left\{\left(\hat{u}_{i,j}^{(3,1)}-O_{u,j}^{(1)}u_{i,j}^{0}\right)'\left(\hat{v}_{t,j}^{(3,1)}-O_{u,j}^{(1)}v_{t,j}^{0}\right)+\left(O_{u,j}^{(1)}u_{i,j}^{0}\right)'\left(\hat{v}_{t,j}^{(3,1)}-O_{u,j}^{(1)}v_{t,j}^{0}\right)+\left(\hat{u}_{i,j}^{(3,1)}-O_{u,j}^{(1)}u_{i,j}^{0}\right)'O_{u,j}^{(1)}v_{t,j}^{0}\right\}\\
&+\frac{1}{2}\left\{\left(\hat{u}_{i,j}^{(3,2)}-O_{u,j}^{(2)}u_{i,j}^{0}\right)'\left(\hat{v}_{t,j}^{(3,2)}-O_{u,j}^{(2)}v_{t,j}^{0}\right)+\left(O_{u,j}^{(2)}u_{i,j}^{0}\right)'\left(\hat{v}_{t,j}^{(3,2)}-O_{u,j}^{(2)}v_{t,j}^{0}\right)+\left(\hat{u}_{i,j}^{(3,2)}-O_{u,j}^{(2)}u_{i,j}^{0}\right)'O_{u,j}^{(1)}v_{t,j}^{0}\right\}\\
&=u_{i,j}^{0\prime}\left(\hat{V}_{v_j,t}^{(3)}\right)^{-1}\frac{1}{N_3}\sum_{i\in I_3}e_{j,it}u_{i,j}^{0}\left(\tau-\mathbf{1}\left\{\epsilon_{it}\le 0\right\}\right)+v_{t,j}^{0\prime}\hat{V}_{u_j}^{-1}\frac{1}{T}\sum_{t=1}^{T}e_{j,it}v_{t,j}^{0}\left(\tau-\mathbf{1}\left\{\epsilon_{it}\le 0\right\}\right)+\mathcal{R}_{it}^{j}\\
&=u_{i,j}^{0\prime}\left(\hat{V}_{v_j,t}^{(3)}\right)^{-1}\frac{1}{N_3}\sum_{i\in I_3}\xi_{j,it}^{0}+v_{t,j}^{0\prime}\hat{V}_{u_j}^{-1}\frac{1}{T}\sum_{t=1}^{T}b_{j,it}^{0}+\mathcal{R}_{it}^{j},
\end{aligned}$$

such that $\max_{i\in I_3,t\in[T]}\left|\mathcal{R}_{it}^{j}\right| = o_p\left((N\vee T)^{-1/2}\right)$ by Theorem 3.3 and the second equality above combines Theorem 3.3 and the fact that $\left\|O_{u,1}^{(1)}-O_{1}^{(1)}\right\|_F = O_p(\eta_N)$. With similar results hold for slope estimators for subsamples I_1 and I_2, and then we obtain the statement (i).

A.4.2 Proof of Statement (ii)

Combining (A.43), (A.47) and Lemma B.13(i), it's clear that

$$\max_{i\in I_3,t\in[T]}\left|\hat{\Theta}_{j,it}-\Theta_{j,it}^{0}\right|=O_p\left(\sqrt{\frac{\log(N\vee T)}{N\wedge T}}\right)\ \forall j\in[p]\quad\text{and}\quad\max_{i\in I_3,t\in[T]}\left|\hat{\Theta}_{0,it}-\Theta_{0,it}^{0}\right|=O_p(\eta_N). \tag{A.48}$$

A.4.3 Proof of Statement (iii)

For $i \in I_a$ and $a \in [3]$, with the distribution theory defined in Theorem 3.3, we notice that

$$\left(\frac{1}{T}v_{t,j}^{0\prime}\Xi_{u_j,i}v_{t,j}^{0}+\frac{1}{N_a}u_{i,j}^{0\prime}\Xi_{v_j}^{a}u_{i,j}^{0}\right)^{-1/2}\left(\hat{\Theta}_{j,it}-\Theta_{j,it}^{0}\right)\rightsquigarrow\mathcal{N}(0,1),$$

which leads to the proof. ■

A.5 Proof of Theorem 4.1

A.5.1 Proof of Statement (i)

The proof is analogous to that in Castagnetti et al. (2015) and Lu and Su (2021). Recall that $S_{u_j} = \max(S_{u_j}^{(1,2)}, S_{u_j}^{(2,3)}, S_{u_j}^{(3,1)})$. For $a \in [3]$ and $b \in [3] \setminus \{a\}$, Theorem 3.3 shows that

$$\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)} u_{i,j}^0 = O_j^{(b)} \hat{V}_{u_j}^{-1} \frac{1}{T} \sum_{t=1}^{T} b_{j,it}^0 + \mathcal{R}_{i,u}^j \; \forall i \in I_a,$$

where $\max_{i \in I_a} \left| \mathcal{R}_{i,u}^j \right| = o_p\left((N \vee T)^{-1/2} \right)$. Recall that $\hat{\bar{u}}_j^{(a)} = \frac{1}{N_a} \sum_{i \in I_a} \hat{u}_{i,j}^{(a,b)}$. Under $H_0^I : u_{i,j}^0 = u_j, \forall i \in [N]$, we have

$$\hat{\bar{u}}_j^{(a,b)} - O_{u,j}^{(b)} u_j = O_j^{(b)} \hat{V}_{u_j}^{-1} \frac{1}{N_a T} \sum_{i \in I_a} \sum_{t=1}^{T} b_{j,it}^0 + \frac{1}{N_a} \sum_{i \in I_a} \mathcal{R}_{i,u}^j = o_p\left((N \vee T)^{-1/2} \right), \tag{A.49}$$

where the last equality holds by a simple application of Bernstein's inequality. Note that

$$\begin{aligned}
&T \left(\hat{u}_{i,j}^{(a,b)} - \hat{\bar{u}}_j^{(a,b)} \right)' \left(\hat{\Sigma}_{u_j} \right)^{-1} \left(\hat{u}_{i,j}^{(a,b)} - \hat{\bar{u}}_j^{(a,b)} \right) \\
&= T \left(\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)} u_j \right)' \left(\hat{\Sigma}_{u_j} \right)^{-1} \left(\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)} u_j \right) + T \left(\hat{\bar{u}}_j^{(a,b)} - O_{u,j}^{(b)} u_j \right)' \left(\hat{\Sigma}_{u_j} \right)^{-1} \left(\hat{\bar{u}}_j^{(a,b)} - O_{u,j}^{(b)} u_j \right) \\
&- 2T \left(\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)} u_j \right)' \left(\hat{\Sigma}_{u_j} \right)^{-1} \left(\hat{\bar{u}}_j^{(a,b)} - O_{u,j}^{(b)} u_j \right) := I_{1ij} + I_{2ij} - 2 I_{3ij}.
\end{aligned} \tag{A.50}$$

For I_{2ij}, we have

$$\max_{i \in I_a} |I_{2ij}| \le \max_{i \in I_a} T \left\| \hat{\bar{u}}_j^{(a,b)} - O_{u,j}^{(b)} u_j \right\|_2^2 \left\{ \lambda_{\min}(\Sigma_{u_j}) + o_p(1) \right\}^{-1} = o_p(1) \tag{A.51}$$

by Lemma B.33, Assumption 11 and (A.49). For I_{3ij}, we have

$$\begin{aligned}
\max_{i \in I_a} |I_{3ij}| &\le T \left(\max_{i \in I_a} \left\| \hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)} u_j \right\|_2 \right) \left(\left\| \hat{\bar{u}}_j^{(a,b)} - O_{u,j}^{(b)} u_j \right\|_2 \right) \left[\lambda_{\min}(\Sigma_{u_j}) + o_p(1) \right]^{-1} \\
&= T O_p\left(\sqrt{\frac{\log(N \vee T)}{T}} \right) o_p\left((N \vee T)^{-1/2} \right) = o_p(1)
\end{aligned} \tag{A.52}$$

by (A.43) and (A.49). It suffices to study I_{1ij} below.

Now, let $\mathbb{Z}_{\mathfrak{B}}^{(1)} = \left(\mathbb{Z}_{\mathfrak{B},1}^{(1)\prime}, \cdots, \mathbb{Z}_{\mathfrak{B},N}^{(1)\prime} \right)'$, where $\mathbb{Z}_{\mathfrak{B},i}^{(1)} \sim \mathcal{N}\left(0, O_j^{(1)\prime} \Sigma_{u_j} O_j^{(1)} \right)$ for $i \in I_3$, $\mathbb{Z}_{\mathfrak{B},i}^{(1)} \sim \mathcal{N}\left(0, O_j^{(3)\prime} \Sigma_{u_j} O_j^{(3)} \right)$ for $i \in I_2$ and $\mathbb{Z}_{\mathfrak{B},i}^{(1)} \sim \mathcal{N}\left(0, O_j^{(2)\prime} \Sigma_{u_j} O_j^{(2)} \right)$ for $i \in I_1$. Note that

$$\begin{aligned}
I_{1ij} &= \left(O_j^{(b)} \mathbb{Z}_{\mathfrak{B},i}^{(1)} + o_{p(1)} + \mathcal{R}_{i,u}^j \right)' \Sigma_{u_j}^{-1} \left(O_j^{(b)} \mathbb{Z}_{\mathfrak{B},i}^{(1)} + o_p(1) + \mathcal{R}_{i,u}^j \right) \\
&+ T \left(\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)} u_j \right)' \left[\hat{\Sigma}_{u_j}^{-1} - \Sigma_{u_j}^{-1} \right] \left(\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)} u_j \right) \\
&= \mathbb{Z}_{\mathfrak{B},i}^{(1)\prime} \left(O_j^{(b)\prime} \Sigma_{u_j} O_j^{(b)} \right)^{-1} \mathbb{Z}_{\mathfrak{B},i}^{(1)} + o_p(1) \quad \text{uniformly over } i \in [N].
\end{aligned} \tag{A.53}$$

It follows that

$$S_{u_j} = \max_{i \in [N]} \begin{cases} \mathbb{Z}_{\mathfrak{B},i}^{(1)\prime} \left(O_j^{(2)\prime} \Sigma_{u_j} O_j^{(2)} \right)^{-1} \mathbb{Z}_{\mathfrak{B},i}^{(1)} + o_p(1), & \forall i \in I_1, \\ \mathbb{Z}_{\mathfrak{B},i}^{(1)\prime} \left(O_j^{(3)\prime} \Sigma_{u_j} O_j^{(3)} \right)^{-1} \mathbb{Z}_{\mathfrak{B},i}^{(1)} + o_p(1), & \forall i \in I_2, \\ \mathbb{Z}_{\mathfrak{B},i}^{(1)\prime} \left(O_j^{(1)\prime} \Sigma_{u_j} O_j^{(1)} \right)^{-1} \mathbb{Z}_{\mathfrak{B},i}^{(1)} + o_p(1), & \forall i \in I_3, \end{cases} \tag{A.54}$$

with $\mathbb{Z}_{\mathfrak{B},i}^{(1)\prime}\left(O_j^{(2)\prime}\Sigma_{u_j}O_j^{(2)}\right)^{-1}\mathbb{Z}_{\mathfrak{B},i}^{(1)} \to \chi^2(1)$ for each $i \in I_1$, $\mathbb{Z}_{\mathfrak{B},i}^{(1)\prime}\left(O_j^{(3)\prime}\Sigma_{u_j}O_j^{(3)}\right)^{-1}\mathbb{Z}_{\mathfrak{B},i}^{(1)} \to \chi^2(1)$ for each $i \in I_2$, and $\mathbb{Z}_{\mathfrak{B},i}^{(1)\prime}\left(O_j^{(1)\prime}\Sigma_{u_j}O_j^{(1)}\right)^{-1}\mathbb{Z}_{\mathfrak{B},i}^{(1)} \to \chi^2(1)$ for each $i \in I_3$. As in Castagnetti et al. (2015), we can conclude that

$$\mathbb{P}\left(\frac{1}{2}S_{u_j} \leq x + \mathsf{b}(N)\right) \to e^{-e^{-x}} \text{ as } (N,T) \to \infty. \tag{A.55}$$

For the test statistic for H_0^{II}, the proof is similar. Let $\mathbb{Z}_{\mathfrak{B}}^{(2)} = \left(\mathbb{Z}_{\mathfrak{B},1}^{(2)\prime}, \cdots, \mathbb{Z}_{\mathfrak{B},T}^{(2)\prime}\right)'$, where $\mathbb{Z}_{\mathfrak{B},t}^{(2)} \sim N\left(0, O_j^{(1)\prime}\Sigma_{v_j}O_j^{(1)}\right)$. As in (A.54), we can show that

$$S_{v_j}^{(3,1)} = \max_{t\in[T]} \mathbb{Z}_{\mathfrak{B},t}^{(2)\prime}\left(O_j^{(1)\prime}\Sigma_{v_j}O_j^{(1)}\right)^{-1}\mathbb{Z}_{\mathfrak{B},t}^{(2)} + o_p(1).$$

By the strong mixing condition in Assumption 1(iii), we have $\max_{j\in[p],i\in[N]}\left\|Cov\left(\mathfrak{b}_{j,it}^{(2)}, \mathfrak{b}_{j,is}^{(2)}\right)\right\|_\infty \log(t-s) = o(1)$ as $t - s \to \infty$ by Davydov's inequality. Then by Theorem 3.5.1 in Leadbetter and Rootzen (1988), we have that

$$\mathbb{P}\left(\frac{1}{2}\left(\max_{t\in[T]}\mathbb{Z}_{\mathfrak{B},t}^{(2)\prime}\left(O_j^{(1)\prime}\Sigma_{v_j}O_j^{(1)}\right)^{-1}\mathbb{Z}_{\mathfrak{B},t}^{(2)}\right) \leq x + \mathsf{b}(T)\right) \to e^{-e^{-x}},$$

which implies that

$$\mathbb{P}\left(\frac{1}{2}S_{v_j}^{(3,1)} \leq x + \mathsf{b}(T)\right) \to e^{-e^{-x}} \text{ as } (N,T) \to \infty.$$

Recall that $\tilde{S}_{v_j}^{(a,b)} = \frac{1}{2}S_{v_j}^{(a,b)} - \mathsf{b}(T)$ and $S_{v_j} = \max(\tilde{S}_{v_j}^{(1,2)}, \tilde{S}_{v_j}^{(2,3)}, \tilde{S}_{v_j}^{(3,1)})$. Noting that S_{v_j} is asymptotically distributed as the maximum of three independent Gumbel random variables under H_0^{II}, we have $\mathbb{P}(S_{v_j} \leq x) \to e^{-3e^{-x}}$ as $(N,T) \to \infty$.

A.5.2 Proof of Statement (ii)

Under H_1^I, we have that $\forall i \in I_a$,

$$\begin{aligned}
\hat{u}_{i,j}^{(a,b)} - \hat{\bar{u}}_j^{(a,b)} &= \left(\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)}u_{i,j}^0\right) + O_{u,j}^{(b)}\left(u_{i,j}^0 - u_j\right) - \left(\hat{\bar{u}}_j^{(a,b)} - O_{u,j}^{(b)}u_j\right) \\
&= \left(\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)}u_{i,j}^0\right) + O_{u,j}^{(b)}c_{i,j}^u - \left(\hat{\bar{u}}_j^{(a,b)} - O_{u,j}^{(b)}u_j\right).
\end{aligned}$$

Then

$$\begin{aligned}
&T\left(\hat{u}_{i,j}^{(a,b)} - \hat{\bar{u}}_j^{(a,b)}\right)'\left(\hat{\Sigma}_{u_j}\right)^{-1}\left(\hat{u}_{i,j}^{(a,b)} - \hat{\bar{u}}_j^{(a,b)}\right) \\
&= T\left(\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)}u_j\right)'\hat{\Sigma}_{u_j}^{-1}\left(\hat{u}_{i,j}^{(1)} - O_{u,j}^{(b)}u_j\right) + T\left(\hat{\bar{u}}_j^{(1)} - O_{u,j}^{(b)}u_j\right)'\hat{\Sigma}_{u_j}^{-1}\left(\hat{\bar{u}}_j^{(a,b)} - O_{u,j}^{(b)}u_j\right) \\
&\quad - 2T\left(\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)}u_j\right)'\hat{\Sigma}_{u_j}^{-1}\left(\hat{\bar{u}}_j^{(1)} - O_{u,j}^{(b)}u_j\right) + 2T\left(O_{u,j}^{(b)}c_{i,j}^u\right)'\hat{\Sigma}_{u_j}^{-1}\left(\hat{u}_{i,j}^{(a,b)} - O_{u,j}^{(b)}u_j\right) \\
&\quad + T\left(O_{u,j}^{(b)}c_{i,j}^u\right)'\hat{\Sigma}_{u_j}^{-1}O_{u,j}^{(b)}c_{i,j}^u - 2T\left(O_{u,j}^{(b)}c_{i,j}^u\right)'\hat{\Sigma}_{u_j}^{-1}\left(\hat{\bar{u}}_j^{(a,b)} - O_{u,j}^{(b)}u_j\right) \\
&:= S_{u_j,i,1}^{(b)} + S_{u_j,2}^{(b)} + S_{u_j,i,3}^{(b)} + S_{u_j,i,4}^{(b)} + S_{u_j,i,5}^{(b)} + S_{u_j,i,6}^{(b)},
\end{aligned}$$

where $\max_{i\in I_a}\left|S_{u_j,i,1}^{(b)}\right| = O_p(\log N)$ by (A.53) and (A.55), $\left|S_{u_j,2}^{(b)}\right| = o_p(1)$ by (A.51), $\max_{i\in I_3}\left|S_{u_j,i,3}^{(b)}\right| = o_p(1)$ by (A.52). Next,

$$\max_{i\in I_a}\left|S_{u_j,i,5}^{(b)}\right| = \max_{i\in I_a} T\left(O_{u,j}^{(b)}c_{i,j}^u\right)'\Sigma_{u_j}^{-1}O_{u,j}^{(b)}c_{i,j}^u\left\{1 + o_p(1)\right\} \gtrsim \left[\max_{i\in I_a}\lambda_{\max}\left(\Sigma_{u,j}\right)\right]^{-1} T\max_{i\in I_a}\left\|c_{i,j}^u\right\|_2^2 \asymp T\max_{i\in I_a}\left\|c_{i,j}^u\right\|_2^2,$$

which diverges to infinity at the rate faster than $\log N$ by condition in statement (ii). By Cauchy-Schwarz inequality, we obtain that

$$\max_{i\in I_a}\left|S^{(b)}_{u_j,i,4}\right| \lesssim \max_{i\in I_a}\left(S^{(b)}_{u_j,i,1}\right)^{1/2}\left(S^{(b)}_{u_j,i,5}\right)^{1/2} = o_p\left(\max_{i\in I_a}\left|S^{(b)}_{u_j,i,5}\right|\right), \text{ and}$$
$$\max_{i\in I_a}\left|S^{(b)}_{u_j,i,6}\right| \lesssim \max_{i\in I_a}\left(S^{a}_{u_j,i,2}\right)^{1/2}\left(S^{(b)}_{u_j,i,5}\right)^{1/2} = o_p\left(\max_{i\in I_a}\left|S^{(b)}_{u_j,i,5}\right|\right).$$

It follows that $\mathbb{P}\left\{S_{u_j} > c_{\alpha,N}\right\} \to 1$ as $c_{\alpha,N} \asymp \log N$, and the final result follows.

The power of the test statistic S_{v_j} can be analyzed analogously. ■

A.6 Proof of Theorem 4.2

We first derive the linear expansion of $\hat{\Theta}^*_{j,it} - \Theta^*_{j,it}$ for $i \in I_3$, and similar results hold for $i \in I_1 \cup I_2$. Let $\bar{v}^0_j := \frac{1}{T}\sum_{t\in[T]} v^0_{t,j}$ and $\bar{u}^{0,I_a}_j := \frac{1}{N_a}\sum_{i\in I_a} u^0_{i,j}$. By Proposition 3.4, we obtain that uniformly in $i \in I_a$,

$$\begin{aligned}
\hat{\bar{\Theta}}_{j,i\cdot} - \bar{\Theta}_{j,i\cdot} &= \frac{1}{T}\sum_{t\in[T]}\left(\hat{\Theta}_{j,it} - \Theta^0_{j,it}\right) \\
&= \frac{1}{T}\sum_{t\in[T]} u^{0\prime}_{i,j}\left(\hat{V}^{(a)}_{v_j,t}\right)^{-1}\frac{1}{N_a}\sum_{i^*\in I_a}\xi^0_{j,i^*t} + \frac{1}{T}\sum_{t\in[T]} v^{0\prime}_{t,j}\hat{V}^{-1}_{u_j}\frac{1}{T}\sum_{t=1}^{T} b^0_{j,it} + \frac{1}{T}\sum_{t\in[T]}\mathcal{R}^j_{it} \\
&= \bar{v}^{0\prime}_j\hat{V}^{-1}_{u_j}\frac{1}{T}\sum_{t=1}^{T} b^0_{j,it} + o_p\left((N\vee T)^{-1/2}\right),
\end{aligned} \tag{A.56}$$

where the third equality holds by the fact that

$$\begin{aligned}
&\max_{i\in I_a}\left|\frac{1}{T}\sum_{t\in[T]} u^{0\prime}_{i,j}\left(\hat{V}^{(a)}_{v_j,t}\right)^{-1}\frac{1}{N_a}\sum_{i^*\in I_a}\xi^0_{j,i^*t}\right| = \max_{i\in I_a}\left|\frac{1}{N_aT}\sum_{i^*\in I_a}\sum_{t\in[T]} u^{0\prime}_{i,j}\left(\hat{V}^{(a)}_{v_j,t}\right)^{-1}\xi^0_{j,i^*t}\right| \\
&= O_p\left(\sqrt{\frac{\log N}{NT}}\xi_N\right) = o_p\left((N\vee T)^{-1/2}\right),
\end{aligned}$$

by conditional Bernstein's inequality given $\mathscr{D}_e$ in Lemma B.12(i), Assumption 1(i)-(ii) and Assumption 1(ix). Similarly, uniformly in $t \in [T]$, we have

$$\begin{aligned}
\hat{\bar{\Theta}}^{I_a}_{j,\cdot t} - \bar{\Theta}^{I_a}_{j,\cdot t} &= \frac{1}{N_a}\sum_{i\in I_3}\left(\hat{\Theta}_{j,it} - \Theta^0_{j,it}\right) \\
&= \frac{1}{N_a}\sum_{i\in I_a} u^{0\prime}_{i,j}\left(\hat{V}^{(a)}_{v_j,t}\right)^{-1}\frac{1}{N_a}\sum_{i\in I_a}\xi^0_{j,it} + \frac{1}{N_a}\sum_{i\in I_a} v^{0\prime}_{t,j}\hat{V}^{-1}_{u_j}\frac{1}{T}\sum_{t^*=1}^{T} b^0_{j,it^*} + \frac{1}{N_a}\sum_{i\in I_a}\mathcal{R}^j_{it} \\
&= \bar{u}^{0,I_a\prime}_j\left(\hat{V}^{(a)}_{v_j,t}\right)^{-1}\frac{1}{N_a}\sum_{i\in I_a}\xi^0_{j,it} + o_p\left((N\vee T)^{-1/2}\right),
\end{aligned} \tag{A.57}$$

and

$$\begin{aligned}
\hat{\bar{\Theta}}^{I_a}_{j} - \bar{\Theta}^{I_a}_{j} &= \frac{1}{N_aT}\sum_{i\in I_a}\sum_{t\in[T]}\left(\hat{\Theta}_{j,it} - \Theta^0_{j,it}\right) \\
&= \frac{1}{N_aT}\sum_{i\in I_a}\sum_{t\in[T]} u^{0\prime}_{i,j}\left(\hat{V}^{(a)}_{v_j,t}\right)^{-1}\frac{1}{N_a}\sum_{i\in I_a}\xi^0_{j,it} + \frac{1}{N_aT}\sum_{i\in I_a}\sum_{t\in[T]} v^{0\prime}_{t,j}\hat{V}^{-1}_{u_j}\frac{1}{T}\sum_{t=1}^{T} b^0_{j,it} + \frac{1}{N_aT}\sum_{i\in I_a}\sum_{t\in[T]}\mathcal{R}^j_{it}
\end{aligned}$$

$$= o_p\left((N \vee T)^{-1/2}\right). \tag{A.58}$$

Combining (A.56)-(A.58), we obtain that $\forall i \in I_a$ and $t \in [T]$,

$$\begin{aligned}\hat{\Theta}^*_{j,it} - \Theta^*_{j,it} &= \hat{\Theta}_{j,it} - \Theta^0_{j,it} - \left(\hat{\bar{\Theta}}_{j,i\cdot} - \bar{\Theta}_{j,i\cdot}\right) - \left(\hat{\bar{\Theta}}^{I_a}_{j,\cdot t} - \bar{\Theta}^{I_a}_{j,\cdot t}\right) + \left(\hat{\bar{\Theta}}^{I_a}_{j} - \bar{\Theta}^{I_a}_{j}\right) \\ &= \left(u^0_{i,j} - \bar{u}^{0,I_a}_j\right)' \left(\hat{V}^{(a)}_{v_j,t}\right)^{-1} \frac{1}{N_a}\sum_{i\in I_a}\xi^0_{j,it} + \left(v^0_{t,j} - \bar{v}^0_j\right)' \hat{V}^{-1}_{u_j}\frac{1}{T}\sum_{t=1}^{T} b^0_{j,it} + \bar{\mathcal{R}}^j_{it},\end{aligned}$$

such that $\max_{i\in I_3, t\in[T]}\left|\bar{\mathcal{R}}^j_{it}\right| = o_p\left((N\vee T)^{-1/2}\right)$. It follows that $\hat{\Theta}^*_{j,it} - \Theta^*_{j,it} \rightsquigarrow \mathcal{N}(0, \Sigma^*_{j,it})$ with

$$\Sigma^*_{j,it} = \sum_{a\in[3]}\frac{1}{N_a}\left(O^{(b)}_j u^0_{i,j} - O^{(b)}_j\bar{u}^0_j\right)'\Sigma_{v_j}\left(O^{(b)}_j u^0_{i,j} - O^{(b)}_j\bar{u}^0_j\right)\mathbf{1}_{ia} + \frac{1}{T}\left(O^{(b)}_j v^0_{t,j} - O^{(b)}_j\bar{v}^0_j\right)'\Sigma_{u_j}\left(O^{(b)}_j v^0_{t,j} - O^{(b)}_j\bar{v}^0_j\right),$$

and Σ_{u_j} and Σ_{v_j} are as defined in Theorem 3.3. The reason why $\Sigma^*_{j,it}$ is not indexed with b is owing to the fact that $O^{(b)}_j$ shown in the right side of the equality can be absorbed by $O^{(b)}_j$ not shown in Σ_{u_j} and Σ_{v_j}.

Define $\hat{\Sigma}^*_{j,it} = \frac{1}{2}\sum_b \hat{\Sigma}^{(b)*}_{j,it}$ with

$$\hat{\Sigma}^{(b)*}_{j,it} = \sum_{a\in[3]}\left[\frac{1}{N_a}\left(\hat{u}^{(a,b)}_{i,j} - \hat{\bar{u}}^{(a,b)}_j\right)'\hat{\Sigma}_{v_j}\left(\hat{u}^{(a,b)}_{i,j} - \bar{u}^{(a,b)}_j\right)\mathbf{1}_{ia} + \frac{1}{T}\left(\hat{v}^{(a,b)}_{t,j} - \hat{\bar{v}}^{(a,b)}_j\right)'\hat{\Sigma}_{u_j}\left(\hat{v}_{t,j} - \hat{\bar{v}}^{(a,b)}_j\right)\right],$$

where $\hat{\bar{u}}^{(a,b)}_j = \frac{1}{N_a}\sum_{i\in I_a}\hat{u}^{(a,b)}_{i,j}$ and $\hat{\bar{v}}^{(a,b)}_j = \frac{1}{T}\sum_{t\in[T]}\hat{v}^{(a,b)}_{t,j}$. By Theorem 3.3 and Lemma B.33, we have

$$\max_{j\in[p], i\in[N], t\in[T]}\left|\hat{\Sigma}^*_{j,it} - \Sigma^*_{j,it}\right| = o_p(1).$$

By arguments as used in the proof of Theorem 3.3, we have that as $(N,T)\to\infty$, $\mathbb{P}\left(\frac{1}{2}S_{NT} \le x + \mathsf{b}(NT)\right) \to e^{-e^{-x}}$ under H^{III}_0, and $\mathbb{P}\left(S_{NT} > c_{\alpha,3\cdot NT}\right) \to 1$ under H^{III}_1 provided $\frac{N\wedge T}{\log NT}\max_{i\in[N],t\in[T]}\left|\Theta^*_{j,it}\right|^2 \to \infty$. ■

B Some Technical Lemmas

In this section we state and prove some technical lemmas that are used in the proofs of the main results in the paper.

B.1 Lemmas for the Proof of Theorem 3.1

Lemma B.1 *Consider a matrix sequence $\{A_i, i = 1, \cdots, N\}$ whose values are symmetric matrices with dimension d,*

(i) suppose $\{A_i, i = 1, \cdots, N\}$ is independent with $\mathbb{E}(A_i) = 0$ and $\|A_i\|_{op} \le M$ a.s. Let $\sigma^2 = \left\|\sum_{i\in[N]}\mathbb{E}\left(A_i^2\right)\right\|_{op}$. Then for all $t > 0$, we have

$$\mathbb{P}\left(\left\|\sum_{i\in[N]}A_i\right\|_{op} > t\right) \le d\cdot\exp\left\{-\frac{t^2/2}{\sigma^2 + Mt/3}\right\}.$$

(ii) suppose $\{A_i, i = 1, \cdots, N\}$ *is sequence of martingale difference matrices with* $\mathbb{E}_{i-1}(A_i) = 0$ *and* $\|A_i\|_{op} \leq M$ *a.s., where* $\mathbb{E}_{i-1}$ *denotes* $\mathbb{E}(\cdot|\mathscr{F}_{i-1})$*, where* $\{\mathscr{F}_i : i \leq N\}$ *denotes the filtration that is clear from the context. Let* $\left\|\sum_{i\in[N]} \mathbb{E}_{i-1}\left(A_i^2\right)\right\|_{op} \leq \sigma^2$*. Then for all* $t > 0$*, we have*

$$\mathbb{P}\left(\left\|\sum_{i\in[N]} A_i\right\|_{op} > t\right) \leq d \cdot \exp\left\{-\frac{t^2/2}{\sigma^2 + Mt/3}\right\}.$$

Proof. Lemma B.1(i) and (ii) are Matrix Bernstein inequality and Matrix Freedman inequality, which are respectively stated in Theorem 1.3 and Corollary 4.2 in Tropp (2011). ■

Lemma B.2 *Consider a specific matrix* $A \in \mathbb{R}^{N\times T}$ *whose rows (denoted as* A_i' *where* $A_i \in \mathbb{R}^T$*) are independent random vectors in* $\mathbb{R}^T$ *with* $\mathbb{E}A_i = 0$ *and* $\Sigma_i = \mathbb{E}(A_iA_i')$*. Suppose* $\max_{i\in[N]} \|A_i\|_2 \leq \sqrt{m}$ *a.s. and* $\max_{i\in[N]} \|\Sigma_i\|_{op} \leq M$ *for some positive constant* M*. Then for every* $t > 0$*, with probability* $1 - 2T\exp\left(-c_1t^2\right)$*, we have*

$$\|A\|_{op} \leq \sqrt{NM} + t\sqrt{m+M},$$

where c_1 *is an absolute constant.*

Proof. The proof follows similar arguments as used in the proof of Theorem 5.41 in Vershynin (2010). Define $Z_i := \frac{1}{N}(A_iA_i' - \Sigma_i) \in \mathbb{R}^{T\times T}$. We notice that $\{Z_i\}$ is an independent sequence with $\mathbb{E}(Z_i) = 0$. To use the matrix Bernstein inequality, we analyze $\|Z_i\|_{op}$ and $\left\|\sum_{i\in[N]} \mathbb{E}\left(X_i^2\right)\right\|_{op}$ as follows. First, note that uniformly over i,

$$\|Z_i\|_{op} \leq \frac{1}{N}\left(\|A_iA_i'\|_{op} + \|\Sigma_i\|_{op}\right) \leq \frac{1}{N}\left(\|A_i\|_2^2 + \|\Sigma_i\|_{op}\right) \leq \frac{m+M}{N}, \quad \text{a.s.} \tag{B.1}$$

Next, noting that $\mathbb{E}\left[(A_iA_i')^2\right] = \mathbb{E}\left[\|A_i\|_2 A_iA_i'\right] \leq m\Sigma_i$ and $Z_i^2 = \frac{1}{N^2}[(A_iA_i')^2 - A_iA_i'\Sigma_i - \Sigma_iA_iA_i' + \Sigma_i^2]$, we have

$$\begin{aligned}\left\|\mathbb{E}\left(Z_i^2\right)\right\|_{op} &= \frac{1}{N^2}\left\|\mathbb{E}\left[(A_iA_i')^2 - \Sigma_i^2\right]\right\|_{op} \leq \frac{1}{N^2}\left\{\left\|\mathbb{E}\left[(A_iA_i')^2\right]\right\|_{op} + \|\Sigma_i\|_{op}^2\right\} \\ &\leq \frac{1}{N^2}\left(m\|\Sigma_i\|_{op} + \|\Sigma_i\|_{op}^2\right) \leq \frac{mM + M^2}{N^2} \text{ uniformly in } i.\end{aligned}$$

It follows that

$$\left\|\sum_{i\in[N]} \mathbb{E}\left(Z_i^2\right)\right\|_{op} \leq N\max_{i\in[N]}\left\|\mathbb{E}\left(Z_i^2\right)\right\|_{op} \leq \frac{mM + M^2}{N}. \tag{B.2}$$

Let $\varepsilon = \max\left(\sqrt{M}\delta, \delta^2\right)$ with $\delta = t\sqrt{\frac{m+M}{N}}$. By (B.1)-(B.2) and the matrix Bernstein inequality in Lemma B.1(i), we have

$$\begin{aligned}\mathbb{P}\left\{\left\|\frac{1}{N}\left(A'A - \sum_{i\in[N]}\Sigma_i\right)\right\|_{op} \geq \varepsilon\right\} &= \mathbb{P}\left(\left\|\sum_{i\in[N]} Z_i\right\|_{op} \geq \varepsilon\right) \\ &\leq 2T\exp\left\{-c_1\min\left(\frac{\varepsilon^2}{\frac{mM+M^2}{N}}, \frac{\varepsilon}{\frac{m+M}{N}}\right)\right\}\end{aligned}$$

$$\leq 2T \exp\left\{-c_1 \min\left(\frac{\varepsilon^2}{M}, \varepsilon\right) \frac{N}{m+M}\right\}$$
$$\leq 2T \exp\left\{-\frac{c_1 \delta^2 N}{m+M}\right\} = 2T \exp\left\{-c_1 t^2\right\},$$

for some positive constant c_1, where the third inequality is due to the fact that

$$\min\left(\frac{\varepsilon^2}{M}, \varepsilon\right) = \min\left(\max\left(\delta^2, \delta^4/M\right), \max\left(\sqrt{M}\delta, \delta\right)\right)$$
$$= \begin{cases} \min\left(\delta^2, \sqrt{M}\delta\right) = \delta^2 & \text{if} \quad \delta^2 \geq \dfrac{\delta^4}{M}, \\ \min\left(\delta^4/M, \delta^2\right) = \delta^2, & \text{if} \quad \delta^2 < \dfrac{\delta^4}{M}. \end{cases}.$$

That is,

$$\left\|\frac{1}{N}A'A - \frac{1}{N}\sum_{i\in[N]} \Sigma_i\right\|_{op} \leq \max\left(\sqrt{M}\delta, \delta^2\right) \tag{B.3}$$

with probability $1 - \exp\left(-c_1 t^2\right)$. Combining the fact that $\|\Sigma_i\| \leq M$ uniformly over i and (B.3), we show that

$$\frac{1}{N}\|A\|_{op}^2 = \left\|\frac{1}{N}A'A\right\|_{op} \leq \left\|\frac{1}{N}\sum_{i\in[N]} \Sigma_i\right\|_{op} + \left\|\frac{1}{N}A'A - \frac{1}{N}\sum_{i\in[N]} \Sigma_i\right\|_{op}$$
$$\leq \max_{i\in[N]} \|\Sigma_i\|_{op} + \sqrt{M}\delta + \delta^2$$
$$\leq M + \sqrt{M}t\sqrt{\frac{m+M}{N}} + t^2\frac{m+M}{N} \leq \left(\sqrt{M} + t\sqrt{\frac{m+M}{N}}\right)^2.$$

It follows that $\|A\|_{op} \leq \sqrt{NM} + t\sqrt{m+M}$. ■

Lemma B.3 *Recall* $a_{it} = \tau - \mathbf{1}\{\epsilon_{it} \leq 0\}$ *and* $a = \{a_{it}\} \in \mathbb{R}^{N\times T}$. *Under Assumption 1, we have* $\|X_j \odot a\|_{op} = O_p\left(\sqrt{N} + \sqrt{T\log T}\right)$ $\forall j \in [p]$ *and* $\|a\|_{op} = O_p\left(\sqrt{N} + \sqrt{T\log T}\right)$.

Proof. We focus on $\|X_j \odot a\|_{op}$ as the result for $||a||_{op}$ can be derived in the same manner. We first note that, conditional $\mathscr{D}$, the i-th row of $X_j \odot a$ only depends on $\{e_{it}, \epsilon_{it}\}_{t\in[T]}$, which are independent across i. Therefore $X_j \odot a$ has independent rows, denoted as $A_i = X_{j,i} \odot a_i$, given $\mathscr{D}$, where $X_{j,i}$ and a_i being the i-th row of matrix X_j and a, respectively. In addition, for the t-th element of A_i, we have

$$\mathbb{E}\left[X_{j,it}a_{it}\middle|\mathscr{D}\right] = \mathbb{E}\left\{X_{j,it}\mathbb{E}\left[a_{it}\middle|\mathscr{D}_e\right]\middle|\mathscr{D}\right\} = 0,$$

where the second equality holds by Assumption 1(ii) and the fact that given $\mathscr{D}_e$, $X_{j,it}$ is known. Therefore, In order to apply Lemma B.1, conditionally on $\mathscr{D}$, we only need to upper bound $\|A_i\|_2$ and $\mathbb{E}\left[A_iA_i'\middle|\mathscr{D}\right]$.

First, under Assumption 1(iv), we have $\frac{1}{T}\sum_{t\in[T]}\left(X_{j,it}a_{it}\right)^2 \leq \frac{1}{T}\sum_{t\in[T]} X_{j,it}^2 \leq c_2$ a.s. for some positive constant c_2, which implies

$$||A_i||_2 = \|X_{j,i} \odot a_i\|_2 \leq c_2\sqrt{T} \quad \text{a.s.} \tag{B.4}$$

Second, let $\Sigma_i = \mathbb{E}\left\{\left[(X_{j,i} \odot a_i)(X_{j,i} \odot a_i)'\right]\middle|\mathscr{D}\right\}$ with $(t,s)^{th}$ element being $\mathbb{E}\left(X_{j,it}X_{j,is}a_{it}a_{is}\middle|\mathscr{D}\right)$. Recall that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are matrix norms induced by 1- and $\infty-$norms, i.e.,

$$\|\Sigma_i\|_1 = \max_{s\in[T]} \sum_{t\in[T]} \left|\mathbb{E}\left(X_{j,it}X_{j,is}a_{it}a_{is}\middle|\mathscr{D}\right)\right|, \quad \|\Sigma_i\|_\infty = \max_{t\in[T]} \sum_{s\in[T]} \left|\mathbb{E}\left(X_{j,it}X_{j,is}a_{it}a_{is}\middle|\mathscr{D}\right)\right|.$$

By Davydov's inequality for conditional strong mixing sequence, we can show that

$$\begin{aligned}
&\max_{s\in[T]} \sum_{t\in[T]} \left|\mathbb{E}\left(X_{j,it}X_{j,is}a_{it}a_{is}\middle|\mathscr{D}\right)\right| \\
&= \max_{s\in[T]} \sum_{t\in[T]} \left|Cov\left(X_{j,it}a_{it}, X_{j,is}a_{is}\middle|\mathscr{D}\right)\right| \\
&\leq \max_{s\in[T]} \sum_{t\in[T]} \left\{\mathbb{E}\left[|X_{j,it}a_{it}|^q\middle|\mathscr{D}\right]\right\}^{1/q} \left\{\mathbb{E}\left[|X_{j,is}a_{is}\}|^q\middle|\mathscr{D}\right]\right\}^{1/q} \times \alpha(t-s)^{(q-2)/q} \\
&\leq \max_{i\in[N],t\in[T]} \left\{\mathbb{E}\left[\left|X_{j,it}^q\right|\middle|\mathscr{D}\right]\right\}^{2/q} \max_{s\in[T]} \sum_{t\in[T]} \alpha(t-s)^{(q-2)/q} \\
&\leq c_3 \ a.s.,
\end{aligned}$$

where c_3 is a positive constant which does not depend on i and the last line is by Assumption 1(iii) and 1(iv). Similarly, we have $\max_{t\in[T]} \sum_{s\in[T]} \left|\mathbb{E}\left(X_{j,it}X_{j,is}a_{it}a_{is}\middle|\mathscr{D}\right)\right| \leq c_3$ a.s. Then by Corollary 2.3.2 in Golub and Van Loan (1996), we have

$$\max_{i\in[N]} \|\Sigma_i\|_{op} \leq \sqrt{\|\Sigma_i\|_1 \|\Sigma_i\|_\infty} \leq c_3 \quad \text{a.s.} \tag{B.5}$$

Combining (B.1), (B.2), and Lemma B.1 with $t = \sqrt{\log T}$, we obtain the desired result. ■

Recall that $\Delta_{\Theta_j} = \Theta_j - \Theta_j^0$ for any Θ_j and define

$$\mathcal{R}(C_1) := \left\{\{\Delta_{\Theta_j}\}_{j=0}^p : \sum_{j=0}^p \left\|\mathcal{P}_j^\perp(\Delta_{\Theta_j})\right\|_* \leq C_1 \sum_{j=0}^p \left\|\mathcal{P}_j(\Delta_{\Theta_j})\right\|_*\right\}.$$

Lemma B.4 *Suppose Assumptions 1-3 hold. Then* $\{\tilde{\Delta}_{\Theta_j}\}_{j=0}^p \in \mathcal{R}(3)$ *w.p.a.1.*

Proof. Define $\mathbb{Q}_\tau\left(\{\Theta_j\}_{j=0}^p\right) = \frac{1}{NT}\sum_{i=1}^N \sum_{t=1}^T \rho_\tau\left(Y_{it} - \Theta_{0,it} - \sum_{j=1}^p X_{j,it}\Theta_{j,it}\right)$ for generic $\{\Theta_j\}_{j=0}^p$. By the definition of the nuclear norm estimator in (2.9), we have

$$\mathbb{Q}_\tau\left(\{\Theta_j^0\}_{j=0}^p\right) - \mathbb{Q}_\tau\left(\left\{\tilde{\Theta}_j\right\}_{j=0}^p\right) + \sum_{j=0}^p \nu_j\left[\left\|\Theta_j^0(\tau)\right\|_* - \left\|\tilde{\Theta}_j(\tau)\right\|_*\right] \geq 0. \tag{B.6}$$

In addition, we have

$$\mathbb{Q}_\tau\left(\{\Theta_j^0\}_{j=0}^p\right) - \mathbb{Q}_\tau\left(\left\{\tilde{\Theta}_j\right\}_{j=0}^p\right)$$

$$
\begin{aligned}
&= -\left\{\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}a_{it}\left(\tilde{\Delta}_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\tilde{\Delta}_{\Theta_{j,it}}\right)+\int_{0}^{\tilde{\Delta}_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\tilde{\Delta}_{\Theta_{j,it}}}\mathbf{1}\left\{\epsilon_{it}\leq s\right\}-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}ds\right\}\\
&\leq\left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}a_{it}\left(\tilde{\Delta}_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\tilde{\Delta}_{\Theta_{j,it}}\right)\right|\\
&\leq\sum_{j=1}^{p}\left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}a_{it}X_{j,it}\tilde{\Delta}_{\Theta_{j,it}}\right|+\left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}a_{it}\tilde{\Delta}_{\Theta_{0,it}}\right|\\
&=\sum_{j=1}^{p}\frac{1}{NT}\left|tr\left[\tilde{\Delta}_{\Theta_j}'\left(X_j\odot a\right)\right]\right|+\frac{1}{NT}\left|tr\left(\tilde{\Delta}_{\Theta_0}'a\right)\right|\\
&\leq\sum_{j=1}^{p}\frac{1}{NT}\left\|\tilde{\Delta}_{\Theta_j}\right\|_{*}\left\|X_j\odot a\right\|_{op}+\frac{1}{NT}\left\|\tilde{\Delta}_{\Theta_0}\right\|_{*}\left\|a\right\|_{op}\\
&\leq c_4\sum_{j=0}^{p}\frac{(\sqrt{N}\vee\sqrt{T\log T})}{NT}\left\|\tilde{\Delta}_{\Theta_j}\right\|_{*},\quad\text{w.p.a.1,}
\end{aligned}
\tag{B.7}
$$

where the first equality holds by Knight's identity in Knight (1998) which states that

$$
\rho_\tau(u-v)-\rho_\tau(u)=v(\tau-\mathbf{1}\left\{u\leq 0\right\})+\int_0^v\left(\mathbf{1}\left\{u\leq s\right\}-\mathbf{1}\{u\leq 0\}\right)ds,
\tag{B.8}
$$

the first inequality is due to the fact that the second term in the bracket of the second line is non-negative, the second inequality holds by triangle inequality, the third inequality is by the fact that $tr(AB)\leq ||A||_{op}||B||_{*}$, and the last inequality holds by Lemma B.3.

Combining (B.6) and (B.7), w.p.a.1, we have

$$
0\leq c_4\sum_{j=0}^{p}\left\{\frac{(\sqrt{N}\vee\sqrt{T\log T})}{NT}||\tilde{\Delta}_{\Theta_j}||_{*}+\nu_j\left(\left\|\Theta_j^0\right\|_{*}-\left\|\tilde{\Theta}_j\right\|_{*}\right)\right\}.
\tag{B.9}
$$

Besides, we can show that

$$
\begin{aligned}
\left\|\tilde{\Theta}_j\right\|_{*}&=\left\|\tilde{\Delta}_{\Theta_j}+\Theta_j^0\right\|_{*}=\left\|\Theta_j^0+\mathcal{P}_j^{\perp}(\tilde{\Delta}_{\Theta_j})+\mathcal{P}_j(\tilde{\Delta}_{\Theta_j})\right\|_{*}\\
&\geq\left\|\Theta_j^0+\mathcal{P}_j^{\perp}(\tilde{\Delta}_{\Theta_j})\right\|_{*}-\left\|\mathcal{P}_j(\tilde{\Delta}_{\Theta_j})\right\|_{*}=\|\Theta_j^0\|_{*}+\left\|\mathcal{P}_j^{\perp}(\tilde{\Delta}_{\Theta_j})\right\|_{*}-\left\|\mathcal{P}_j(\tilde{\Delta}_{\Theta_j})\right\|_{*}
\end{aligned}
\tag{B.10}
$$

where the second equality holds by Lemma D.2(i) in Chernozhukov et al. (2019), the first inequality holds by triangle inequality, and the last equality is by the construction of the linear space $\mathcal{P}_j^{\perp}$ and $\mathcal{P}_j$. Then combining (B.9) and (B.10), we obtain

$$
\begin{aligned}
\sum_{j=0}^{p}\nu_j\left\{\left\|\mathcal{P}_j^{\perp}\left(\tilde{\Delta}_{\Theta_j}\right)\right\|_{*}-\left\|\mathcal{P}_j\left(\tilde{\Delta}_{\Theta_j}\right)\right\|_{*}\right\}&\leq c_4\sum_{j=0}^{p}\frac{(\sqrt{N}\vee\sqrt{T\log T})}{NT}\left\|\tilde{\Delta}_{\Theta_j}\right\|_{*}\\
&=c_4\sum_{j=0}^{p}\frac{(\sqrt{N}\vee\sqrt{T\log T})}{NT}\left\{\left\|\mathcal{P}_j\left(\tilde{\Delta}_{\Theta_j}\right)\right\|_{*}+\left\|\mathcal{P}_j^{\perp}\left(\tilde{\Delta}_{\Theta_j}\right)\right\|_{*}\right\},\ \text{w.p.a.1}
\end{aligned}
$$

By setting $\nu_j=\frac{2c_4(\sqrt{N}\vee\sqrt{T\log T})}{NT}$, we obtain $\sum_{j=0}^{p}\left\|\mathcal{P}_j^{\perp}\left(\tilde{\Delta}_{\Theta_j}\right)\right\|_{*}\leq 3\sum_{j=0}^{p}\left\|\mathcal{P}_j\left(\tilde{\Delta}_{\Theta_j}\right)\right\|_{*}$ w.p.a.1. ■

Recall $\mathscr{G}_{i,t-1}$ is the σ-field generated by $\{V_j^0\}_{j\in[p]\cup\{0\}}$, $\{W_j^0\}_{j\in[p]}$, $\{e_{ls}\}_{l\leq i-1,s\in[T]}$, $\{e_{is}\}_{s\leq t}$, $\{\epsilon_{ls}\}_{l\leq i-1,s\in[T]}$, and $\{\epsilon_{is}\}_{s\leq t-1}$ and $\mathsf{F}_{it}(\cdot)$ and $\mathsf{f}_{it}(\cdot)$ are the conditional CDF and PDF of ϵ_{it} given $\mathscr{G}_{i,t-1}$, respectively. Specifically, we note that $(\{X_{j,it}\}_{j\in[p]},\{\Theta_{j,it}^0\}_{j\in[p]\cup\{0\}})$ are measurable w.r.t. $\mathscr{G}_{i,t-1}$.

Lemma B.5 *For all $u_1, u_2 \in \mathbb{R}$ and all $c_5 \in (0,1]$, we have*

(i) $\int_0^{u_2} (\mathbf{1}\{u_1 \le z\} - \mathbf{1}\{u_1 \le 0\})\, dz \ge \int_0^{c_5 u_2} (\mathbf{1}\{u_1 \le z\} - \mathbf{1}\{u_1 \le 0\})\, dz \ge 0$,

(ii) $\int_0^{u_2} \{\mathsf{F}_{it}(u_1+z) - \mathsf{F}_{it}(u_1)\}\, dz \ge \int_0^{c_5 u_2} \{\mathsf{F}_{it}(u_1+z) - \mathsf{F}_{it}(u_1)\}\, dz \ge 0$.

Proof. Statement (i) is just Feng (2019, Lemma A2). To prove statement (ii), notice that if $u_2 \ge 0$, then $z \ge 0$ and $\mathsf{F}_{it}(u_1+z) - \mathsf{F}_{it}(u_1) \ge 0$ for all $z \in [0, u_2]$, which leads to the existence of the second inequality naturally:

$$\begin{aligned}&\int_0^{u_2} \{\mathsf{F}_{it}(u_1+z) - \mathsf{F}_{it}(u_1)\}\, dz - \int_0^{c_5 u_2} \{\mathsf{F}_{it}(u_1+z) - \mathsf{F}_{it}(u_1)\}\, dz \\ &= \int_{c_5 u_2}^{u_2} \{\mathsf{F}_{it}(u_1+z) - \mathsf{F}_{it}(u_1)\}\, dz \ge 0.\end{aligned}$$

On the other hand, if $u_2 < 0$, we have

$$\begin{aligned}&\int_0^{u_2} \{\mathsf{F}_{it}(u_1+z) - \mathsf{F}_{it}(u_1)\}\, dz - \int_0^{c_5 u_2} \{\mathsf{F}_{it}(u_1+z) - \mathsf{F}_{it}(u_1)\}\, dz \\ &= \int_{u_2}^{0} \{\mathsf{F}_{it}(u_1) - \mathsf{F}_{it}(u_1+z)\}\, dz - \int_{c_5 u_2}^{0} \{\mathsf{F}_{it}(u_1) - \mathsf{F}_{it}(u_1+z)\}\, dz \\ &= \int_{u_2}^{c_5 u_2} \{\mathsf{F}_{it}(u_1) 5\mathsf{F}_{it}(u_1+z)\}\, dz \ge 0,\end{aligned}$$

where the last inequality holds as the same reason for $u_2 \ge 0$ case. ∎

Lemma B.6 *Under Assumptions 1-4, for any $\{\Delta_{\Theta_j}\}_{j=0}^p \in \mathcal{R}(3, C_2)$ such that $||\Delta_{\Theta_j}||_{\max} \le M$ for some constant $M > 0$, we have*

$$Q_\tau\left(\{\Theta_j^0 + \Delta_{\Theta_j}\}_{j=0}^p\right) - Q_\tau\left(\{\Theta_j^0\}_{j=0}^p\right) \ge \frac{c_7 C_3}{NT\xi_N^2} \sum_{j=0}^p \left\|\Delta_{\Theta_j}\right\|_F^2 - \frac{c_7 C_4}{NT\xi_N^2}(N+T) \ \ w.p.a.1,$$

where $Q_\tau(\{\Theta_j\}_{j=0}^p) = \frac{1}{NT}\sum_{i=1}^N \sum_{t=1}^T \mathbb{E}\left[\rho_\tau\left(Y_{it} - \Theta_{0,it} - \sum_{j=1}^p X_{j,it}\Theta_{j,it}\right)\Big|\mathscr{G}_{i,t-1}\right]$, $c_7 = \frac{\mathsf{f}c_6^2}{4}$ with c_6 being a positive constant between 0 and 1.

Proof. We can choose a sufficiently large constant M such that $c_6 := \frac{3\underline{\mathsf{f}}}{2\bar{\mathsf{f}}' M(1+p)} \in (0,1]$. Then we have

$$\begin{aligned}&Q_\tau\left(\{\Theta_j^0 + \Delta_{\Theta_j}\}_{j=0}^p\right) - Q_\tau\left(\{\Theta_j^0\}_{j=0}^p\right) \\ &= \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \mathbb{E}\left\{\rho_\tau\left(Y_{it} - \Theta_{0,it} - \sum_{j=1}^p X_{j,it}\Theta_{j,it}\right) - \rho_\tau\left(Y_{it} - \Theta_{0,it}^0 - \sum_{j=1}^p X_{j,it}\Theta_{j,it}^0\right)\Bigg|\mathscr{G}_{i,t-1}\right\} \\ &= \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \mathbb{E}\left\{\left(\Delta_{\Theta_{0,it}} + \sum_{j=1}^p X_{j,it}\Delta_{\Theta_{j,it}}\right)(\tau - \mathbf{1}\{\epsilon_{it} \le 0\})\Bigg|\mathscr{G}_{i,t-1}\right\} \\ &+ \frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \mathbb{E}\left\{\int_0^{\Delta_{\Theta_{0,it}} + \sum_{j=1}^p X_{j,it}\Delta_{\Theta_{j,it}}} (\mathbf{1}\{\epsilon_{it} \le s\} - \mathbf{1}\{\epsilon_{it} \le 0\})\, ds\Bigg|\mathscr{G}_{i,t-1}\right\}\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}\left\{\int_0^{\Delta_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\Delta_{\Theta_{j,it}}}\left(\mathbf{1}_{\epsilon_{it}\le s}-\mathbf{1}_{\epsilon_{it}\le 0}\right)ds\middle|\mathscr{G}_{i,t-1}\right\}\\
&\ge \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbb{E}\left\{\int_0^{c_6\xi_N^{-1}\left(\Delta_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\Delta_{\Theta_{j,it}}\right)}\left(\mathbf{1}_{\epsilon_{it}\le s}-\mathbf{1}_{\epsilon_{it}\le 0}\right)ds\middle|\mathscr{G}_{i,t-1}\right\}\\
&= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\int_0^{c_6\xi_N^{-1}\left(\Delta_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\Delta_{\Theta_{j,it}}\right)}\left[\mathsf{F}_{it}(s)-\mathsf{F}_{it}(0)\right]ds\\
&= \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\int_0^{c_6\xi_N^{-1}\left(\Delta_{\Theta_{0,it}}+\sum_{j=1}^{0}X_{j,it}\Delta_{\Theta_{j,it}}\right)}\left[s\mathsf{f}_{it}(0)+\frac{s^2}{2}\mathsf{f}'_{it}(\tilde{s})\right]ds\\
&\ge \frac{\underline{\mathsf{f}}c_6^2}{2NT\xi_N^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\Delta_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\Delta_{\Theta_{j,it}}\right)^2-\frac{\bar{\mathsf{f}}'c_6^3}{6NT\xi_N^3}\sum_{i=1}^{N}\sum_{t=1}^{T}\left|\Delta_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\Delta_{\Theta_{j,it}}\right|^3\\
&= \frac{\underline{\mathsf{f}}c_6^2}{4NT\xi_N^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(\Delta_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\Delta_{\Theta_{j,it}}\right)^2\\
&+\frac{1}{NT\xi_N^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\frac{\underline{\mathsf{f}}c_6^2}{4}\left(\Delta_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\Delta_{\Theta_{j,it}}\right)^2\left(1-\frac{2c_6\bar{\mathsf{f}}'}{3\underline{\mathsf{f}}\xi_N}\left|\Delta_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\Delta_{\Theta_{j,it}}\right|\right)\right\}\\
&\ge \frac{\underline{\mathsf{f}}c_6^2}{4NT\xi_N^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\left\{\Delta_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\Delta_{\Theta_{j,it}}\right\}^2\\
&= \frac{\underline{\mathsf{f}}c_6^2}{4NT\xi_N^2}\left\|\Delta_{\Theta_0}+\sum_{j=1}^{p}X_j\odot\Delta_{\Theta_j}\right\|_F^2\\
&\ge \frac{\underline{\mathsf{f}}c_6^2}{4NT\xi_N^2}\left\{C_3\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F^2-C_4(N+T)\right\}\quad \text{w.p.a.1}\\
&= \frac{c_7C_3}{NT\xi_N^2}\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F^2-\frac{c_7C_4}{NT\xi_N^2}(N+T) \qquad \text{(B.11)}
\end{aligned}
$$

where the second equality is by (B.8), the first inequality is by Lemma B.5 and the fact that $c_6/\xi_N \le 1$, the fifth equality is by the mean-value theorem, the third inequality is by the fact that

$$
1-\frac{2c_6\bar{\mathsf{f}}'}{3\underline{\mathsf{f}}\xi_N}\left|\Delta_{\Theta_{0,it}}+\sum_{j=1}^{p}X_{j,it}\Delta_{\Theta_{j,it}}\right|\ge 1-\frac{2c_6\bar{\mathsf{f}}'}{3\underline{\mathsf{f}}\xi_N}M(1+p)\xi_N\ge 0,
$$

and the fourth inequality holds under Assumption 4. This concludes the proof. ■

Lemma B.7 *Under Assumptions 1–4, for any* $\left\{\Delta_{\Theta_j}\right\}_{j=0}^{p}\in\mathcal{R}(3,C_2)$, *we have* $\left\|\Delta_{\Theta_j}\right\|_*\le c_8\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F\ \ \forall j\in\{0,\cdots,p\}$ *where* $c_8=4\sqrt{2\bar{K}}$.

Proof. For $\forall j\in\{0,\cdots,p\}$, we obtain that

$$
\left\|\Delta_{\Theta_j}\right\|_*=\left\|\mathcal{P}_j(\Delta_{\Theta_j})\right\|_*+\left\|\mathcal{P}_j^{\perp}(\Delta_{\Theta_j})\right\|_*\le\left\|\mathcal{P}_j(\Delta_{\Theta_j})\right\|_*+3\sum_{j=0}^{p}\left\|\mathcal{P}_j(\Delta_{\Theta_j})\right\|_*
$$

$$
\begin{aligned}
&\leq 4\sum_{j=0}^{p}\left\|\mathcal{P}_j(\Delta_{\Theta_j})\right\|_* \leq 4\sum_{j=0}^{p}\sqrt{2K_j}\left\|\mathcal{P}_j\left(\Delta_{\Theta_j}\right)\right\|_F \\
&\leq 4\sqrt{2\bar{K}}\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F := c_8\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F,
\end{aligned}
$$

where the first equality is by Chernozhukov et al. (2019, Lemma D.2(i)), the first inequality is by the definition of $\mathcal{R}(C_1, C_2)$, and the last two inequalities follow the facts that $\|A\|_* \leq \sqrt{\text{rank}(A)}\,\|A\|_F$ for any matrix $A \in \mathbb{R}^{N\times T}$ and $\text{rank}\left(\mathcal{P}_j\left(\Delta_{\Theta_j}\right)\right) \leq 2K_j$, which hold by Chernozhukov et al. (2019, Lemma D.2.(iii)). ■

Let $\mathcal{Z}$ be a separable metric space, $\{Z_1, \cdots, Z_n\}$ be a sequence of random variables in $\mathcal{Z}$ adapted to the filtration $\{\mathscr{F}_t\}_{t\in[n]}$, $\mathcal{F} = \{f : \mathcal{Z} \to \mathbb{R}\}$ be a set of bounded real valued functions on $\mathbb{R}$, and $u_1, \cdots, u_n$ be i.i.d. Rademacher random variables. Here we allow the dependence of sequence $\{Z_1, \cdots, Z_n\}$. Similarly as in Rakhlin et al. (2015), we define a $\mathcal{Z}$-valued tree $\mathbf{z}$ of depth n with the sequence $(\mathbf{z}_1, \cdots, \mathbf{z}_n)$ such that $\mathbf{z}_t : \{u_1, \cdots, u_{t-1}\} \to \mathcal{Z}$. For simplicity, we denote as $\mathbf{z}_t(u) = \mathbf{z}_t(u_1, \cdots, u_{t-1})$ and $\mathbb{E}_t[\cdot] = \mathbb{E}\left[\cdot\middle|\mathscr{F}_t\right]$ for short. Also, denote $u_{1:t} := (u_1, \cdots, u_t)$ and similarly for $Z_{1:t}$.

Lemma B.8 *Let $\mathcal{F}$ be a class of functions. For any $\alpha > 0$, it holds that*

$$
\beta_n \mathbb{P}\left\{\sup_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{t=1}^{n}\left(f\left(Z_t\right) - \mathbb{E}_{t-1}\left[f\left(Z_t\right)\right]\right)\right| > \alpha\right\} \leq 2\sup_{\mathbf{z}}\mathbb{P}\left\{\sup_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{t=1}^{n}u_t f\left(\mathbf{z}_t(u)\right)\right| > \frac{\alpha}{4}\right\}.
$$

where $\beta_n \geq 1 - \sup_{f\in\mathcal{F}}\frac{4\sum_{t=1}^{n}Var\left(f(Z_t)\middle|\mathscr{F}_{t-1}\right)}{n^2\alpha^2}$ and the outer supremum is taken over all $\mathcal{Z}$-valued tree of depth n.[1]

Proof. Let $Z'_{1:n}$ be a decoupled sequence tangent to $Z_{1:n}$. For the sequence of random variables $\{Z_t : t \in [n]\}$ adapted to the filtration $\{\mathscr{F}_t : t \in [n]\}$, the sequence $Z'_{1:n} = \{Z_t^*, t \in [n]\}$ is said to be a decoupled sequence tangent to $\{Z_t : t \in [n]\}$ if for each $t \in [n]$, Z_t^* is generated from the conditional distribution of Z_t given $\mathscr{F}_{t-1}$ and independent of everything else. This means the sequence $Z'_{1:n}$ is conditionally independent given $\mathscr{F}_n$ and for any measurable function f of Z_t^*,

$$
\mathbb{E}(f(Z_t^*)|\mathscr{F}_n) = \mathbb{E}(f(Z_t^*)|\mathscr{F}_{t-1})\ a.s. \tag{B.12}
$$

For $\forall f \in \mathcal{F}$, with Chebyshev's inequality, we have

$$
\begin{aligned}
&\mathbb{P}\left(\frac{1}{n}\left|\sum_{t=1}^{n}f\left(Z_t^*\right) - \mathbb{E}_{t-1}\left[f\left(Z_t^*\right)\right]\right| > \alpha/2\middle|\mathscr{F}_n\right) \\
&\leq \frac{4\mathbb{E}\left\{\left(\sum_{t=1}^{n}f\left(Z_t^*\right) - \mathbb{E}_{t-1}\left[f\left(Z_t^*\right)\right]\right)^2\middle|\mathscr{F}_n\right\}}{n^2\alpha^2} \\
&= \frac{4\sum_{t=1}^{n}\mathbb{E}\left\{\left(f\left(Z_t^*\right) - \mathbb{E}_{t-1}\left[f\left(Z_t^*\right)\right]\right)^2\middle|\mathscr{F}_n\right\}}{n^2\alpha^2}
\end{aligned}
$$

[1] This means the supremum is taken over all $\mathbf{z} = \{\mathbf{z}_t(\cdot)\}_{t\in[n]}$.

$$= \frac{4 \sum_{t=1}^{n} Var\left(f\left(Z_t\right) \middle| \mathscr{F}_{t-1}\right)}{n^2 \alpha^2},$$

where the first equality holds by the fact that, given $\mathscr{F}_n$, $\{Z_t^*\}_{t\in[n]}$ are independent and the last equality holds by (B.12). This implies

$$\begin{aligned}
\beta_n &:= \inf_{f\in\mathcal{F}} \mathbb{P}\left(\frac{1}{n}\left|\sum_{t=1}^{n} f\left(Z_t^*\right) - \mathbb{E}_{t-1}\left[f\left(Z_t^*\right)\right]\right| \leq \alpha/2 \middle| \mathscr{F}_n\right) \\
&= 1 - \sup_{f\in\mathcal{F}} \mathbb{P}\left(\frac{1}{n}\left|\sum_{t=1}^{n} f\left(Z_t^*\right) - \mathbb{E}_{t-1}\left[f\left(Z_t^*\right)\right]\right| > \alpha/2 \middle| \mathscr{F}_n\right) \\
&\geq 1 - \sup_{f\in\mathcal{F}} \frac{4 \sum_{t=1}^{n} Var\left(f\left(Z_t\right) \middle| \mathscr{F}_{t-1}\right)}{n^2 \alpha^2}.
\end{aligned}$$

Let function f^* be the function that maximizes $\frac{1}{n}\left|\sum_{t=1}^{n}\left[f\left(Z_t^*\right) - \mathbb{E}_{t-1}\left(f\left(Z_t^*\right)\right)\right]\right|$ condition on $\mathscr{F}_n$, and define the event $A_1 = \left\{\sup_{f\in\mathcal{F}} \frac{1}{n}\left|\sum_{t=1}^{n}\left[f\left(Z_t\right) - \mathbb{E}_{t-1}\left(f\left(Z_t\right)\right)\right]\right| > \alpha\right\}$. Then we obtain that

$$\beta_n \leq \mathbb{P}\left(\frac{1}{n}\left|\sum_{t=1}^{n} f^*\left(Z_t^*\right) - \mathbb{E}_{t-1}\left[f^*\left(Z_t\right)\right]\right| \leq \alpha/2 \middle| \mathscr{F}_n\right),$$

where the inequality follows by the definition of β_n and the fact that $\mathbb{E}_{t-1}\left[f^*\left(Z_t^*\right)\right] = \mathbb{E}_{t-1}\left[f^*\left(Z_t\right)\right]$. As $A_1 \in \mathscr{F}_n$, we have

$$\beta_n \leq \mathbb{P}\left(\frac{1}{n}\left|\sum_{t=1}^{n} f^*\left(Z_t^*\right) - \mathbb{E}_{t-1}\left[f^*\left(Z_t^*\right)\right]\right| \leq \alpha/2 \middle| A_1\right).$$

It follows that

$$\begin{aligned}
&\beta_n \mathbb{P}\left\{\sup_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{t=1}^{n}\left(f\left(Z_t\right) - \mathbb{E}_{t-1}\left[f\left(Z_t\right)\right]\right)\right| > \alpha\right\} \\
&\leq \mathbb{P}\left(\frac{1}{n}\left|\sum_{t=1}^{n} f^*\left(Z_t^*\right) - \mathbb{E}_{t-1}\left[f^*\left(Z_t^*\right)\right]\right| \leq \alpha/2 \middle| A_1\right) \mathbb{P}\left(A_1\right) \\
&= \mathbb{P}\left(\left\{\frac{1}{n}\left|\sum_{t=1}^{n} f^*\left(Z_t^*\right) - \mathbb{E}_{t-1}\left[f^*\left(Z_t^*\right)\right]\right| \leq \alpha/2\right\} \bigcup A_1\right) \\
&\leq \mathbb{P}\left(\frac{1}{n}\left|\sum_{t=1}^{n} f^*\left(Z_t\right) - f^*\left(Z_t^*\right)\right| \geq \alpha/2\right) \\
&\leq \mathbb{P}\left(\frac{1}{n}\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^{n} f\left(Z_t\right) - f\left(Z_t^*\right)\right| \geq \alpha/2\right), \qquad \text{(B.13)}
\end{aligned}$$

where the second inequality holds by the implication rule. Let $\phi(\cdot) = \mathbf{1}\{\cdot > n\alpha/2\}$. By Lemma 18 in Rakhlin et al. (2015), we have

$$\begin{aligned}
&\mathbb{P}\left(\frac{1}{n}\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^{n} f\left(Z_t\right) - f\left(Z_t^*\right)\right| \geq \alpha/2\right) \\
&= \mathbb{E}\mathbf{1}\left\{\frac{1}{n}\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^{n} f\left(Z_t\right) - f\left(Z_t^*\right)\right| \geq \alpha/2\right\} \\
&\leq \sup_{z_1,z_1'} \mathbb{E}_{u_1} \cdots \sup_{z_n,z_n'} \mathbb{E}_{u_n} \mathbf{1}\left\{\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^{n} u_t\left[f\left(z_t\right) - f\left(Z_t^*\right)\right]\right| \geq n\alpha/2\right\}
\end{aligned}$$

$$\leq \sup_{z_1,z_1'} \mathbb{E}_{u_1} \cdots \sup_{z_n,z_n'} \mathbb{E}_{u_n} \mathbf{1}\left\{\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^n u_t f(z_t)\right| \geq n\alpha/4\right\} + \sup_{z_1,z_1'} \mathbb{E}_{u_1} \cdots \sup_{z_n,z_n'} \mathbb{E}_{u_n} \mathbf{1}\left\{\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^n u_t f(Z_t^*)\right| \geq n\alpha/4\right\}$$
$$= 2\sup_{z_1,z_1'} \mathbb{E}_{u_1} \cdots \sup_{z_n,z_n'} \mathbb{E}_{u_n} \mathbf{1}\left\{\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^n u_t f(z_t)\right| \geq n\alpha/4\right\}$$
$$= 2\sup_{\mathbf{z}} \mathbb{P}\left\{\sup_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{t=1}^n u_t f(\mathbf{z}_t(u))\right| > \frac{\alpha}{4}\right\}, \tag{B.14}$$

where the standard text z_t and Z_t^* are $\mathcal{Z}$-valued, the bold text $\mathbf{z}_t = \mathbf{z}_t(u)$ is the t-th root of a tree $\mathbf{z}(\cdot)$ (i.e., a function of $u_{1:t-1}$), and the outer supremum in the last line is taken over all $\mathcal{Z}$-valued tree of depth n. Combining (B.13) and (B.14), we can conclude that

$$\beta_n \mathbb{P}\left\{\sup_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{t=1}^n \left(f(Z_t) - \mathbb{E}_{t-1}[f(Z_t)]\right)\right| > \alpha\right\} \leq 2\sup_{\mathbf{z}} \mathbb{P}\left\{\sup_{f\in\mathcal{F}}\left|\frac{1}{n}\sum_{t=1}^n u_t f(\mathbf{z}_t(u))\right| > \frac{\alpha}{4}\right\}.$$

■

The next lemma is an extension of the contraction principle, i.e., Ledoux and Talagrand (1991, Theorem 4.12), to the case with sequential symmetrization.

Lemma B.9 *Let function $F : \mathbb{R}_+ \to \mathbb{R}_+$ be convex and increasing and $\phi_t : \mathbb{R} \to \mathbb{R}$ be contractions such that $\phi_t(0) = 0$, $\mathbf{z}_t$ is the t-th root of a tree $(\mathbf{z})$ which depends on $\{u_1, \cdots, u_{t-1}\}$. Then we have*

$$\mathbb{E}F\left\{\frac{1}{2}\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^n u_t\phi_t(f(\mathbf{z}_t(u)))\right|\right\} \leq \mathbb{E}F\left\{\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^n u_t f(\mathbf{z}_t(u))\right|\right\}.$$

Proof. We first consider the statement without the absolute value. Let function $G : \mathbb{R} \to \mathbb{R}$ be convex and increasing. We observe that

$$\mathbb{E}G\left\{\sup_{f\in\mathcal{F}}\sum_{t=1}^n u_t\phi_t(f(\mathbf{z}_t(u)))\right\} = \mathbb{E}\left\{\mathbb{E}\left[G\left\{\sup_{f\in\mathcal{F}}\sum_{t=1}^n u_t\phi_t(f(\mathbf{z}_t(u)))\right\}\middle| u_{1:n-1}\right]\right\}$$

and

$$\mathbb{E}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^n u_t\phi_t(f(\mathbf{z}_t(u)))\middle| u_{1:n-1}\right)\right\} = \mathbb{E}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n-1} u_t\phi_t(f(\mathbf{z}_t(u))) + u_n\phi_n(f(\mathbf{z}_n(u)))\middle| u_{1:n-1}\right)\right\}$$
$$= \mathbb{E}\left\{G\left(\sup_{k_1,k_2\in\mathcal{K}}(k_1 + u_n\phi_n(k_2))\right)\middle| u_{1:n-1}\right\},$$

where $k_1 = \sum_{t=1}^{n-1} u_t\phi_t(f(\mathbf{z}_t(u)))$, $k_2 = f(\mathbf{z}_n(u))$, and $\mathcal{K} = \{(k_1, k_2) : f \in \mathcal{F}\} \subset \mathbb{R}^2$. We also note that k_1 and k_2 only depend on $u_{1:n-1}$ and is independent of u_n. The proof in Ledoux and Talagrand (1989, Theorem 4.12) shows

$$\mathbb{E}G\left\{\sup_{k_1,k_2\in\mathcal{K}} k_1 + u_2\phi_2(k_2)\right\} \leq \mathbb{E}G\left\{\sup_{k_1,k_2\in\mathcal{K}} k_1 + u_2 k_2\right\},$$

which implies

$$\mathbb{E}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^n u_t\phi_t(f(\mathbf{z}_t(u)))\middle| u_{1:n-1}\right)\right\} \leq \mathbb{E}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n-1} u_t\phi_t(f(\mathbf{z}_t(u))) + u_n f(\mathbf{z}_n(u))\middle| u_{1:n-1}\right)\right\}.$$

Taking expectation on both sides, we have

$$\mathbb{E}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)\right)\right\}\leq\mathbb{E}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n-1}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)+u_nf\left(\mathbf{z}_n(u)\right)\right)\right\}.$$

Next, let $k_1=\sum_{t=1}^{n-2}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)$, $k_2=f(\mathbf{z}_{n-1}(u))$ and $k_3(u_{n-1})=u_nf(\mathbf{z}_n(u))$. We emphasize that k_1 and k_2 only depend on $u_{1:n-2}$ and u_n while k_3 also depends on u_{n-1}. The dependence of $(k_1,k_2,k_3(\cdot))$ on f is made implicit for notation simplicity. Furthermore, given the fact that u_{n-1} only takes values {-1,1}, we have

$$k_3(u_{n-1})=\frac{u_{n-1}+1}{2}k_3(1)+\frac{1-u_{n-1}}{2}k_3(-1)=\frac{k_3(1)+k_3(-1)}{2}+u_{n-1}\frac{k_3(1)-k_3(-1)}{2}.$$

Given these notation and conditioning on $(u_{1:n-2},u_n)$, we have

$$\begin{aligned}
&\mathbb{E}_{u_{n-1}}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n-1}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)+u_nf\left(\mathbf{z}_n(u)\right)\right)\right\}\\
&=\mathbb{E}_{u_{n-1}}\left[G\left(\sup_{f\in\mathcal{F}}k_1+u_{n-1}\phi_{n-1}(k_2)+k_3(u_{n-1})\right)\right]\\
&=\mathbb{E}_{u_{n-1}}\left[G\left(\sup_{f\in\mathcal{F}}\left(k_1+\frac{k_3(1)-k_3(-1)}{2}\right)+u_{n-1}\left(\phi_{n-1}(k_2)+\frac{k_3(1)-k_3(-1)}{2}\right)\right)\right]\\
&=\mathbb{E}_{u_{n-1}}\left[G\left(\sup_{(h_1,h_2,h_3)\in\mathcal{H}}h_1+u_{n-1}\left(\phi_{n-1}(h_2)+h_3\right)\right)\right],
\end{aligned}$$

where $\mathbb{E}_{u_{n-1}}$ means the expectation is taken conditionally on $(u_{1:n-2},u_n)$, $h_1=\sum_{t=1}^{n-2}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)+\frac{u_n(f(\mathbf{z}_n(u_{1:n-2},1))+f(\mathbf{z}_n(u_{1:n-2},-1)))}{2}$, $h_2=k_2$, $h_3=\frac{u_n(f(\mathbf{z}_n(u_{1:n-2},1))-f(\mathbf{z}_n(u_{1:n-2},-1)))}{2}$, and $\mathcal{H}=((h_1,h_2,h_3):f\in\mathcal{F})\in\mathbb{R}^3$. Suppose $(h_1^*,h_2^*,h_3^*)\in\mathcal{H}$ and $(h_1^\dagger,h_2^\dagger,h_3^\dagger)\in\mathcal{H}$ achieve the supremum of

$$h_1+\left(\phi_{n-1}(h_2)+h_3\right)\quad\text{and}\quad h_1-\left(\phi_{n-1}(h_2)+h_3\right),\quad\text{respectively.}$$

Then, we have

$$\begin{aligned}
&\mathbb{E}_{u_{n-1}}\left[G\left(\sup_{(h_1,h_2,h_3)\in\mathcal{H}}h_1+u_{n-1}\left(\phi_{n-1}(h_2)+h_3\right)\right)\right]\\
&=\frac{1}{2}G((h_1^*+h_3^*)+\phi_{n-1}(h_2^*))+\frac{1}{2}G((h_1^\dagger-h_3^\dagger)-\phi_{n-1}(h_2^\dagger))\\
&\leq\frac{1}{2}G((h_1^*+h_3^*)+h_2^*)+\frac{1}{2}G((h_1^\dagger-h_3^\dagger)-h_2^\dagger)\\
&\leq\mathbb{E}_{u_{n-1}}\left[G\left(\sup_{(h_1,h_2,h_3)\in\mathcal{H}}h_1+u_{n-1}\left(h_2+h_3\right)\right)\right],
\end{aligned}$$

where the first inequality is by the fact proved in the proof of Ledoux and Talagrand (1991, Theorem 4.12) that for any t_1,s_1,t_2,s_2,

$$\frac{1}{2}G(t_1+\phi_{n-1}(t_2))+\frac{1}{2}G(s_1-\phi_{n-1}(s_2))\leq\frac{1}{2}G(t_1+t_2)+\frac{1}{2}G(s_1-s_2).$$

Plugging back the definition of h_1, h_2, h_3, we have

$$\mathbb{E}_{u_{n-1}}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n-1}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)+u_nf\left(\mathbf{z}_n(u)\right)\right)\right\}\leq\mathbb{E}_{u_{n-1}}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n-2}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)+\sum_{t=n-1}^{n}u_tf\left(\mathbf{z}_t(u)\right)\right)\right\}.$$

Taking expectation on both sides, we have

$$\mathbb{E}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n-1}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)+u_nf\left(\mathbf{z}_n(u)\right)\right)\right\}\leq\mathbb{E}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n-2}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)+\sum_{t=n-1}^{n}u_tf\left(\mathbf{z}_t(u)\right)\right)\right\}.$$

We can repeat a similar argument by taking conditional expectations given $(u_{1:t-1}, u_{t+1:n})$ and removing ϕ_t for all $t = n-2, \cdots, 2$. This leads to the result that

$$\mathbb{E}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)\right)\right\}\leq\mathbb{E}\left\{G\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n}u_tf\left(\mathbf{z}_t(u)\right)\right)\right\}. \tag{B.15}$$

Next, we come back to the case with the absolute value. Note that

$$\begin{aligned}
&\mathbb{E}F\left[\frac{1}{2}\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^{n}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)\right|\right]\\
&\leq\frac{1}{2}\left\{\mathbb{E}F\left[\sup_{f\in\mathcal{F}}\left(\sum_{t=1}^{n}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)\right)^+\right]+\mathbb{E}F\left[\sup_{f\in\mathcal{F}}\left(\sum_{t=1}^{n}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)\right)^-\right]\right\}\\
&=\frac{1}{2}\left\{\mathbb{E}F\left[\sup_{f\in\mathcal{F}}\left(\sum_{t=1}^{n}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)\right)^+\right]+\mathbb{E}F\left[\sup_{f\in\mathcal{F}}\left(\sum_{t=1}^{n}u_t^*\phi_t\left(f\left(\mathbf{z}_t^*(u^*)\right)\right)\right)^+\right]\right\}\\
&=\frac{1}{2}\left\{\mathbb{E}F\left[\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n}u_t\phi_t\left(f\left(\mathbf{z}_t(u)\right)\right)\right)^+\right]+\mathbb{E}F\left[\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n}u_t^*\phi_t\left(f\left(\mathbf{z}_t^*(u^*)\right)\right)\right)^+\right]\right\}\\
&\leq\frac{1}{2}\left\{\mathbb{E}F\left[\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n}u_t\left(f\left(\mathbf{z}_t(u)\right)\right)\right)^+\right]+\mathbb{E}F\left[\left(\sup_{f\in\mathcal{F}}\sum_{t=1}^{n}u_t^*f\left(\mathbf{z}_t^*(u^*)\right)\right)^+\right]\right\}\\
&=\frac{1}{2}\left\{\mathbb{E}F\left[\sup_{f\in\mathcal{F}}\left(\sum_{t=1}^{n}u_t\left(f\left(\mathbf{z}_t(u)\right)\right)\right)^+\right]+\mathbb{E}F\left[\sup_{f\in\mathcal{F}}\left(\sum_{t=1}^{n}u_t^*f\left(\mathbf{z}_t^*(u^*)\right)\right)^+\right]\right\}\\
&=\frac{1}{2}\left\{\mathbb{E}F\left[\sup_{f\in\mathcal{F}}\left(\sum_{t=1}^{n}u_t\left(f\left(\mathbf{z}_t(u)\right)\right)\right)^+\right]+\mathbb{E}F\left[\sup_{f\in\mathcal{F}}\left(\sum_{t=1}^{n}u_tf\left(\mathbf{z}_t(u)\right)\right)^-\right]\right\}\\
&\leq\mathbb{E}F\left[\sup_{f\in\mathcal{F}}\left|\sum_{t=1}^{n}u_tf\left(\mathbf{z}_t(u)\right)\right|\right],
\end{aligned} \tag{B.16}$$

where $u_t^* = -u_t$, $\mathbf{z}_t^*(u) = \mathbf{z}_t(-u)$, the first inequality is by the convexity of F, the first and second equalities are by the fact that $(v)^- = (-v)^+$ for any v, and the second inequality is by (B.15) and the fact that $F((\cdot)^+)$ is convex and increasing. This leads to the desired result. ■

Define Rademacher sequence $u = (u_{11}, \cdots, u_{1T}, \cdots, u_{N1}, \cdots, u_{NT}) = \left(u_{(1)}, \cdots, u_{(NT)}\right) \in \mathbb{R}^{NT\times 1}$. In the matrix notation, let $U = \{u_{it}\} \in \mathbb{R}^{N\times T}$. By vectorization, for a sequence of independent variables

$X_{j,it}$, we define

$$\left(X_{j,11},\cdots,X_{j,1T},\cdots,X_{j,N1},\cdots,X_{j,NT}\right)=\left(X_{j,(1)},\cdots,X_{j,(NT)}\right),$$
$$\left(\epsilon_{11},\cdots,\epsilon_{1T},\cdots,\epsilon_{N1},\cdots,\epsilon_{NT}\right)=\left(\epsilon_{(1)},\cdots,\epsilon_{(NT)}\right),$$

$\forall j\in[p]$. Using the binary tree representation, let $x^*_{j,(l)}$ be $(l)^{th}$ root of the tree which takes values in the support of $X_{j,(l)}$, i.e., $x^*_{j,(l)}:u^{l-1}\mapsto[-\xi_N,\xi_N]$ for $u^{l-1}:=\left(u_{(1)},\cdots,u_{(l-1)}\right)$ such that $\max_{i\in[N]}\sum_{t\in[T]}\left(x^*_{j,it}\right)^2\leq MT$ and $\max_{t\in[T]}\sum_{i\in[N]}\left(x^*_{j,it}\right)^{2\ell}\leq MN$ for some fixed constant $M<\infty$ and $\ell=1,2$. Similar notation follows for $\epsilon^*_{(l)}=\epsilon^*_{(l)}\left(u^{l-1}\right)$. In the matrix notation, let $x^*_j=\left\{x^*_{j,it}\right\}\in\mathbb{R}^{N\times T}$ such that $x^*_{j,it}=x_{j,(l)}$ with $i=\lceil\frac{l}{T}\rceil$ and $t=l-(i-1)T$.

Lemma B.10 *Under Assumption 1, for $j\in[p]$, there exists an absolute constant C that is independent of the trees (x^*,ϵ^*) such that when $\log(N\vee T)\geq 2$,*

$$\mathbb{E}\exp\left(\frac{\|U\|_{op}}{\sqrt{(N\vee T)\log(N\vee T)}}\right)\leq C\quad\textit{and}\quad\mathbb{E}\exp\left(\frac{\left\|U\odot x^*_j\right\|_{op}}{\sqrt{(N\vee T)\log(N\vee T)}}\right)\leq C.$$

Proof. The proof here follows similarly as Lemma B.2 except that we have martingale difference matrices rather than independent matrices. For a specific j, let $A=U\odot x^*_j=(A_1,\cdots,A_N)'\in\mathbb{R}^{N\times T}$, $\mathscr{F}_i$ be the σ-field generated by $\{u_{i^*t}\}_{i^*\leq i,t\in[T]}$, $\mathbb{E}_i(\cdot)=\mathbb{E}(\cdot|\mathscr{F}_i)$, $\Sigma_i=\mathbb{E}_{i-1}\left(A_iA_i'\right)$ and $Z_i=\frac{1}{N}\left(A_iA_i'-\Sigma_i\right)$, with $\frac{1}{N}\left(A'A-\sum_{i\in[N]}\Sigma_i\right)=\sum_{i\in[N]}Z_i$. Note that $\mathbb{E}_{i-1}\left(Z_i\right)=0$,

$$\max_{i\in[N]}\|A_i\|_2=\max_{i\in[N]}\sqrt{\sum_{t\in[T]}\left(x^*_{j,it}u_{it}\right)^2}=\max_{i\in[N]}\sqrt{\sum_{t\in[T]}\left(x^*_{j,it}\right)^2}\leq\sqrt{MT}\ a.s.,\tag{B.17}$$

$$\max_{i\in[N]}\|\Sigma_i\|_{op}=\max_{i\in[N]}\left\|\text{diag}\left(\left(x^*_{j,i1}\right)^2,\cdots,\left(x^*_{j,iT}\right)^2\right)\right\|_{op}\leq\xi_N^2\ a.s.,\tag{B.18}$$

and for $\ell=1,2$,

$$\left\|\sum_{i\in[N]}\Sigma_i^\ell\right\|_{op}=\left\|\text{diag}\left(\sum_{i\in[N]}\left(x^*_{j,i1}\right)^{2\ell},\cdots,\sum_{i\in[N]}\left(x^*_{j,iT}\right)^{2\ell}\right)\right\|_{op}\leq MN\ a.s..\tag{B.19}$$

Combining (B.17) and (B.18) yields

$$\begin{aligned}\max_{i\in[N]}\|Z_i\|_{op}&\leq\max_{i\in[N]}\frac{1}{N}\left(\|A_iA_i'\|_{op}+\|\Sigma_i\|_{op}\right)\\&\leq\frac{1}{N}\left(\max_{i\in[N]}\|A_i\|_2^2+\max_{i\in[N]}\|\Sigma_i\|_{op}\right)\leq\frac{MT+\xi_N^2}{N}\ a.s.\end{aligned}\tag{B.20}$$

In addition,

$$\begin{aligned}\left\|\sum_{i\in[N]}\mathbb{E}_{i-1}\left(Z_i^2\right)\right\|_{op}&=\left\|\sum_{i\in[N]}\mathbb{E}_{i-1}\left\{\frac{1}{N^2}\left[\left(A_iA_i'\right)^2-\Sigma_i^2\right]\right\}\right\|_{op}\\&\leq\frac{1}{N^2}\left[\sum_{i\in[N]}\left\|\mathbb{E}_{i-1}\left(\|A_i\|_2^2A_iA_i'\right)\right\|_{op}+\left\|\sum_{i\in[N]}\Sigma_i^2\right\|_{op}\right]\end{aligned}$$

$$\leq \frac{1}{N^2}\left[MT\left\|\sum_{i\in[N]}\Sigma_i\right\|_{op}+\left\|\sum_{i\in[N]}\Sigma_i^2\right\|_{op}\right]$$

$$\leq \frac{(MT+1)\,M}{N}, \tag{B.21}$$

where the last inequality holds by (B.17) and (B.19).

Combining (B.20) and (B.21), by matrix Bernstein's inequality in Lemma B.1(ii), for some sufficiently large constant $\bar{c}$ that depends on M, we have

$$\mathbb{P}\left(\left\|\sum_{i\in[N]}Z_i\right\|_{op}>\bar{c}\frac{(N\vee T)}{N}\log(N\vee T)\right)$$
$$\leq T\exp\left\{-\frac{\frac{1}{2}\bar{c}^2\left(\frac{(N\vee T)}{N}\right)^2(\log(N\vee T))^2}{\frac{(MT+1)M}{N}+\frac{\frac{MT+\xi_N^2}{N}\bar{c}\frac{(N\vee T)}{N}\log(N\vee T)}{3}}\right\}=\exp\left(-\left(\frac{\bar{c}}{2}-1\right)\log((N\vee T))\right),$$

which implies that with probability greater than $1-\exp(-(\bar{c}/2-1)\log(N\vee T))$,

$$\left\|\frac{1}{N}\left(A'A-\sum_{i\in[N]}\Sigma_i\right)\right\|_{op}\leq\bar{c}\left(\frac{(N\vee T)}{N}\log(N\vee T)\right),$$

$$\left\|\frac{1}{N}A'A\right\|_{op}\leq\left\|\frac{1}{N}\sum_{i\in[N]}\Sigma_i\right\|_{op}+\left\|\frac{1}{N}\left(A'A-\sum_{i\in[N]}\Sigma_i\right)\right\|_{op}\leq M+\bar{c}\left(\frac{(N\vee T)}{N}\log(N\vee T)\right),\quad\text{and}$$

$$\|A\|_{op}=\left\|U\odot x_j^*\right\|_{op}\leq\sqrt{(1+\bar{c})(N\vee T)\log(N\vee T)}.$$

Consequently, when $\log(N\vee T)\geq 2$, we have

$$\mathbb{E}\exp\left(\frac{\left\|U\odot x_j^*\right\|_{op}}{\sqrt{(N\vee T)\log(N\vee T)}}\right)$$
$$=\int_0^\infty\exp(u)\mathbb{P}\left(\frac{\left\|U\odot x_j^*\right\|_{op}}{\sqrt{(N\vee T)\log(N\vee T)}}\geq u\right)du$$
$$=\left(\int_0^2+\int_2^\infty\right)\exp(u)\mathbb{P}\left(\left\|U\odot x_j^*\right\|_{op}\geq u\sqrt{(N\vee T)\log(N\vee T)}\right)du$$
$$\leq\int_0^2\exp(u)du+\int_2^\infty\exp(u)\exp\left(-\frac{(u^2-3)}{2}\log(N\vee T)\right)du$$
$$\leq\int_0^2\exp(u)du+\int_2^\infty\exp\left(-\left(u-\frac{1}{2}\right)^2+\frac{13}{4}\right)du$$
$$\leq\exp(2)-1+\sqrt{2\pi}\exp\left(\frac{13}{4}\right):=C,$$

where the first inequality is by the fact that

$$\mathbb{P}\left(\frac{\left\|U\odot x_j^*\right\|_{op}}{\sqrt{(N\vee T)\log(N\vee T)}}\geq\sqrt{1+\bar{c}}\right)\leq\exp\left(-\left(\frac{\bar{c}}{2}-1\right)\log((N\vee T))\right)$$

and by letting $u = \sqrt{1+\bar{c}} \geq 2$ for large $\bar{c}$. Similarly, we can show that when $\log(N \vee T) \geq 2$, $\mathbb{E}\exp\left(\frac{\|U\|_{op}}{\sqrt{(N\vee T)\log(N\vee T)}}\right) \leq C$ for some absolute constant C. ■

Recall that $\tilde{\rho}_{it}\left(\left\{\Delta_{\Theta_j,it}, X_{j,it}\right\}_{j=0}^{p}, \epsilon_{it}\right)$ is defined in (A.4) and $\mathscr{G}_{i,t}$ is defined in Assumption 1.

Lemma B.11 *If Assumptions 1-4 hold, then we have*

$$\sup_{\left\{\Delta_{\Theta_j}\right\}_{j=0}^{p} \in \mathcal{R}(3,C_2)} \frac{\left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{\rho}_{it}\left(\left\{\Delta_{\Theta_j,it}, X_{j,it}\right\}_{j=0}^{p}, \epsilon_{it}\right)\right|}{\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F} = O_p\left(a_{NT}\right),$$

where $a_{NT} = \frac{\sqrt{(N\vee T)\log(N\vee T)}}{NT}$.

Proof. Let $n = NT$ and for $l \in [n]$, $Z_l = (\{X_{j,it}\}_{j\in[p]}, \epsilon_{it})$ with $i = \lceil \frac{l}{T} \rceil$ and $t = l-(i-1)T$, and $\mathscr{F}_l = \mathscr{G}_{it}$. Then, Z_l is adapted to the filtration $\{\mathscr{F}_l\}_{l\in[n]}$. Lemma B.8 implies

$$\begin{aligned}
&\beta_{NT}\mathbb{P}\left\{\sup_{\left\{\Delta_{\Theta_j}\right\}_{j=0}^{p} \in \mathcal{R}(3,C_2)} \left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\tilde{\rho}_{it}\left(\left\{\Delta_{\Theta_j,it}, X_{j,it}\right\}_{j=0}^{p}, \epsilon_{it}\right)}{\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F}\right| > C_5 a_{NT}\right\} \\
\leq & 2\sup_{x^*,\epsilon^*}\mathbb{P}\left\{\sup_{\left\{\Delta_{\Theta_j}\right\}_{j=0}^{p} \in \mathcal{R}(3,C_2)} \left|\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\frac{u_{it}\rho_{it}\left(\left\{\Delta_{\Theta_j,it}, x^*_{j,it}\right\}_{j=0}^{p}, \epsilon^*_{it}\right)}{\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F}\right| > \frac{C_5 a_{NT}}{4}\right\}, \qquad \text{(B.22)}
\end{aligned}$$

for some positive constant C_5, where the outer supremum on the RHS of the above display is taken over $[-\xi_N, \xi_N]^p \times \mathbb{R}$-valued trees with depth n and

$$\beta_{NT} = 1 - \sup_{\left\{\Delta_{\Theta_j}\right\}_{j=0}^{p} \in \mathcal{R}(3,C_2)} \frac{4}{C_5^2 (NT)^2 a_{NT}^2}\sum_{(i,t)\in[N]\times[T]} Var\left(\frac{\tilde{\rho}_{it}\left(\left\{\Delta_{\Theta_j,it}, X_{j,it}\right\}_{j=0}^{p}, \epsilon_{it}\right)}{\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F}\middle|\mathscr{G}_{i,t-1}\right).$$

We first note

$$\begin{aligned}
&\sum_{(i,t)\in[N]\times[T]} Var\left(\frac{\tilde{\rho}_{it}\left(\left\{\Delta_{\Theta_j,it}, X_{j,it}\right\}_{j=0}^{p}, \epsilon_{it}\right)}{\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F}\middle|\mathscr{G}_{i,t-1}\right) \\
\leq & \sum_{(i,t)\in[N]\times[T]} \mathbb{E}\left[\left(\frac{\rho_{it}\left(\left\{\Delta_{\Theta_j,it}, X_{j,it}\right\}_{j=0}^{p}, \epsilon_{it}\right)}{\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F}\right)^2\middle|\mathscr{G}_{i,t-1}\right] \\
\leq & 2\sum_{(i,t)\in[N]\times[T]}\left(\frac{\Delta_{\Theta_0,it} + \sum_{j=1}^{p}X_{j,it}\Theta_{j,it}}{\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F}\right)^2 \\
\lesssim & \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\sum_{j=1}^{p}\left(X_{j,it}\Delta_{\Theta_j,it}\right)^2 + \Delta_{\Theta_0,it}^2}{\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F^2} \leq c_{10}\xi_N^2
\end{aligned}$$

with some positive constant c_{10}, where the first inequality holds by Jensen inequality, the second inequality is by $|\rho_\tau(u) - \rho_\tau(v)| \leq 2|u-v|$, and the last line holds by Assumption 1(iv) and the fact that $\left(\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F\right)^2 = O\left(\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F^2\right)$. Therefore, we have

$$\beta_{NT} \geq 1 - \sup_{\left\{\Delta_{\Theta_j}\right\}_{j=0}^{p} \in \mathcal{R}(3,C_2)} \frac{4}{C_5^2 (NT)^2 a_{NT}^2}\sum_{(i,t)\in[N]\times[T]} Var\left(\frac{\tilde{\rho}_{it}\left(\left\{\Delta_{\Theta_j,it}, X_{j,it}\right\}_{j=0}^{p}, \epsilon_{it}\right)}{\sum_{j=0}^{p}\left\|\Delta_{\Theta_j}\right\|_F}\right)$$

$$\geq 1-\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p\in\mathcal{R}(3,C_2)}\frac{4c_{10}\xi_N^2}{C_5^2\,(N\vee T)\log(N\vee T)}$$

$$\geq 1-O\left(\frac{\xi_N^2}{(N\vee T)\log(N\vee T)}\right)\to 1, \tag{B.23}$$

where the last line is by Assumption 1(ix).

Define

$$\mathcal{A}_0=\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p\in\mathcal{R}(3,C_2)}\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\frac{u_{it}\Delta_{\Theta_0,it}}{\sum_{j=0}^p\left\|\Delta_{\Theta_j}\right\|_F}\right|,$$

$$\mathcal{A}_j=\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p\in\mathcal{R}(3,C_2)}\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\frac{u_{it}x_{j,it}^*\Delta_{\Theta_j,it}}{\sum_{j=0}^p\left\|\Delta_{\Theta_j}\right\|_F}\right|,\quad \forall j\in[p],$$

$$\mathcal{A}_{p+1}=\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p\in\mathcal{R}(3,C_2)}\left|\frac{\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T u_{it}\phi_{it}\left(\Delta_{\Theta_0,it}+\sum_{j=1}^p x_{j,it}^*\Delta_{\Theta_j,it}\right)}{\sum_{j=0}^p\left\|\Delta_{\Theta_j}\right\|_F}\right|,$$

where $\phi_{it}(u)=(\epsilon_{it}^*-u)^- -(\epsilon_{it}^*)^-$. Notice that

$$\rho_{it}\left(\left\{\Delta_{\Theta_j,it},x_{j,it}^*\right\}_{j=0}^p,\epsilon_{it}^*\right)=\tau\left(\Delta_{\Theta_0,it}+\sum_{j=1}^p x_{j,it}^*\Delta_{\Theta_j,it}\right)+\phi_{it}\left(\Delta_{\Theta_0,it}+\sum_{j=1}^p x_{j,it}^*\Delta_{\Theta_j,it}\right),$$

we obtain that

$$\sup_{x^*,\epsilon^*}\mathbb{P}\left\{\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p\in\mathcal{R}(3,C_2)}\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T\frac{u_{it}\rho_{it}\left(\left\{\Delta_{\Theta_j,it},x_{j,it}^*\right\}_{j=0}^p,\epsilon_{it}^*\right)}{\sum_{j=0}^p\left\|\Delta_{\Theta_j}\right\|_F}\right|>\frac{C_5a_{NT}}{4}\right\}$$

$$\leq\sum_{j=0}^p\sup_{x^*}\mathbb{P}\left\{\tau\mathcal{A}_j>\frac{C_5a_{NT}}{4\,(p+2)}\right\}+\sup_{x^*,\epsilon^*}\mathbb{P}\left\{\mathcal{A}_{p+1}>\frac{C_5a_{NT}}{4\,(p+2)}\right\}. \tag{B.24}$$

We first bound $\mathcal{A}_j$ for $j\in[p]$. We have

$$\mathcal{A}_j=\frac{1}{NT}\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p\in\mathcal{R}(3,C_2)}\left|\sum_{i=1}^N\sum_{t=1}^T\frac{u_{it}x_{j,it}^*\Delta_{\Theta_j,it}}{\sum_{j=0}^p\left\|\Delta_{\Theta_j}\right\|_F}\right|\leq\frac{1}{NT}\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p\in\mathcal{R}(3,C_2)}\frac{\left|tr\left[\Delta_{\Theta_j}'\left(U\odot x_j^*\right)\right]\right|}{\sum_{j=0}^p\left\|\Delta_{\Theta_j}\right\|_F}$$

$$\leq\frac{1}{NT}\left\|U\odot x_j^*\right\|_{op}\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p\in\mathcal{R}(3,C_2)}\frac{\left\|\Delta_{\Theta_j}\right\|_*}{\sum_{j=0}^p\left\|\Delta_{\Theta_j}\right\|_F}\leq\frac{c_8}{NT}\left\|U\odot x_j^*\right\|_{op}, \tag{B.25}$$

where the first inequality holds by $tr(AB)\leq\|A\|_{op}\|B\|_*$ and the second inequality holds by Lemma B.7. Then

$$\sup_{x^*,\epsilon^*}\mathbb{P}\left\{\tau\mathcal{A}_j>\frac{C_5a_{NT}}{4\,(p+2)}\right\}\leq\sup_{x^*}\mathbb{P}\left\{\frac{\left\|U\odot x_j^*\right\|}{\sqrt{(N\vee T)\log(N\vee T)}}>\frac{C_5}{4c_8\tau\,(p+2)}\right\}$$

$$\leq\sup_{x^*}\left\{\exp\left(-\frac{C_5}{4c_8\tau\,(p+2)}\right)\mathbb{E}\left[\exp\left(\frac{\left\|U\odot x_j^*\right\|}{\sqrt{(N\vee T)\log(N\vee T)}}\right)\right]\right\}$$

$$\leq C\exp\left(-\frac{C_5}{4c_8\tau\,(p+2)}\right) \tag{B.26}$$

for some absolute constant C that independent of (x^*, ϵ^*), where the last inequality holds by Lemma B.10. Similarly, we can establish

$$\sup_{x^*,\epsilon^*} \mathbb{P}\left\{\tau \mathcal{A}_0 > \frac{C_5 a_{NT}}{4\,(p+2)}\right\} \leq C \exp\left(-\frac{C_5}{4c_8 \tau\,(p+2)}\right). \tag{B.27}$$

Next, we turn to $\mathcal{A}_{p+1}$. We have

$$\sup_{x^*,\epsilon^*} \mathbb{P}\left\{\mathcal{A}_{p+1} > \frac{C_5 a_{NT}}{4\,(p+2)}\right\} \leq \exp\left(-\frac{C_5}{4c_8 (p+1)\,(p+2)}\right) \sup_{x^*,\epsilon^*} \mathbb{E}\left\{\exp\left(\frac{NT\mathcal{A}_{p+1}}{c_8(p+1)\sqrt{(N\vee T)\log\,(N \vee T)}}\right)\right\}. \tag{B.28}$$

Because $\phi_{it}(\cdot)$ is a contraction, Lemma B.9 implies

$$\begin{aligned}
&\mathbb{E}\left\{\exp\left(\frac{NT\mathcal{A}_{p+1}}{c_8(p+1)\sqrt{(N\vee T)\,(\log(N\vee T))}}\right)\right\} \\
&= \mathbb{E}\left\{\exp\left[\frac{1}{c_8(p+1)\sqrt{(N\vee T)\,(\log(N\vee T))}} \sup_{\{\Delta_{\Theta_j}\}_{j=0}^p \in \mathcal{R}(3,C_2)} \left|\frac{\sum_{i=1}^N \sum_{t=1}^T u_{it}\phi_{it}\left(\Delta_{\Theta_0,it} + \sum_{j=1}^p x^*_{j,it}\Delta_{\Theta_j,it}\right)}{\sum_{j=0}^p \left\|\Delta_{\Theta_j}\right\|_F}\right|\right]\right\} \\
&\leq \mathbb{E}\left\{\exp\left[\frac{1}{c_8(p+1)\sqrt{(N\vee T)\,(\log(N\vee T))}} \sup_{\{\Delta_{\Theta_j}\}_{j=0}^p \in \mathcal{R}(3,C_2)} \left|\sum_{i=1}^N \sum_{t=1}^T \frac{u_{it}(\Delta_{\Theta_0,it} + \sum_{j=1}^p x^*_{j,it}\Delta_{\Theta_j,it})}{\sum_{j=0}^p \left\|\Delta_{\Theta_j}\right\|_F}\right|\right]\right\} \\
&\leq \mathbb{E}\left\{\exp\left[\frac{\left(\|U\|_{op} + \sum_{j\in[p]} \left\|U \odot x^*_j\right\|_{op}\right)}{(p+1)\sqrt{(N\vee T)\,(\log(N\vee T))}}\right]\right\} \\
&\leq \mathbb{E}\left\{\exp\left[\frac{\left(\|U\|_{op}\right)}{\sqrt{(N\vee T)\,(\log(N\vee T))}}\right]\right\}^{1/(1+p)} \Pi_{j\in[p]}\left[\mathbb{E}\left\{\exp\left[\frac{\left(\left\|U\odot x^*_j\right\|_{op}\right)}{\sqrt{(N\vee T)\,(\log(N\vee T))}}\right]\right\}^{1/(1+p)}\right] \leq C
\end{aligned} \tag{B.29}$$

for some absolute constant C, where the first inequality is by Lemma B.9, the second inequality is by (B.25), the third inequality is due to the fact that, for random variables $\{A_i\}_{i\in[p+1]}$,

$$\mathbb{E}(\Pi_{i\in[p+1]}|A_i|) \leq \Pi_{i\in[p+1]}[\mathbb{E}|A_i^{1+p}|]^{1/(1+p)},$$

and the final inequality is by Lemma B.10 with an absolute constant C that is independent of (x^*, ϵ^*).

Combining (B.28) and (B.29), we have

$$\sup_{x^*,\epsilon^*} \mathbb{P}\left\{\mathcal{A}_{p+1} > \frac{C_5 a_{NT}}{4\,(p+2)}\right\} \leq C \exp\left(-\frac{C_5}{4c_8\,(p+1)\,(P+2)}\right),$$

which, combined with (B.22), (B.24), (B.26), and (B.27), further implies that

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p \in \mathcal{R}(3,C_2)} \left|\frac{1}{NT}\sum_{i=1}^N \sum_{t=1}^T \frac{\tilde{\rho}_{it}\left(\left\{\Delta_{\Theta_j,it}, X_{j,it}\right\}_{j=0}^p, \epsilon_{it}\right)}{\sum_{j=0}^p \left\|\Delta_{\Theta_j}\right\|_F}\right| > C_5 a_{NT}\right\} \\
&\leq 2\beta_{NT}^{-1} \sup_{x^*,\epsilon^*} \mathbb{P}\left\{\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p \in \mathcal{R}(3,C_2)} \left|\frac{1}{NT}\sum_{i=1}^N \sum_{t=1}^T \frac{u_{it}\rho_{it}\left(\left\{\Delta_{\Theta_j,it}, x^*_{j,it}\right\}_{j=0}^p, \epsilon^*_{it}\right)}{\sum_{j=0}^p \left\|\Delta_{\Theta_j}\right\|_F}\right| > \frac{C_5 a_{NT}}{4}\right\}
\end{aligned}$$

$$\leq C(1+o(1))\left[(p+1)\exp\left(-\frac{C_5}{4c_8\tau\,(p+2)}\right)+\exp\left(-\frac{C_5}{4c_8(p+1)\,(p+2)}\right)\right].$$

The RHS of the last inequality converges to zero as $C_5 \to \infty$, which implies that

$$\sup_{\{\Delta_{\Theta_j}\}_{j=0}^p \in \mathcal{R}(3,C_2)} \frac{\left|\frac{1}{NT}\sum_{i=1}^N\sum_{t=1}^T \tilde{\rho}_{it}\left(\{\Delta_{\Theta_j,it}, X_{j,it}\}_{j=0}^p, \epsilon_{it}\right)\right|}{\sum_{j=0}^p \left\|\Delta_{\Theta_j}\right\|_F} = O_p(a_{NT}).$$

■

B.2 Lemmas for the Proof of Theorem 3.2

Lemma B.12 *Let $\{\Upsilon_t, t=1,\cdots,T\}$ be a zero-mean strong mixing process, not necessarily stationary, with the mixing coefficients satisfying $\alpha(z)\leq c_\alpha\rho^z$ for some $c_\alpha>0$ and $\rho\in(0,1)$. If $\sup_{1\leq t\leq T}|\Upsilon_t|\leq M_T$, then there exist a constant c_9 depending on c_α and ρ such that for any $T\geq 2$ and $d>0$,*

(i) $\mathbb{P}\left\{\left|\sum_{t=1}^T\Upsilon_t\right|>d\right\}\leq\exp\left\{-\frac{c_9d^2}{M_T^2T+dM_T(\log T)(\log\log T)}\right\}$.

(ii) $\mathbb{P}\left\{\left|\sum_{t=1}^T\Upsilon_t\right|>d\right\}\leq\exp\left\{-\frac{c_9d^2}{v_0^2T+M_T^2+dM_T(\log T)^2}\right\}$,

with $v_0^2=\sup_{t\in[T]}\left[Var(\Upsilon_t)+2\sum_{s>t}|Cov(\Upsilon_t,\Upsilon_s)|\right]$.

Proof. The proof is the same as that of Theorems 1 and 2 in Merlevède et al. (2009)) with the condition assumed $\alpha(a)\leq\exp\{-2ca\}$ for some $c>0$ changed with $c_\alpha=1$ and $\rho=\exp\{-2c\}$ in our lemma instead.

■

Lemma B.13 *Suppose Assumptions 1-4 hold. Then, for $j\in\{0,\cdots,p\}$, we have*

(i) $\max_{i\in[N]}\left\|u_{i,j}^0\right\|_2\leq M$ *and* $\max_{t\in[T]}\left\|v_{t,j}^0\right\|_2\leq\frac{M}{\sigma_{K_j,j}}\leq\frac{M}{c_\sigma}$.

(ii) $\max_{t\in[T]}\left\|O_j'\tilde{v}_{t,j}\right\|_2\leq\frac{2M}{\sigma_{K_j,j}}\leq\frac{2M}{c_\sigma}$ *and* $\max_{t\in[T]}\left\|O_j^{(1)\prime}\tilde{v}_{t,j}^{(1)}\right\|_2\leq\frac{2M}{\sigma_{K_j,j}}\leq\frac{2M}{c_\sigma}$ *w.p.a.1.*

(iii) $\max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\phi}_{it}^{(1)}\right\|_2^2\leq\frac{4M^2}{c_\sigma}+\frac{4M^2pC}{c_\sigma}$ *w.p.a.1.*

Proof. (i) Recall that $\frac{1}{\sqrt{NT}}\Theta_j^0=\mathcal{U}_j^0\Sigma_j^0\mathcal{V}_j^{0\prime}$, $U_j^0=\sqrt{N}\mathcal{U}_j^0\Sigma_j^0$ and $V_j=\sqrt{T}\mathcal{V}_j$. Then we have

$$\frac{1}{\sqrt{T}}\Theta_j^0\mathcal{V}_j^0=\sqrt{N}\mathcal{U}_j^0\Sigma_j^0=U_j^0 \quad\text{and}\quad \frac{1}{\sqrt{N}}\mathcal{U}_j^{0\prime}\Theta_j^0=\sqrt{T}\Sigma_j^0\mathcal{V}_j^{0\prime}=\Sigma_j^0V_j^{0\prime}. \tag{B.30}$$

Hence, it's natural to see that

$$\left\|u_{i,j}^0\right\|_2=\frac{1}{\sqrt{T}}\left\|\left[\Theta_j^0\mathcal{V}_j^0\right]_{i.}\right\|_2\leq\frac{1}{\sqrt{T}}\left\|[\Theta_j]_{i.}\right\|_2\leq M,$$

where the first inequality is due to the fact that $\mathcal{V}_j$ is the unitary matrix and the last inequality holds by Assumption 2. Since the upper bound M is not dependent on i, this result holds uniformly in i. Analogously,

$$\left\|v_{t,j}^0\right\|_2\leq\frac{1}{\sqrt{N}}c_\sigma^{-1}\left\|\left[\mathcal{U}_j^{0\prime}\Theta_j^0\right]_{.t}\right\|_2\leq\frac{1}{\sqrt{N}}c_\sigma^{-1}\left\|\left[\Theta_j^0\right]_{.t}\right\|_2\leq\frac{M}{c_\sigma}.$$

(ii) As in (B.30), we have

$$\frac{1}{\sqrt{N}}\tilde{\mathcal{U}}_j'\tilde{\Theta}_j = \sqrt{T}\tilde{\Sigma}_j\tilde{\mathcal{V}}_j' = \tilde{\Sigma}_j\tilde{V}_j'.$$

It follows that

$$\left\|O_j'\tilde{v}_{t,j}\right\|_2 \le \frac{1}{\sqrt{N}}\tilde{\sigma}_{K_j,j}^{-1}\left\|\left[\tilde{\mathcal{U}}_j'\tilde{\Theta}_j\right]_{.t}\right\|_2 \le \frac{1}{\sqrt{N}}\tilde{\sigma}_{K_j,j}^{-1}\left\|\left[\tilde{\Theta}_j\right]_{.t}\right\|_2 \le \frac{2M}{c_\sigma},$$

where the last inequality holds due to the fact that $\max_{k\in[K_j]}\left|\tilde{\sigma}_{k,j}^{-1}-\Sigma_{k,j}^{-1}\right| \le \Sigma_{K_j,j}^{-1}$ w.p.a.1. and the bounded parameter space where $\tilde{\Theta}_j$ lies in by Assumption 2 and ADMM algorithm proposed in the last section. The upper bound of $\max_{t\in[T]}\left\|O_j^{(1)\prime}\tilde{v}_{t,j}^{(1)}\right\|_2$ follows the same argument as above.

(iii) We observe that

$$\max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2^2 \le \frac{1}{T}\sum_{t\in[T]}\left\|O_0'\tilde{v}_{t,0}^{(1)}\right\|_2^2 + \max_{i\in I_2}\sum_{j\in[p]}\frac{1}{T}\sum_{t\in[T]}\left\|O_j'\tilde{v}_{t,j}^{(1)}\right\|_2^2|X_{j,it}|^2 \le \frac{4M^2}{c_\sigma^2}+\frac{4M^2pC}{c_\sigma^2} \quad \text{w.p.a.1},$$

where the last inequality holds by Lemma B.13(ii) and Assumption 1(iv). ■

Lemma B.14 *Under Assumptions 1–5, we have*

(i) $\min_{i\in I_2}\lambda_{\min}\left(\tilde{\Phi}_i^{(1)}\right) \ge \frac{c_\phi}{2}$, $\max_{i\in I_2}\lambda_{\max}\left(\tilde{\Phi}_i^{(1)}\right) \le 2C_\phi$ *w.p.a.1,*

(ii) For $\forall j\in[p]$, $\max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left[X_{j,it}^2-\mathbb{E}\left(X_{j,it}^2\middle|\mathscr{D}_{\{e_{is}\}_{s<t}}^{I_1}\right)\right]\left\|\tilde{O}_j^{(1)\prime}\tilde{v}_{t,j}^{(1)}-v_{t,j}^0\right\|_2^2 = O_p(\eta_N^2)$,

(iii) $\max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}-\Phi_{it}^0\right\|_2^2 = O_p(\eta_N^2)$.

Proof. (i) Recall that

$$\Phi_i = \frac{1}{T}\sum_{t=1}^T\Phi_{it}^0\Phi_{it}^{0\prime} \quad \text{with} \quad \Phi_{it}^0 = (v_{t,0}^{0\prime}, v_{t,1}^{0\prime}X_{1,it}, ..., v_{t,p}^{0\prime}X_{p,it})', \quad \text{and}$$

$$\tilde{\Phi}_i^{(1)} = \frac{1}{T}\sum_{t=1}^T\tilde{\Phi}_{it}^{(1)}\tilde{\Phi}_{it}^{(1)\prime} \quad \text{with} \quad \tilde{\Phi}_{it}^{(1)} = \left[\left(O_0^{(1)\prime}\tilde{v}_{t,0}^{(1)}\right)', \left(O_1^{(1)\prime}\tilde{v}_{t,1}^{(1)}X_{1,it}\right)', \cdots, \left(O_p^{(1)\prime}\tilde{v}_{t,p}^{(1)}X_{p,it}\right)'\right]'.$$

Uniformly over $i\in I_2$, it is clear that

$$\begin{aligned}
&\left\|\tilde{\Phi}_i^{(1)}-\Phi_i\right\|_F \\
&\lesssim \frac{4M}{c_\sigma T}\sum_{t=1}^T\left\|O_0^{(1)\prime}\tilde{v}_{t,0}^{(1)}-v_{t,0}^0\right\|_2 + \frac{4M}{c_\sigma T}\sum_{j=1}^p\sum_{t=1}^T\left\|O_j^{(1)\prime}\tilde{v}_{t,j}^{(1)}-v_{t,j}^0\right\|_2|X_{j,it}| \\
&\le \frac{4M}{c_\sigma}\frac{1}{\sqrt{T}}\left\|O_0^{(1)\prime}\tilde{V}_0^{(1)}-V_0^0\right\|_F + \frac{4M^2}{c_\sigma}\sum_{j=1}^p\frac{1}{\sqrt{T}}\left\|O_j^{(1)\prime}\tilde{V}_j^{(1)}-V_j^0\right\|_F\left(\frac{1}{T}\sum_{t\in[T]}|X_{j,it}|^2\right)^{1/2} = O_p(\eta_N),
\end{aligned}$$

where the third line holds by Lemma B.13(i) and Assumption 1(iv). It follows that

$$\min_{i\in I_2}\lambda_{\min}\left[\tilde{\Phi}_i^{(1)}\right] \ge \min_{i\in I_2}\lambda_{\min}[\Phi_i] - O(\eta_N) \ge \frac{c_\phi}{2}, \quad \text{w.p.a.1}$$

and

$$\max_{i\in I_2}\lambda_{\max}\left[\tilde{\Phi}_i^{(1)}\right] \le \max_{i\in I_2}\lambda_{\max}[\Phi_i] + O(\eta_N) \le 2C_\phi, \quad \text{w.p.a.1.}$$

(ii) Let $I_{j,i} := \frac{1}{T}\sum_{t\in[T]} I_{j,it}$ such that $I_{j,it} = \left[X_{j,it}^2 - \mathbb{E}\left(X_{j,it}^2 \middle| \mathscr{D}^{I_1}_{\{e_{is}\}_{s<t}}\right)\right] \left\|\tilde{O}_j^{(1)\prime}\tilde{v}_{t,j}^{(1)} - v_{t,j}^0\right\|_2^2$. Then for a constant c, we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_2}\left|\sum_{t\in[T]} I_{j,it}\right| > cT\eta_N^2\right) \leq \sum_{i\in I_2}\mathbb{P}\left(\left|\sum_{t\in[T]} I_{j,it}\right| > cT\eta_N^2\right)\\
&= \sum_{i\in I_2}\mathbb{E}\mathbb{P}\left(\left|\sum_{t\in[T]} I_{j,it}\right| > cT\eta_N^2 \middle| \mathscr{D}^{I_1}_{\{e_{is}\}_{s<T}}\right)\\
&\leq 2\sum_{i\in I_2}\exp\left\{-\frac{2(cT\eta_N^2)^2}{\sum_{t\in[T]}\left[2\xi_N^2\left\|\tilde{O}_j^{(1)\prime}\tilde{v}_{t,j}^{(1)} - v_{t,j}^0\right\|_2^2\right]^2}\right\}\\
&\leq 2\exp\left\{-\frac{2(cT\eta_N^2)^2}{4\xi_N^4\left[\frac{M^2}{c_\sigma^2}+\frac{4M^2}{c_\sigma^2}\right]\sum_{t\in[T]}\left\|\tilde{O}_j^{(1)\prime}\tilde{v}_{t,j}^{(1)} - v_{t,j}^0\right\|_2^2} + \log N\right\} = o(1),
\end{aligned}$$

where the first inequality combines the fact that $I_{j,it}$ is the martingale difference sequence, Assumption 1(v) and the Azuma-Hoeffding inequality in Wainwright (2019, Corollary 2.20). The last inequality is by Lemma B.13(i) and B.13(ii), and the finial result is by the definition of η_N.

(iii) Note that

$$\begin{aligned}
&\max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^0\right\|_2^2 \leq \frac{1}{T}\sum_{t\in[T]}\left\|\tilde{O}_0^{(1)\prime}\tilde{v}_{t,0}^{(1)} - v_{t,0}^0\right\|_2^2 + p\max_{i\in I_2, j\in[p]}\frac{1}{T}\sum_{t\in[T]}|X_{j,it}|^2\left\|\tilde{O}_j^{(1)\prime}\tilde{v}_{t,j}^{(1)} - v_{t,j}^0\right\|_2^2\\
&= \frac{1}{T}\sum_{t\in[T]}\left\|\tilde{O}_0^{(1)\prime}\tilde{v}_{t,0}^{(1)} - v_{t,0}^0\right\|_2^2 + p\max_{i\in I_2, j\in[p]}\frac{1}{T}\sum_{t\in[T]}\left[X_{j,it}^2 - \mathbb{E}\left(X_{j,it}^2\middle|\mathscr{D}^{I_1}_{\{e_{is}\}_{s<t}}\right)\right]\left\|\tilde{O}_j^{(1)\prime}\tilde{v}_{t,j}^{(1)} - v_{t,j}^0\right\|_2^2\\
&+ \max_{i\in I_2, j\in[p]}\frac{1}{T}\sum_{t\in[T]}\mathbb{E}\left(X_{j,it}^2\middle|\mathscr{D}^{I_1}_{\{e_{is}\}_{s<t}}\right)\left\|\tilde{O}_j^{(1)\prime}\tilde{v}_{t,j}^{(1)} - v_{t,j}^0\right\|_2^2\\
&\leq \frac{1}{T}\sum_{t\in[T]}\left\|\tilde{O}_0^{(1)\prime}\tilde{v}_{t,0}^{(1)} - v_{t,0}^0\right\|_2^2 + O_p(\eta_N^2) + pM\max_{j\in[p]}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{O}_j^{(1)\prime}\tilde{v}_{t,j}^{(1)} - v_{t,j}^0\right\|_2^2\\
&= \frac{1}{T}\left\|O_0^{(1)}\tilde{V}_0^{(1)} - V_0^0\right\|_F^2 + p\max_{j\in[p]}\frac{1}{T}\left\|O_j^{(1)}\tilde{V}_j^{(1)} - V_j^0\right\|_F^2 + O_p(\eta_N^2)\\
&= O_p(\eta_N^2),
\end{aligned}$$

where the the second inequality holds by Assumption 1(iv) and Lemma B.14(ii), and the last equality holds by the Theorem 3.1(ii). ■

Lemma B.15 *Recall $\{A_{1,i},\cdots,A_{7,i}\}_{i\in I_2}$ and q_i^I defined in (A.5) and (A.6), respectively. Suppose Assumptions 1–5 hold. Then for any constant $c_{11} < \min(\frac{3\underline{\mathfrak{f}}}{\bar{\mathfrak{f}}'}, 1)$, we have*

$$\max_{i\in I_2}(|A_{m,i}|/\left\|\dot{\Delta}_{i,u}\right\|_2) = O_p(\eta_N), \forall m\in\{1,2,3,5,6,7\} \quad \textit{and}$$

$$|A_{4,i}| \geq \min\left(\frac{\left(3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'\right)c_\phi\left\|\dot{\Delta}_{i,u}\right\|_2^2}{12}, \frac{\left(3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'\right)\sqrt{c_\phi}q_i^I\left\|\dot{\Delta}_{i,u}\right\|_2}{6\sqrt{2}}\right), \quad \forall i\in I_2, \quad w.p.a.1.$$

Proof. Recall that $w_{1,it} = \epsilon_{it} - u_i^{0\prime}\left(\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^0\right)$. Let $w_{2,it} = \tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}$. For some positive constant $c_{11} \in (0,1]$, we observe that

$$\begin{aligned}
A_{4,i} &= \frac{1}{T}\sum_{t=1}^{T}\int_0^{w_{2,it}} \mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it} \leq s\right\} - \mathbf{1}\left\{\epsilon_{it} \leq 0\right\}\middle| \mathscr{D}_{e_i}^{I_1}\right) ds \\
&= \frac{1}{T}\sum_{t=1}^{T}\int_0^{w_{2,it}} \left[\mathfrak{F}_{it}(s) - \mathfrak{F}_{it}(0)\right] ds \geq \frac{1}{T}\sum_{t=1}^{T}\int_0^{c_{11}w_{2,it}} \left[\mathfrak{F}_{it}(s) - \mathfrak{F}_{it}(0)\right] ds \\
&= \frac{1}{T}\sum_{t=1}^{T}\int_0^{c_{11}w_{2,it}} \left[s\mathfrak{f}_{it}(0) + \frac{s^2}{2}\mathfrak{f}_{it}'(\tilde{s})\right] ds \\
&\geq \frac{1}{T}\sum_{t=1}^{T}\left[\frac{c_{11}^2\underline{\mathfrak{f}}\left(\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}\right)^2}{2} - \frac{c_{11}^3\bar{\mathfrak{f}}'\left|\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}\right|^3}{6}\right]
\end{aligned} \tag{B.31}$$

where $\tilde{s} \in (0,s)$. Here, due to Assumption 1(i), the conditional CDF of ϵ_{it} given $\mathscr{D}_{e_i}^{I_1}$ is the same as that given the σ-field generated by $\left\{\{e_{it}\}_{t\in[T]} \bigcup \left\{V_j^0\right\}_{j\in[p]\cup\{0\}} \bigcup \left\{W_j^0\right\}_{j\in[p]}\right\}$, which leads to the second equality of the above display by Assumption 1(vii); the first inequality holds by Lemma B.5(ii) for any $c_{11} \in (0,1]$. Here, we choose c_{11} such that $c_{11} < \frac{3\underline{\mathfrak{f}}}{\bar{\mathfrak{f}}'}$; and the last inequality holds by Assumption 1(vii).

Let $q_i^{II} = \left[\frac{1}{T}\sum_{t\in[T]}\left(\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}\right)^2\right]^{\frac{1}{2}}$ and recall that $q_i^I = \inf_{\Delta} \frac{\left[\frac{1}{T}\sum_{t\in[T]}\left(\tilde{\Phi}_{it}^{(1)\prime}\Delta\right)^2\right]^{\frac{3}{2}}}{\frac{1}{T}\sum_{t\in[T]}\left|\tilde{\Phi}_{it}^{(1)\prime}\Delta\right|^3}$. If $q_i^I \geq q_i^{II}$, we notice that $\frac{1}{T}\sum_{t\in[T]}\left|\tilde{\Phi}_{it}^{(1)\prime}\Delta\right|^3 < \left(q_i^{II}\right)^2$ and $A_{4,i} \geq \frac{c_{11}^2\underline{\mathfrak{f}}\left(q_i^{II}\right)^2}{2} - \frac{c_{11}^3\bar{\mathfrak{f}}'\left(q_i^{II}\right)^2}{6} = \frac{3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'}{6}\left(q_i^{II}\right)^2$. If $q_i^I < q_i^{II}$, we have $\left[\frac{1}{T}\sum_{t\in[T]}\left(\tilde{\Phi}_{it}^{(1)\prime}\Delta_{i,u}^*\right)^2\right]^{\frac{1}{2}} = q_i^I$ with $\Delta_{i,u}^* = \frac{q_i^{II}\dot{\Delta}_{i,u}}{q_i^I}$. Define the function

$$F(\Delta) = \frac{1}{T}\sum_{t=1}^{T}\int_0^{\tilde{\Phi}_{it}^{(1)\prime}\Delta} \left[\mathfrak{F}_{it}(s) - \mathfrak{F}_{it}(0)\right] ds.$$

Note that the second-order derivative of function $F(\Delta)$ is no less than zero, which implies $F(\Delta)$ is convex. Therefore, we have

$$F\left(\dot{\Delta}_{i,u}\right) = F\left(\frac{q_i^I\Delta_{i,u}^*}{q_i^{II}}\right) \geq \frac{q_i^I}{q_i^{II}}F\left(\Delta_{i,u}^*\right) \geq \frac{q_i^{II}}{q_i^I}\frac{3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'}{6}\frac{1}{T}\sum_{t\in[T]}\left(\tilde{\Phi}_{it}^{(1)\prime}\Delta_{i,u}^*\right)^2 = \frac{\left(3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'\right)q_i^Iq_i^{II}}{6}.$$

Combining these two cases, we have

$$\begin{aligned}
A_{4,i} &\geq \min\left(\frac{3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'}{6}\left(q_i^{II}\right)^2, \frac{\left(3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'\right)q_i^Iq_i^{II}}{6}\right) \\
&\geq \min\left(\frac{\left(3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'\right)c_\phi\left\|\dot{\Delta}_{i,u}\right\|_2^2}{12}, \frac{\left(3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'\right)\sqrt{c_\phi}q_i^I\left\|\dot{\Delta}_{i,u}\right\|_2}{6\sqrt{2}}\right),
\end{aligned} \tag{B.32}$$

where the second inequality holds by Lemma B.14(i).

As for $|A_{1,i}|$, we notice that

$$\max_{i\in I_2}\left(|A_{1,i}|/\left\|\dot{\Delta}_{i,u}\right\|_2\right) = \max_{i\in I_2}\frac{\left|\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left\{\tilde{\Phi}_{it}^{(1)\prime}\left(\tau - \mathbf{1}\left\{\epsilon_{it} \leq u_i^{0\prime}\left(\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^0\right)\right\}\right)\middle| \mathscr{D}_{e_i}^{I_1}\right\}\dot{\Delta}_{i,u}\right|}{\left\|\dot{\Delta}_{i,u}\right\|_2}$$

$$
\begin{aligned}
&= \max_{i\in I_2} \frac{\left|\frac{1}{T}\sum_{t=1}^{T} \tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}\left(\tau - \mathfrak{F}_{it}\left[u_i^{0\prime}\left(\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^{0}\right)\right]\right)\right|}{\left\|\dot{\Delta}_{i,u}\right\|_2} \\
&= \max_{i\in I_2} \frac{\left|\frac{1}{T}\sum_{t=1}^{T} \tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}\left[\mathfrak{f}_{it}(s_{it})u_i^{0\prime}\left(\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^{0}\right)\right]\right|}{\left\|\dot{\Delta}_{i,u}\right\|_2} \\
&\le \max_{i\in I_2} \frac{\bar{\mathfrak{f}}}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2 \left\|\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^{0}\right\|_2 \left\|u_i^0\right\|_2 \\
&\le \max_{i\in I_2}\sqrt{\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|^2}\sqrt{\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^{0}\right\|^2}\left\|u_i^0\right\|_2 \\
&\le O_p(\eta_N),
\end{aligned} \tag{B.33}
$$

where the second and third equalities hold by Assumption 1(vii) and mean-value theorem with some $|s_{it}| \in \left(0, \left|u_i^{0\prime}(\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^{0})\right|\right)$, the second inequality holds by Cauchy-Schwarz inequality, and the third inequality holds by Lemmas B.13(i), B.13(iii) and B.14(ii).

For $A_{3,i}$, note that

$$
\begin{aligned}
A_{3,i} &= \frac{1}{T}\sum_{t\in[T]}\int_0^{\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}} \left(\mathbf{1}\left\{\epsilon_{it}\le s\right\} - \mathbf{1}\left\{\epsilon_{it}\le 0\right\}\right) - \left\{\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}\le s\right\} - \mathbf{1}\left\{\epsilon_{it}\le 0\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right\} ds \\
&= \int_0^1 \left\{\frac{1}{T}\sum_{t\in[T]}\tilde{\Phi}_{it}^{(1)}\left(\mathbf{1}\left\{\epsilon_{it}\le \tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}s^*\right\} - \mathbf{1}\left\{\epsilon_{it}\le 0\right\}\right)\right. \\
&\quad \left.- \left[\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}\le \tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}s^*\right\} - \mathbf{1}\left\{\epsilon_{it}\le 0\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right]\right\}^{\prime}\dot{\Delta}_{i,u}ds^* \\
&\le \sup_{s\in\mathbb{R}}\left\|A_{3,i}^I(s)\right\|_2\left\|\dot{\Delta}_{i,u}\right\|_2
\end{aligned} \tag{B.34}
$$

by change of variables with $s^* = \frac{s}{\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}}$, $A_{3,i}^I(s) = \frac{1}{T}\sum_{t\in[T]} A_{3,it}^I(s)$ and

$$
A_{3,it}^I(s) = \tilde{\Phi}_{it}^{(1)}\left[\left(\mathbf{1}\left\{\epsilon_{it}\le s\right\} - \mathbf{1}\left\{\epsilon_{it}\le 0\right\}\right) - \mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}\le s\right\} - \mathbf{1}\left\{\epsilon_{it}\le 0\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right].
$$

Below, we aim to show $\sup_{s\in(-\infty,+\infty)}\max_{i\in I_2}\left\|A_{3,i}^I(s)\right\|_2 = O_p(\eta_N)$.

When $|s| > T^{1/4}$, we notice that

$$
\begin{aligned}
\sup_{|s|>T^{1/4}}\max_{i\in I_2}\left\|A_{3,i}^I(s)\right\|_2 &\le \sup_{|s|>T^{1/4}}\max_{i\in I_2}\left\|\frac{1}{T}\sum_{t\in[T]}\tilde{\Phi}_{it}^{(1)}\left[\mathbf{1}\left\{\epsilon_{it}\le s\right\} - \mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}\le s\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right]\right\|_2 \\
&\quad + \max_{i\in I_2}\left\|\frac{1}{T}\sum_{t\in[T]}\tilde{\Phi}_{it}^{(1)}\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\} - \mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}\le 0\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right]\right\|_2 \\
&\le \max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\mathbf{1}\left\{\epsilon_{it} > T^{1/4}\right\} + \max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it} > T^{1/4}\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right) \\
&\quad + \max_{i\in I_2}\left\|\frac{1}{T}\sum_{t\in[T]}\tilde{\Phi}_{it}^{(1)}\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\} - \mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}\le 0\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right]\right\|_2
\end{aligned}
$$

$$
\begin{aligned}
&\leq 2\max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}>T^{1/4}\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\\
&+\max_{i\in I_2}\left|\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\left[\mathbf{1}\left\{\epsilon_{it}>T^{1/4}\right\}-\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}>T^{1/4}\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right]\right|\\
&+\max_{i\in I_2}\left|\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right]\right|,
\end{aligned}
$$

where the first and third inequalities hold by triangle inequality. Besides, we observe that

$$
\begin{aligned}
&\max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}>T^{1/4}\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\\
&=\max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\mathbb{P}\left(\epsilon_{it}>T^{1/4}\middle|\mathscr{D}_{e_i}^{I_1}\right)\leq\max_{i\in I_2}\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\frac{\mathbb{E}\left(\epsilon_{it}^2\middle|\mathscr{D}_{e_i}^{I_1}\right)}{\sqrt{T}}\\
&\leq T^{-1/2}\max_{i\in I_2}\sqrt{\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2^2}\max_{i\in I_2}\sqrt{\frac{1}{T}\sum_{t\in[T]}\left[\mathbb{E}\left(\epsilon_{it}^2\middle|\mathscr{D}_{e_i}\right)\right]^2}=O_p\left(T^{-1/2}\right),
\end{aligned}
\tag{B.35}
$$

where the last equality holds by Lemma B.13(iii) and Assumption 1(iv). For a positive constant c_{12},

$$
\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_2}\left|\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\left[\mathbf{1}\left\{\epsilon_{it}>T^{1/4}\right\}-\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}>T^{1/4}\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right]\right|>c_{12}\frac{\xi_N\sqrt{\log(N\vee T)}}{\sqrt{T}}\right)\\
&\leq\sum_{i\in I_2}\mathbb{E}\mathbb{P}\left(\left|\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\left[\mathbf{1}\left\{\epsilon_{it}>T^{1/4}\right\}-\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}>T^{1/4}\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right]\right|>c_{12}\frac{\xi_N\sqrt{\log(N\vee T)}}{\sqrt{T}}\middle|\mathscr{D}_{e_i}^{I_1}\right)\\
&\leq\sum_{i\in I_2}\exp\left\{-\frac{c_9c_{12}^2T\xi_N^2\log(N\vee T)}{\frac{16TM^2}{c_\sigma^2}\left(1+p\xi_N^2\right)+\frac{4Mc_{12}}{c_\sigma}\sqrt{1+p\xi_N}\sqrt{T}\xi_N\sqrt{\log(N\vee T)}\log T\left(\log\log T\right)}\right\}=o(1),
\end{aligned}
$$

where the first inequality holds by the union bound and Assumption 1(i), and the second inequality holds by Assumption 1(iii) and the conditional Bernstein's inequality in Lemma B.12(i) with the fact that $\max_{i\in I_2,t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\leq\frac{2M}{c_\sigma}\sqrt{1+p\xi_N^2}$ w.p.a.1. Here, we can apply the conditional Bernstein's inequality because $\tilde{\Phi}_{it}^{(1)}$ and $\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}>T^{1/4}\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)$ are deterministic given $\left\{\mathscr{D}_{e_i}^{I_1}\right\}_{i\in I_2,t\in[T]}$ so that the only randomness comes from $\{\epsilon_{it}\}_{t\in[T]}$. Furthermore, the joint distribution of $\{\epsilon_{it}\}_{t\in[T]}$ given $\mathscr{D}_{e_i}^{I_1}$ is the same as that given the σ-field generated by $\mathscr{D}_{e_i}$ due to the independence structure assumed in Assumption 1(i). Last, given the σ-field generated by $\mathscr{D}_{e_i}$, $\{\epsilon_{it}\}_{t\in[T]}$ is strong mixing with mixing coefficient $\alpha_i(\cdot)$ as assumed by Assumption 1(iii). Similarly, we obtain that

$$
\max_{i\in I_2}\left|\frac{1}{T}\sum_{t\in[T]}\left\|\tilde{\Phi}_{it}^{(1)}\right\|_2\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\middle|\mathscr{D}_{e_i}^{I_1}\right)\right]\right|=O_p\left(\frac{\sqrt{\log(N\vee T)}\xi_N}{\sqrt{T}}\right),
$$

which implies

$$
\sup_{|s|>T^{1/4}}\max_{i\in I_2}\left|A_{3,i}^I(s)\right|=O_p(\eta_N).
\tag{B.36}
$$

For $|s| \le T^{1/4}$, let $\mathbb{S} = [-T^{1/4}, T^{1/4}]$ and divide $\mathbb{S}$ into $\mathbb{S}_m$ for $m = 1, \cdots, n_{\mathbb{S}}$ such that $|s - \bar{s}| < \frac{\varepsilon}{T}$ for s and $\bar{s} \in \mathbb{S}_m$ and $n_{\mathbb{S}} \asymp T^{5/4}$. Let $s_m \in \mathbb{S}_m$. For any $s \in \mathbb{S}_m$, we have

$$\left\| \frac{1}{T} \sum_{t\in[T]} A^I_{3,it}(s) \right\|_2 \le \left\| \frac{1}{T} \sum_{t\in[T]} A^I_{3,it}(s_m) \right\|_2 + \left\| \frac{1}{T} \sum_{t\in[T]} \left[A^I_{3,it}(s) - A^I_{3,it}(s_m) \right] \right\|_2, \tag{B.37}$$

such that

$$\begin{aligned}
&\max_{i\in I_2, m\in[n_{\mathbb{S}}]} \sup_{s\in\mathbb{S}_m} \left\| \frac{1}{T} \sum_{t\in[T]} \left[A^I_{3,it}(s) - A^I_{3,it}(s_m) \right] \right\|_2 \\
&= \max_{i\in I_2, m\in[n_{\mathbb{S}}]} \sup_{s\in\mathbb{S}_m} \left\| \frac{1}{T} \sum_{t\in[T]} \tilde{\Phi}^{(1)}_{it} \left(\mathbf{1}\{\epsilon_{it} \le s\} - \mathbf{1}\{\epsilon_{it} \le s_m\} \right) - \tilde{\Phi}^{(1)}_{it} \left[\mathbb{E}\left(\mathbf{1}\{\epsilon_{it} \le s\} - \mathbf{1}\{\epsilon_{it} \le s_m\} \middle| \mathscr{D}^{I_1}_{e_i} \right) \right] \right\|_2 \\
&\le \max_{i\in I_2, m\in[n_{\mathbb{S}}]} \frac{2}{T} \sum_{t\in[T]} \left\| \tilde{\Phi}^{(1)}_{it} \right\|_2 \mathbb{E}\left(\mathbf{1}\left\{ \epsilon_{it} \in \left[s_m - \frac{\varepsilon}{T}, s_m + \frac{\varepsilon}{T} \right] \right\} \middle| \mathscr{D}^{I_1}_{e_i} \right) \\
&+ \max_{i\in I_2, m\in[n_{\mathbb{S}}]} \left| \frac{1}{T} \sum_{t\in[T]} \left\| \tilde{\Phi}^{(1)}_{it} \right\|_2 \left[\left(\mathbf{1}\left\{ \epsilon_{it} \in \left[s_m - \frac{\varepsilon}{T}, s_m + \frac{\varepsilon}{T} \right] \right\} \right) - \mathbb{E}\left(\mathbf{1}\left\{ \epsilon_{it} \in \left[s_m - \frac{\varepsilon}{T}, s_m + \frac{\varepsilon}{T} \right] \right\} \middle| \mathscr{D}^{I_1}_{e_i} \right) \right] \right| \\
&:= \max_{i\in I_2, m\in[n_{\mathbb{S}}]} \frac{2}{T} \sum_{t\in[T]} \left\| \tilde{\Phi}^{(1)}_{it} \right\|_2 \mathbb{E}\left(\mathbf{1}\left\{ \epsilon_{it} \in \left[s_m - \frac{\varepsilon}{T}, s_m + \frac{\varepsilon}{T} \right] \right\} \middle| \mathscr{D}^{I_1}_{e_i} \right) + \max_{i\in I_2, m\in[n_{\mathbb{S}}]} \left| A^{II}_{3,i}(m) \right|.
\end{aligned} \tag{B.38}$$

For the first term in (B.38), we notice that

$$\begin{aligned}
&\max_{i\in I_2, m\in[n_{\mathbb{S}}]} \frac{2}{T} \sum_{t\in[T]} \left\| \tilde{\Phi}^{(1)}_{it} \right\|_2 \mathbb{E}\left(\mathbf{1}\left\{ \epsilon_{it} \in \left[s_m - \frac{\varepsilon}{T}, s_m + \frac{\varepsilon}{T} \right] \right\} \middle| \mathscr{D}^{I_1}_{e_i} \right) \\
&= \max_{i\in I_2, m\in[n_{\mathbb{S}}]} \frac{2}{T} \sum_{t\in[T]} \left\| \tilde{\Phi}^{(1)}_{it} \right\|_2 \left(\mathfrak{F}_{it}(s_m + \frac{\varepsilon}{T}) - \mathfrak{F}_{it}(s_m - \frac{\varepsilon}{T}) \right) \\
&\le \max_{i\in I_2} \frac{2}{T} \sum_{t\in[T]} \left\| \tilde{\Phi}^{(1)}_{it} \right\|_2 \frac{2\varepsilon}{T} \\
&\le \frac{6M\varepsilon}{c_\sigma T} \left(1 + \sqrt{p} C \right) \quad \text{w.p.a.1},
\end{aligned} \tag{B.39}$$

where the first inequality is by mean-value theorem, and the second inequality is due to the fact that

$$\begin{aligned}
\max_{i\in I_2} \frac{1}{T} \sum_{t\in[T]} \left\| \tilde{\Phi}^{(1)}_{it} \right\|_2 &\le \max_{i\in I_2} \frac{1}{T} \sum_{t\in[T]} \sqrt{ \frac{4M^2}{c_\sigma} \left(1 + p \max_{j\in[p]} |X_{j,it}|^2 \right) } \\
&\le \max_{i\in I_2} \frac{2M}{c_\sigma} \frac{1}{T} \sum_{t\in[T]} \left(1 + \sqrt{p} \max_{j\in[p]} |X_{j,it}| \right) \\
&\le \frac{2M\left(1 + \sqrt{p} C \right)}{c_\sigma} \quad \text{w.p.a.1}.
\end{aligned}$$

For $A^{II}_{3,i}(m)$, let $A^{II}_{3,i}(m) = \frac{1}{T} \sum_{t\in[T]} A^{II}_{3,it}(m)$ with

$$A^{II}_{3,it}(m) = \left\| \tilde{\Phi}^{(1)}_{it} \right\|_2 \left[\left(\mathbf{1}\left\{ \epsilon_{it} \in \left[s_m - \frac{\varepsilon}{T}, s_m + \frac{\varepsilon}{T} \right] \right\} \right) - \mathbb{E}\left(\mathbf{1}\left\{ \epsilon_{it} \in \left[s_m - \frac{\varepsilon}{T}, s_m + \frac{\varepsilon}{T} \right] \right\} \middle| \mathscr{D}^{I_1}_{e_i} \right) \right].$$

We first observe that

$$\max_{i\in I_2,m\in[n_{\mathbb{S}}],t\in[T]} Var\left(A^{II}_{3,it}(m)\middle|\mathscr{D}^{I_1}_{e_i}\right) \leq \max_{i\in I_2,t\in[T]}\left\|\tilde{\Phi}^{(1)}_{it}\right\|_2^2 \mathbb{E}\left(\mathbf{1}\left\{\epsilon_{it}\in\left[s_m-\frac{\varepsilon}{T}, s_m+\frac{\varepsilon}{T}\right]\right\}\middle|\mathscr{D}^{I_1}_{e_i}\right)$$

$$\leq \max_{i\in I_2,t\in[T]}\left\|\tilde{\Phi}^{(1)}_{it}\right\|_2^2 \frac{2\varepsilon}{T} \lesssim \frac{\xi_N^2\varepsilon}{T}, \quad w.p.a.1,$$

where the first inequality is by $Var(x)\leq\mathbb{E}(x^2)$ for any random variable x and the second inequality is by the mean-value theorem and Assumption 1(v). Similarly, we have

$$\begin{aligned}
&\max_{i\in I_2,m\in[n_{\mathbb{S}}],t\in[T]}\sum_{s=t+1}^{T}\left|Cov\left(A^{II}_{3,it}(m), A^{II}_{3,is}(m)\middle|\mathscr{D}^{I_1}_{e_i}\right)\right| \\
&\leq \max_{i\in I_2,m\in[n_{\mathbb{S}}],t\in[T]}\sum_{s=t+1}^{T}\mathbb{E}\left(\left|A^{II}_{3,it}(m)\right|^3\middle|\mathscr{D}^{I_1}_{e_i}\right)^{1/3}\mathbb{E}\left(\left|A^{II}_{3,is}(m)\right|^3\middle|\mathscr{D}^{I_1}_{e_i}\right)^{1/3}[\alpha(t-s)]^{1/3} \\
&\lesssim \max_{i\in I_2,m\in[n_{\mathbb{S}}],t\in[T]}\mathbb{E}\left(\left|A^{II}_{3,it}(m)\right|^3\middle|\mathscr{D}^{I_1}_{e_i}\right)^{2/3}\lesssim \xi_N^2\left(\frac{\varepsilon}{T}\right)^{2/3}, \quad w.p.a.1.
\end{aligned}$$

It follows that

$$\max_{i\in I_2,m\in[n_{\mathbb{S}}],t\in[T]}\left\{Var\left(A^{II}_{3,it}(m)\middle|\mathscr{D}^{I_1}_{e_i}\right)+2\sum_{s=t+1}^{T}\left|Cov\left(A^{II}_{3,it}(m),A^{II}_{3,is}(m)\middle|\mathscr{D}^{I_1}_{e_i}\right)\right|\right\}\leq c_{13}\xi_N^2\left(\frac{\varepsilon}{T}\right)^{2/3}, \quad w.p.a.1.$$

and $\max_{i\in I_2,m\in[n_{\mathbb{S}}],t\in[T]}\left|A^{II}_{3,it}(m)\right|\leq C_{13}\xi_N$ w.p.a.1 for positive constants c_{13} and c_{14}. Denote the events

$$\mathscr{A}_{3,N}=\left(\begin{array}{c}\max_{i\in I_2,m\in[n_{\mathbb{S}}],t\in[T]}\left\{Var\left(A^{II}_{3,it}(m)\middle|\mathscr{D}^{I_1}_{e_i}\right)+2\sum_{s=t+1}^{T}\left|Cov\left(A^{II}_{3,it}(m),A^{II}_{3,is}(m)\middle|\mathscr{D}^{I_1}_{e_i}\right)\right|\right\}\leq c_{13}\xi_N^2\left(\frac{\varepsilon}{T}\right)^{2/3}\\ \text{and}\quad \max_{i\in I_2,t\in[T]}\left\|\tilde{\Phi}^{(1)}_{it}\right\|_2\leq c_{14}\xi_N\end{array}\right)$$

$$\mathscr{A}_{3,N,i}=\left(\begin{array}{c}\max_{m\in[n_{\mathbb{S}}],t\in[T]}\left\{Var\left(A^{II}_{3,it}(m)\middle|\mathscr{D}^{I_1}_{e_i}\right)+2\sum_{s=t+1}^{T}\left|Cov\left(A^{II}_{3,it}(m),A^{II}_{3,is}(m)\middle|\mathscr{D}^{I_1}_{e_i}\right)\right|\right\}\leq c_{13}\xi_N^2\left(\frac{\varepsilon}{T}\right)^{2/3}\\ \text{and}\quad \max_{t\in[T]}\left\|\tilde{\Phi}^{(1)}_{it}\right\|_2\leq c_{14}\xi_N\end{array}\right).$$

Then, we have $\mathbb{P}(\mathscr{A}^c_{3,N})\to 0$ and

$$\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_2,m\in[n_{\mathbb{S}}]}\left|A^{II}_{3,i}(m)\right|>c_{12}\eta_N\right)\\
&\leq \mathbb{P}\left(\max_{i\in I_2,m\in[n_{\mathbb{S}}]}\left|A^{II}_{3,i}(m)\right|>c_{12}\eta_N,\mathscr{A}_{3,N}\right)+\mathbb{P}(\mathscr{A}^c_{3,N})\\
&\leq \sum_{i\in[I_2]}\mathbb{P}\left(\max_{m\in[n_{\mathbb{S}}]}\left|A^{II}_{3,i}(m)\right|>c_{12}\eta_N,\mathscr{A}_{3,N}\right)+\mathbb{P}(\mathscr{A}^c_{3,N})\\
&\leq \sum_{i\in[I_2]}\mathbb{P}\left(\max_{m\in[n_{\mathbb{S}}]}\left|A^{II}_{3,i}(m)\right|>c_{12}\eta_N,\mathscr{A}_{3,N,i}\right)+\mathbb{P}(\mathscr{A}^c_{3,N})\\
&= \sum_{i\in[I_2]}\mathbb{E}\mathbb{P}\left(\max_{m\in[n_{\mathbb{S}}]}\left|A^{II}_{3,i}(m)\right|>c_{12}\eta_N\middle|\mathscr{D}^{I_1}_{e_i}\right)\mathbf{1}\{\mathscr{A}_{3,N,i}\}+\mathbb{P}(\mathscr{A}^c_{3,N})\\
&\leq \sum_{i\in I_2,m\in[n_{\mathbb{S}}]}\exp\left(-\frac{c_9c_{12}^2T^2\eta_N^2}{c_{13}T\xi_N^2\left(\frac{\varepsilon}{T}\right)^{2/3}+c_{14}^2\xi_N^2+c_{12}c_{14}T\eta_N\xi_{N,1}(\log T)^2}\right)+o(1)=o(1),
\end{aligned}$$

where second inequality is by the union bound, the first inequality is by $\mathscr{A}_{3,N} \subset \mathscr{A}_{3,N,i}$, the equality is by the fact that $\mathscr{A}_{3,N,i}$ is $\mathscr{D}_{e_i}^{I_1}$ measurable, and the last inequality is by Lemma B.12(ii), the definition of $\mathscr{A}_{3,N,i}$, and the fact that $\{\epsilon_{it}\}_{t\in[T]}$ is strong mixing given $\mathscr{D}_{e_i}^{I_1}$. This implies

$$\max_{i\in I_2, m\in[n_{\mathbb{S}}]} \sup_{s\in\mathbb{S}_m} \left\| \frac{1}{T}\sum_{t\in[T]} \left[A_{3,it}^{I}(s) - A_{3,it}^{I}(s_m)\right] \right\|_2 = O_p(\eta_N).$$

Last, we turn to the first term of (B.37). Denote $A_{3,i}^{I,k}(s_m)$ as the k^{th} element of $A_{3,i}^{I,k}(s_m)$ and the event set $\mathscr{A}_{4,N,i} = \{\max_{t\in[T]} \left\|\tilde{\Phi}_{it}^{(1)}\right\|_2 \leq c_{14}\xi_N\}$. Similarly, we have $\mathbb{P}(\cap_{i\in I_2}\mathscr{A}_{4,N,i}^{c}) = \mathbb{P}(\max_{i\in I_2, t\in[T]} \left\|\tilde{\Phi}_{it}^{(1)}\right\|_2 > c_{14}\xi_N) = o(1)$. Following the same argument as above, we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_2, m\in[n_{\mathbb{S}}]} \left|A_{3,i}^{I,k}(s_m)\right| > c_{12}\eta_N\right) \\
&\leq \sum_{m\in[n_{\mathbb{S}}], i\in I_2} \mathbb{E}\mathbb{P}\left(\left|A_{3,i}^{I,k}(s_m)\right| > \frac{c_{12}}{2}\eta_N \middle| \mathscr{D}_{e_i}^{I_1}\right) 1\{\mathscr{A}_{4,N,i}\} + o(1) \\
&\leq \sum_{m\in[n_{\mathbb{S}}], i\in I_2} \exp\left(-\frac{c_9 c_{12}^2 T^2 \eta_N^2/4}{c_{14}^2 T\xi_N^2 + c_{12}c_{14}T\eta_N\xi_N \log T\left(\log\log T\right)/2}\right) + o(1) = o(1), \qquad \text{(B.40)}
\end{aligned}$$

where the second inequality combines Lemma B.12(i). Combining (B.34), (B.36) and (B.40), we have

$$\max_{i\in I_2} \frac{|A_{3,i}|}{\left\|\dot{\Delta}_{i,u}\right\|_2} = O_p(\eta_N).$$

We now turn to $A_{2,i}$. Let $A_{2,it}^{I} = \left(\tau - \mathbf{1}\{w_{1,it} \leq 0\}\right)\tilde{\Phi}_{it}^{(1)} - \mathbb{E}\left(\left(\tau - \mathbf{1}\{w_{1,it}\leq 0\}\right)\tilde{\Phi}_{it}^{(1)} \middle| \mathscr{D}_{e_i}^{I_1}\right)$. Then $A_{2,i} = \frac{1}{T}\sum_{t=1}^{T}\left(A_{2,it}^{I}\right)'\dot{\Delta}_{i,u}$. By conditional Bernstein's inequality and similarly as (B.40), we can show that

$$\mathbb{P}\left\{\max_{i\in I_2}\left\|\frac{1}{T}\sum_{t=1}^{T} A_{2,it}^{I}\right\|_2 \geq c_{12}\eta_N\right\} = o(1),$$

which implies $\max_{i\in I_2}\frac{|A_{2,i}|}{\left\|\dot{\Delta}_{i,u}\right\|_2} = O_p(\eta_N)$. By similar arguments for $A_{3,i}$, we can also show that $\max_{i\in I_2}\frac{|A_{5,i}|}{\left\|\dot{\Delta}_{i,u}\right\|_2} = O_p(\eta_N)$. For $A_{6,i}$, we note that

$$\begin{aligned}
\max_{i\in I_2}\frac{|A_{6,i}|}{\left\|\dot{\Delta}_{i,u}\right\|_2} &= \max_{i\in I_2}\frac{\left|\frac{1}{T}\sum_{t=1}^{T}\int_0^{\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}} \mathbb{E}\left(\mathbf{1}\{w_{1,it}\leq s\} - \mathbf{1}\{\epsilon_{it}\leq s\}\middle|\mathscr{D}_{e_i}^{I_1}\right)ds\right|}{\left\|\dot{\Delta}_{i,u}\right\|_2} \\
&= \max_{i\in I_2}\frac{\left|\frac{1}{T}\sum_{t=1}^{T}\int_0^{\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}} \left[\mathfrak{F}_{it}\left(u_i^{0\prime}\left(\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^{0}\right) + s\right) - \mathfrak{F}_{it}(s)\right]ds\right|}{\left\|\dot{\Delta}_{i,u}\right\|_2} \\
&\leq \max_{i\in I_2}\frac{\left|\frac{1}{T}\sum_{t=1}^{T}\int_0^{\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}} u_i^{0\prime}\left(\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^{0}\right)\mathfrak{f}_{it}(s)ds\right|}{\left\|\dot{\Delta}_{i,u}\right\|_2} \\
&\leq \max_{i\in I_2}\frac{\frac{\bar{\mathfrak{f}}}{T}\sum_{t=1}^{T}\left|\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}\right|\left|u_i^{0\prime}\left(\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^{0}\right)\right|}{\left\|\dot{\Delta}_{i,u}\right\|_2}
\end{aligned}$$

$$\leq \bar{\mathfrak{f}} \max_{i\in I_2} \sqrt{\frac{1}{T}\sum_{t\in[T]} \left\|\tilde{\Phi}_{it}^{(1)}\right\|^2} \sqrt{\frac{1}{T}\sum_{t\in[T]} \left\|\tilde{\Phi}_{it}^{(1)} - \Phi_{it}^0\right\|^2} \left\|u_i^0\right\|_2$$
$$= O_p(\eta_N),$$

where the first inequality is by mean-value theorem and the other inequalities holds by similar reasons as those in (B.33).

Last, for $A_{7,i}$, we have

$$\begin{aligned}
&\frac{1}{T}\sum_{t\in[T]} \left[\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}\left(\mathbf{1}\{\epsilon_{it}\leq 0\} - \mathbf{1}\{w_{1,it}\leq 0\}\right)\right] \\
&= \frac{1}{T}\sum_{t\in[T]} \mathbb{E}\left[\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}\left(\mathbf{1}\{\epsilon_{it}\leq 0\} - \mathbf{1}\{w_{1,it}\leq 0\}\right)\middle|\mathscr{D}_{e_i}^{I_1}\right] \\
&+ \frac{1}{T}\sum_{t\in[T]} \left\{\left[\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}\left(\mathbf{1}\{\epsilon_{it}\leq 0\} - \mathbf{1}\{w_{1,it}\leq 0\}\right)\right] - \mathbb{E}\left[\tilde{\Phi}_{it}^{(1)\prime}\dot{\Delta}_{i,u}\left(\mathbf{1}\{\epsilon_{it}\leq 0\} - \mathbf{1}\{w_{1,it}\leq 0\}\right)\middle|\mathscr{D}_{e_i}^{I_1}\right]\right\} \\
&:= A_{7,i}^{I} + A_{7,i}^{II},
\end{aligned}$$

Then, following the same analyses of $A_{1,i}$ and $A_{3,i}$, we have

$$\max_{i\in I_2} \frac{\left|A_{7,i}^{I}\right|}{\left\|\dot{\Delta}_{i,u}\right\|_2} = O_p(\eta_N) \quad \text{and} \quad \max_{i\in I_2} \frac{\left|A_{7,i}^{II}\right|}{\left\|\dot{\Delta}_{i,u}\right\|_2} = O_p(\eta_N),$$

which implies $\max_{i\in I_2} \frac{|A_{7,i}|}{\left\|\dot{\Delta}_{i,u}\right\|_2} = O_p(\eta_N)$. ■

Recall that event $\mathscr{A}_4 = \left\{\max_{i\in I_2}\left\|O_j^{(1)\prime}\dot{u}_{i,j}^{(1)} - u_{i,j}^0\right\|_2 = O(\eta_N), \forall j\in[p]\cup\{0\}\right\}$ where we define $q_i^{III} = \inf_{\Delta} \frac{\left[\frac{1}{N_2}\sum_{i\in I_2}\left(\tilde{\Psi}_{it}^{(1)\prime}\Delta\right)^2\right]^{\frac{3}{2}}}{\frac{1}{N_2}\sum_{i\in I_2}\left|\tilde{\Psi}_{it}^{(1)\prime}\Delta\right|^3}$.

Lemma B.16 . *Suppose Assumptions 1–5 hold. Then for $\{B_{1,t},\cdots,B_{6,t}\}_{t\in[T]}$ defined in (A.8), for any constant $0 < c_{11} < \min(\frac{3\underline{\mathfrak{f}}}{\bar{\mathfrak{f}}'}, 1)$, we have*

$$\max_{t\in[T]} \frac{|B_{m,t}|}{\left\|\dot{\Delta}_{t,v}\right\|_2} = O_p(\eta_N) \quad \forall m\in\{1,2,3,5,6\},$$
$$|B_{4,t}| \geq \min\left(\frac{\left(3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'\right)c_\psi\left\|\dot{\Delta}_{t,v}\right\|_2^2}{12}, \frac{\left(3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'\right)\sqrt{c_\psi}q_i^{III}\left\|\dot{\Delta}_{t,v}\right\|_2}{6\sqrt{2}}\right), \quad \forall t\in[T].$$

Proof. We first deal with $B_{4,t}$. Let $w_{4,it} = \dot{\Psi}_{it}^{(1)\prime}\dot{\Delta}_{t,v}$. Following the same arguments as used to derive the lower bound for $A_{4,i}$ in the proof of Lemma B.15 by replacing $w_{2,it}$ with $w_{4,it}$, we can show that, for $t\in[T]$ and any constants $c_{11} < \min(\frac{3\underline{\mathfrak{f}}}{\bar{\mathfrak{f}}'}, 1)$,

$$|B_{4,t}| \geq \min\left(\frac{\left(3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'\right)c_\psi\left\|\dot{\Delta}_{t,v}\right\|_2^2}{12}, \frac{\left(3c_{11}^2\underline{\mathfrak{f}} - c_{11}^3\bar{\mathfrak{f}}'\right)\sqrt{c_\psi}q_i^{III}\left\|\dot{\Delta}_{t,v}\right\|_2}{6\sqrt{2}}\right).$$

For $B_{1,t}$, we have

$$B_{1,t}=\left(\frac{1}{N_2}\sum_{i\in I_2}\Psi_{it}^{0}\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right)'\dot{\Delta}_{t,v}:=\left(\frac{1}{N_2}\sum_{i\in I_2}B_{1,it}^{I}\right)'\dot{\Delta}_{t,v}.$$

Conditional on the fixed effects $\left\{V_j^0\right\}_{j\in[p]\cup\{0\}}$ and $\left\{W_j^0\right\}_{j\in[p]}$, the randomness in $B_{1,t}$ is from $\{\epsilon_{it}\}_{i\in I_2,t\in[T]}$ and $\{e_{j,it}\}_{j\in[p],i\in I_2,t\in[T]}$, which are independent across i. Owing to this, by conditional Hoeffding's inequality, we can show that

$$\begin{aligned}&\mathbb{P}\left\{\max_{t\in[T]}\frac{1}{N_2}\left\|\sum_{i\in I_2}B_{1,it}^{I,k}\right\|_2>c_{15}\eta_N\middle|\mathscr{D}\right\}\leq\sum_{t\in[T]}\mathbb{P}\left\{\left\|\sum_{i\in I_2}B_{1,it}^{I,k}\right\|_2>c_{15}N_2\eta_N\middle|\mathscr{D}\right\}\\&\leq 2\sum_{t\in[T]}\exp\left(-\frac{c_{15}^2N_2^2\eta_N^2}{4M^2\xi_N^2N_2}\right)=o(1),\end{aligned}\tag{B.41}$$

with $B_{1,it}^{I,k}$ being the k^{th} element in $B_{1,it}^{I}$, c_{15} is a positive constant, where the second inequality is by Hoeffding's inequality with the fact that $\max_{i\in I_2,t\in[T]}\left|B_{1,it}^{I,k}\right|\leq M\xi_N$ a.s. by Assumption 1(v) and Lemma B.13(i). It follows that $\max_{t\in[T]}\frac{|B_{1,t}|}{\left\|\dot{\Delta}_{t,v}\right\|_2}=O_p\left(\eta_N\right)$. If we use the conditional Bernstein's inequality for the independent sequence rather than the Hoeffding's inequality above, we can show that $\max_{t\in[T]}\frac{1}{N_2}\left\|\sum_{i\in I_2}B_{1,it}^{I,k}\right\|_2=O_p\left(\sqrt{\frac{\log N\vee T}{N}}\right)$, but here we only need to show the uniform convergence rate to be η_N.

Let $X_{0,it}=1$. As for $B_{2,t}$, note that

$$\begin{aligned}\max_{t\in[T]}\frac{|B_{2,t}|}{\left\|\dot{\Delta}_{t,v}\right\|_2}&=\max_{t\in[T]}\frac{\left|\left[\frac{1}{N_2}\sum_{i\in I_2}\left(\dot{\Psi}_{it}^{(1)}-\Psi_{it}^{0}\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right]'\dot{\Delta}_{t,v}\right|}{\left\|\dot{\Delta}_{t,v}\right\|_2}\\&\leq\max_{t\in[T]}\left\|\frac{1}{N_2}\sum_{i\in I_2}\left(\dot{\psi}_{it}^{(1)}-\Psi_{it}^{0}\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_2\\&\leq\max_{t\in[T]}\frac{1}{N_2}\sum_{i\in I_2}\left\|\dot{\Psi}_{it}^{(1)}-\Psi_{it}^{0}\right\|_2=\max_{t\in[T]}\frac{1}{N_2}\sum_{i\in I_2}\sqrt{\sum_{j=0}^{p}\left\|O_0^{(1)\prime}\dot{u}_{i,j}^{(1)}-u_{i,j}^{0}\right\|_2^2X_{j,it}^2}\\&\leq\max_{i\in I_2}\max_{j\in[p]\cup\{0\}}\left(||O_0^{(1)\prime}\dot{u}_{i,j}^{(1)}-u_{i,j}^{0}||_2\right)\left[\frac{1}{N_2}\sum_{i\in I_2}\sum_{j=0}^{p}X_{j,it}^2\right]^{1/2}=O_p(\eta_N),\end{aligned}\tag{B.42}$$

where the first inequality is by Cauchy's inequality, the second inequality is by Jensen's inequality, and the last equality is by Theorem 3.2(i) and Assumption 1(iv).

Next, we deal with $B_{3,t}$. Following similar arguments as used in (B.42), we obtain that

$$\max_{t\in[T]}\frac{\left|\frac{1}{N_2}\sum_{i\in I_2}\left(\dot{\Psi}_{it}^{(1)}-\Psi_{it}^{0}\right)'\dot{\Delta}_{t,v}\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{w_{3,it}\leq 0\right\}\right]\right|}{\left\|\dot{\Delta}_{t,v}\right\|_2}=O_p\left(\eta_N\right),$$

which implies

$$\max_{t\in[T]}\frac{|B_{3,t}|}{2\left\|\dot{\Delta}_{t,v}\right\|_2}=\max_{t\in[T]}\frac{\left|\left\{\frac{1}{N_2}\sum_{i\in I_2}\Psi_{it}^0\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{w_{3,it}\leq 0\right\}\right]\right\}'\dot{\Delta}_{t,v}\right|}{\left\|\dot{\Delta}_{t,v}\right\|_2}+O_p\left(\eta_N\right)$$
$$\leq\max_{t\in[T]}\left\|\frac{1}{N_2}\sum_{i\in I_2}\Psi_{it}^0\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{w_{3,it}\leq 0\right\}\right]\right\|_2+O_p\left(\eta_N\right)$$
$$\leq\max_{t\in[T]}\frac{1}{N_2}\sum_{i\in I_2}\left\|\Psi_{it}^0\right\|_2\mathbf{1}\left\{|\epsilon_{it}|\leq\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\left\|\dot{u}_{i,j}^{(1)}-O_j^{(1)}u_{i,j}^0\right\|_2\right\}+O_p\left(\eta_N\right)$$
$$\equiv\Xi_{NT}+O_p\left(\eta_N\right).$$

Now define the event set $\mathscr{B}_{N,1}(M)=\{\max_{i\in I_2,j\in[p]\cup\{0\}}||\dot{u}_{i,j}^{(1)}-O_j^{(1)}u_{i,j}^0||\leq M\eta_N\}$. Then, Theorem 3.2(i) implies, for any $e>0$, there is a sufficiently large M such that $\mathbb{P}(\mathscr{B}_{N,1}^c(M))\leq e$. Recall that $\mathscr{D}_{e_{it}}$ is the σ-field generated by $e_{it}\cup\left\{V_j^0\right\}_{j\in[p]\cup\{0\}}\cup\left\{W_j^0\right\}_{j\in[p]}$. Then, we have

$$\mathbb{P}\left(\Xi_{NT}\geq C\eta_N\right)\leq\mathbb{P}\left(\Xi_{NT}\geq C\eta_N,\mathscr{B}_{N,1}\right)+e$$
$$\leq\mathbb{P}\left(\frac{1}{N_2}\sum_{i\in I_2}\left\|\Psi_{it}^0\right\|_2\mathbf{1}\left\{|\epsilon_{it}|\leq M\eta_N\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\right\}\geq C\eta_N\right)+e$$
$$\leq\mathbb{P}\left(\max_{t\in[T]}B_{3,t}^I\geq C\eta_N+\max_{t\in T}B_{3,t}^{II}\right)+e,\tag{B.43}$$

where

$$B_{3,t}^I=\frac{1}{N_2}\sum_{i\in I_2}\left[\left\|\Psi_{it}^0\right\|_2\mathbf{1}\left\{|\epsilon_{it}|\leq M\eta_N\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\right\}-B_{3,t}^{II}\right]\quad\text{and}$$
$$B_{3,t}^{II}=\frac{1}{N_2}\sum_{i\in I_2}\mathbb{E}\left(\left\|\Psi_{it}^0\right\|_2\mathbf{1}\left\{|\epsilon_{it}|\leq M\eta_N\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\right\}\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right)$$
$$=\frac{1}{N_2}\sum_{i\in I_2}\mathbb{E}\left(\left\|\Psi_{it}^0\right\|_2\mathbf{1}\left\{|\epsilon_{it}|\leq M\eta_N\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\right\}\middle|\mathscr{D}_{e_{it}}\right),$$

and the second equality for $B_{3,t}^{II}$ holds by Assumption 1(i). Following this, we show that

$$\max_{t\in[T]}B_{3,t}^{II}=\max_{t\in[T]}\frac{1}{N_2}\sum_{i\in I_2}\left\|\Psi_{it}^0\right\|_2\left[\mathfrak{F}_{it}\left(M\eta_N\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\right)-\mathfrak{F}_{it}\left(-M\eta_N\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\right)\right]$$
$$\leq\max_{t\in[T]}\frac{1}{N_2}\sum_{i\in I_2}\sqrt{1+p\max_{j\in[p]}|X_{j,it}|^2}\frac{2(1+p)M^2}{c_\sigma}\eta_N$$
$$\leq\frac{2(1+p)^2M^2C}{c_\sigma}\eta_N\quad\text{a.s.},\tag{B.44}$$

where $\mathfrak{F}_{it}(\cdot)$ is the conditional CDF of ϵ_{it} given $\mathscr{D}_{e_i}$ who also has bounded PDF by Assumption 1(vii), the first inequality is by mean-value theorem and facts that $\left\|\Psi_{it}^0\right\|_2^2\leq M^2\left(1+p\max_{j\in[p]}|X_{j,it}|^2\right)$ together with Lemma B.13(i), and the second inequality is due to Assumption 1(iv). In addition, given $\{\mathscr{D}_{e_{it}}\}_{i\in I_2}$, $\{\epsilon_{it}\}_{i\in I_2}$ are still independent across i. Therefore, by Hoeffding's inequality and similar arguments for term $B_{1,t}$ in (B.41), we can show that

$$\begin{aligned}
\mathbb{P}\left(\max_{t\in[T]}\left\|B^I_{3,t}\right\|_2 > c_{12}\eta_N\right) &\leq \sum_{t\in[T]}\mathbb{P}\left(\left\|B^I_{3,t}\right\|_2 > c_{12}\eta_N\right) \\
&= \sum_{t\in[T]}\mathbb{E}\mathbb{P}\left(\left\|B^I_{3,t}\right\|_2 > c_{12}\eta_N\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right) = o(1). \qquad \text{(B.45)}
\end{aligned}$$

Combining (B.43)-(B.45), we obtain that $\max_{t\in[T]}\frac{|B_{3,t}|}{\left\|\dot{\Delta}_{t,v}\right\|_2} = O_p(\eta_N)$.

For $B_{5,t}$, we observe that

$$\begin{aligned}
B_{5,t} &= \frac{1}{N_2}\sum_{i\in I_2}\left\{\int_0^{\Psi^{0\prime}_{it}\dot{\Delta}_{t,v}}\left[(\mathbf{1}\{\epsilon_{it}\leq s\}-\mathbf{1}\{\epsilon_{it}\leq 0\}) - \mathbb{E}\left((\mathbf{1}\{\epsilon_{it}\leq s\}-\mathbf{1}\{\epsilon_{it}\leq 0\})\middle|\mathscr{D}_{e_{it}}\right)\right]ds\right\} \\
&\quad + \frac{1}{N_2}\sum_{i\in I_2}\left\{\int_{\Psi^{0\prime}_{it}\dot{\Delta}_{t,v}}^{\dot{\Psi}^{(1)\prime}_{it}\dot{\Delta}_{t,v}}\left[(\mathbf{1}\{\epsilon_{it}\leq s\}-\mathbf{1}\{\epsilon_{it}\leq 0\}) - \mathbb{E}\left((\mathbf{1}\{\epsilon_{it}\leq s\}-\mathbf{1}\{\epsilon_{it}\leq 0\})\middle|\mathscr{D}_{e_{it}}\right)\right]ds\right\} \\
&:= B^I_{5,t} + B^{II}_{5,t}.
\end{aligned}$$

Following similar arguments for $A_{3,i}$ with Bernstein's inequality replacing by Hoeffding's inequality, we can show that $\max_{t\in[T]}\frac{\left|B^I_{5,t}\right|}{\left\|\dot{\Delta}_{t,v}\right\|_2} = O_p(\eta_N)$. Besides, we obtain that

$$\max_{t\in[T]}\frac{\left|B^{II}_{5,t}\right|}{\left\|\dot{\Delta}_{t,v}\right\|_2} \leq \max_{t\in[T]}\frac{4}{N_2}\sum_{i\in I_2}\left\|\dot{\Psi}^{(1)}_{it}-\Psi^0_{it}\right\|_2 = O_p(\eta_N),$$

which implies $\max_{t\in[T]}\frac{|B_{5,t}|}{\left\|\dot{\Delta}_{t,v}\right\|_2} = O_p(\eta_N)$.

For $B_{6,t}$, we first note that

$$\max_{t\in[T]}\frac{\left|\frac{1}{N_2}\sum_{i\in I_2}\int_{\Psi^{0\prime}_{it}\dot{\Delta}_{t,v}}^{\dot{\Psi}^{(1)\prime}_{it}\dot{\Delta}_{t,v}}\left[\mathbf{1}\{\epsilon_{it}\leq s\}-\mathbf{1}\{w_{3,it}\leq s\}\right]\right|}{\left\|\dot{\Delta}_{t,v}\right\|_2} \leq \max_{t\in[T]}\frac{2}{N_2}\sum_{i\in I_2}\left\|\dot{\Psi}^{(1)}_{it}-\Psi^0_{it}\right\|_2 = O_p(\eta_N),$$

and this implies

$$\begin{aligned}
\max_{t\in[T]}\frac{|B_{6,t}|}{\left\|\dot{\Delta}_{t,v}\right\|_2} &= \max_{t\in[T]}\frac{\left|\frac{1}{N_2}\sum_{i\in I_2}\int_0^{\dot{\Psi}^{(1)\prime}_{it}\dot{\Delta}_{t,v}}(\mathbf{1}\{w_{3,it}\leq s\}-\mathbf{1}\{\epsilon_{it}\leq s\})\,ds\right|}{\left\|\dot{\Delta}_{t,v}\right\|_2} \\
&= \max_{t\in[T]}\frac{\left|\frac{1}{N_2}\sum_{i\in I_2}\int_0^{\Psi^{0\prime}_{it}\dot{\Delta}_{t,v}}(\mathbf{1}\{w_{3,it}\leq s\}-\mathbf{1}\{\epsilon_{it}\leq s\})\,ds\right|}{\left\|\dot{\Delta}_{t,v}\right\|_2} + O_p(\eta_N).
\end{aligned}$$

In addition, for the first term on the RHS of the last equality, we have

$$\max_{t\in[T]}\frac{\left|\frac{1}{N_2}\sum_{i\in I_2}\int_0^{\Psi^{0\prime}_{it}\dot{\Delta}_{t,v}}(\mathbf{1}\{w_{3,it}\leq s\}-\mathbf{1}\{\epsilon_{it}\leq s\})\,ds\right|}{\left\|\dot{\Delta}_{t,v}\right\|_2}$$

$$\leq \max_{t\in[T]} \frac{\frac{1}{N_2}\sum_{i\in I_2}\int_0^{\left\|\Psi_{it}^0\right\|_2\left\|\dot{\Delta}_{t,v}\right\|_2}\left(\mathbf{1}\left\{|\epsilon_{it}-s|\leq \sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\left\|\dot{u}_{i,j}^{(1)}-O_j^{(1)}u_{i,j}^0\right\|_2\right\}\right)ds}{\left\|\dot{\Delta}_{t,v}\right\|_2}$$

$$= \max_{t\in[T]}\frac{1}{N_2}\sum_{i\in I_2}\int_0^1 \left\|\Psi_{it}^0\right\|_2 \mathbf{1}\left\{\left|\epsilon_{it}-\left\|\Psi_{it}^0\right\|_2\left\|\dot{\Delta}_{t,v}\right\|_2 s\right|\leq \sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\left\|\dot{u}_{i,j}^{(1)}-O_j^{(1)}u_{i,j}^0\right\|_2\right\}ds$$

$$\leq \sup_{s\geq 0,t\in[T]}\frac{1}{N_2}\sum_{i\in I_2}\left\|\Psi_{it}^0\right\|_2 \mathbf{1}\left\{|\epsilon_{it}-s|\leq \sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\left\|\dot{u}_{i,j}^{(1)}-O_j^{(1)}u_{i,j}^0\right\|_2\right\}.$$

Following the same argument for $B_{3,t}$, we only need to upper bound the RHS of the last display by $\sup_{s\geq 0,t\in[T]} B_{6,t}^I(s)$ on $\mathscr{B}_{N,1}(M)$ for some sufficiently large but fixed constant M, where

$$B_{6,t}^I(s) = \frac{1}{N_2}\sum_{i\in I_2}\left\|\Psi_{it}^0\right\|_2 \mathbf{1}\left\{|\epsilon_{it}-s|\leq M\eta_N\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\right\}.$$

Let

$$B_{6,t}^{II}(s) = \mathbb{E}(B_{6,t}^I(s)|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}) = \frac{1}{N_2}\sum_{i\in I_2}\mathbb{E}\left(\left\|\Psi_{it}^0\right\|_2 \mathbf{1}\left\{|\epsilon_{it}-s|\leq M\eta_N\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\right\}\middle|\mathscr{D}_{e_{it}}\right).$$

We note that $\sup_{s\geq 0,t\in[T]} B_{6,t}^{II}(s) = O_p(\eta_N)$. Similar to the arguments in (B.44), to show $\max_{t\in[T]}\frac{|B_{6,t}|}{\left\|\dot{\Delta}_{t,v}\right\|_2} = O_p(\eta_N)$, it suffices to show

$$\mathbb{P}\left(\sup_{s\geq 0,t\in[T]}\left|B_{6,t}^I(s)-B_{6,t}^{II}(s)\right|>c_{12}\eta_N\right)=o(1). \tag{B.46}$$

Further denote

$$B_{6,t}^{III}(s) = \frac{1}{N_2}\sum_{i\in I_2}\left\|\Psi_{it}^0\right\|_2 \mathbf{1}\left\{\epsilon_{it}>s-M\eta_N\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\right\},$$
$$B_{6,t}^{IV}(s) = \frac{1}{N_2}\sum_{i\in I_2}\left\|\Psi_{it}^0\right\|_2 \mathbf{1}\left\{\epsilon_{it}\geq s+M\eta_N\sum_{j\in[p]\cup\{0\}}\left\|v_{t,j}^0\right\|_2\right\}.$$

Then, we have $B_{6,t}^I(s) = B_{6,t}^{III}(s) - B_{6,t}^{IV}(s)$ and thus,

$$\sup_{s>T^{1/4},t\in[T]}\left|B_{6,t}^I(s)-B_{6,t}^{II}(s)\right|$$
$$\leq \sup_{s>T^{1/4},t\in[T]}\left|B_{6,t}^{III}(s)-\mathbb{E}\left(B_{6,t}^{III}(s)\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right)\right| + \sup_{s>T^{1/4},t\in[T]}\left|B_{6,t}^{IV}(s)-\mathbb{E}\left(B_{6,t}^{IV}(s)\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right)\right|$$
$$\leq \max_{t\in[T]}B_{6,t}^{III}(T^{1/4}) + \max_{t\in[T]}\mathbb{E}\left(B_{6,t}^{III}(T^{1/4})\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right) + \max_{t\in[T]}B_{6,t}^{IV}(T^{1/4}) + \max_{t\in[T]}\mathbb{E}\left(B_{6,t}^{IV}(T^{1/4})\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right)$$
$$\leq \max_{t\in[T]}\left|B_{6,t}^{III}(T^{1/4})-\mathbb{E}\left(B_{6,t}^{III}(T^{1/4})\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right)\right| + 2\max_{t\in[T]}\mathbb{E}\left(B_{6,t}^{III}(T^{1/4})\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right)$$
$$+ \max_{t\in[T]}\left|B_{6,t}^{IV}(T^{1/4})-\mathbb{E}\left(B_{6,t}^{IV}(T^{1/4})\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right)\right| + 2\max_{t\in[T]}\mathbb{E}\left(B_{6,t}^{IV}(T^{1/4})\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right),$$

where the second inequality holds because both $B_{6,t}^{III}(s)$ and $B_{6,t}^{IV}(s)$ are non-decreasing in s. Further note that

$$\left\|\Psi_{it}^0\right\|_2 \mathbf{1}\left\{\epsilon_{it} > s - M\eta_N \sum_{j\in[p]\cup\{0\}} \left\|v_{t,j}^0\right\|_2\right\} \quad \text{and} \quad \left\|\Psi_{it}^0\right\|_2 \mathbf{1}\left\{\epsilon_{it} > s + M\eta_N \sum_{j\in[p]\cup\{0\}} \left\|v_{t,j}^0\right\|_2\right\}$$

are independent across $i \in I_2$ given $\{\mathscr{D}_{e_{it}}\}_{i\in I_2}$. Therefore, following the same argument in the analysis of $A_{6,t}$ with the Bernstein's inequality replaced by the Hoeffding's inequality, we have

$$\max_{t\in[T]}\left|B_{6,t}^{III}(T^{1/4}) - \mathbb{E}\left(B_{6,t}^{III}(T^{1/4})\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right)\right| = O_p(\eta_N), \quad \max_{t\in[T]}\mathbb{E}\left(B_{6,t}^{III}(T^{1/4})\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right) = O_p(\eta_N)$$
$$\max_{t\in[T]}\left|B_{6,t}^{IV}(T^{1/4}) - \mathbb{E}\left(B_{6,t}^{IV}(T^{1/4})\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right)\right| = O_p(\eta_N), \quad \max_{t\in[T]}\mathbb{E}\left(B_{6,t}^{IV}(T^{1/4})\middle|\{\mathscr{D}_{e_{it}}\}_{i\in I_2}\right) = O_p(\eta_N),$$

which implies

$$\sup_{s>T^{1/4},t\in[T]}\left|B_{6,t}^{I}(s) - B_{6,t}^{II}(s)\right| = O_p(\eta_N).$$

In addition, following the same analysis in (B.37) and (B.38), we have

$$\sup_{s\in[0,T^{1/4}],t\in[T]}\left|B_{6,t}^{I}(s) - B_{6,t}^{II}(s)\right| = O_p(\eta_N),$$

which leads to the desired result that $\max_{t\in[T]} \frac{|B_{6,t}|}{\left\|\dot{\Delta}_{t,v}\right\|_2} = O_p(\eta_N)$. ■

Recall for $i \in I_3$, $\dot{\mathcal{H}}_i\left(\{u_{i,j}\}_{j\in[p]\cup\{0\}}\right) = \frac{1}{T}\sum_{t=1}^T\left\{\left[\tau - \mathfrak{F}_{it}\left(g_{it}(\{u_{i,j}\}_{j\in[p]\cup\{0\}})\right)\right]\dot{\varpi}_{it}\right\}$, where

$$g_{it}(\{u_{i,j}\}_{j\in[p]\cup\{0\}}) = u_{i,0}'\dot{v}_{t,0}^{(1)} + \sum_{j\in[p]} u_{i,j}'\dot{v}_{t,j}^{(1)}X_{j,it} - u_{i,0}^{0\prime}v_{t,0}^0 - \sum_{j\in[p]} u_{i,j}^{0\prime}v_{t,j}^0 X_{j,it},$$

and $\dot{\varpi}_{it} = \left(\dot{v}_{t,0}^{(1)\prime}, X_{1,it}\dot{v}_{t,1}^{(1)\prime}, \cdots, X_{p,it}\dot{v}_{t,p}^{(1)\prime}\right)'$.

Lemma B.17 *Under Assumptions 1–5, the second-order derivative of $\dot{\mathcal{H}}_i\left(\{u_{i,j}\}_{j=0}^p\right)$ is bounded in probability.*

Proof. Noted that

$$\frac{\dot{\mathcal{H}}_i\left(\{u_{i,j}\}_{j=0}^p\right)}{\partial u_i'} = \frac{1}{T}\sum_{t=1}^T -\mathfrak{f}_{it}\left(u_{i,0}'\dot{v}_{t,0}^{(1)} - u_{i,0}^{0\prime}v_{t,0}^0 + X_{1,it}u_{i,1}'\dot{v}_{t,1}^{(1)} - X_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right)\dot{\varpi}_{it}\dot{\varpi}_{it}'.$$

For notation simplicity, we focus on the case with $p = 1$ and denote $u_i = \left(u_{i,0}', u_{i,1}'\right)'$. Further denote $\dot{\mathcal{H}}_{i,k}\left(\{u_{i,j}\}_{j=0}^p\right)$ as the k^{th} element in $\dot{\mathcal{H}}_i\left(\{u_{i,j}\}_{j=0}^p\right)$ and $\dot{v}_{t,0,k}^{(1)}$ as the k^{th} element in $\dot{v}_{t,0}^{(1)}$. For $k \in [K_0]$, we have

$$\frac{\partial\dot{\mathcal{H}}_{i,k}\left(\{u_{i,j}\}_{j=0}^p\right)}{\partial u_i} = \frac{1}{T}\sum_{t=1}^T -\mathfrak{f}_{it}\left(u_{i,0}'\dot{v}_{t,0}^{(1)} - u_{i,0}^{0\prime}v_{t,0}^0 + X_{1,it}u_{i,1}'\dot{v}_{t,1}^{(1)} - X_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right)\dot{v}_{t,0,k}^{(1)}\dot{\varpi}_{it} \quad \text{and}$$
$$\frac{\partial^2\dot{\mathcal{H}}_{i,k}\left(\{u_{i,j}\}_{j=0}^p\right)}{\partial u_i\partial u_i'} = \frac{1}{T}\sum_{t=1}^T -\mathfrak{f}_{it}'\left(u_{i,0}'\dot{v}_{t,0}^{(1)} - u_{i,0}^{0\prime}v_{t,0}^0 + X_{1,it}u_{i,1}'\dot{v}_{t,1}^{(1)} - X_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right)\dot{v}_{t,0,k}^{(1)}\dot{\varpi}_{it}\dot{\varpi}_{it}'.$$

Therefore, we have

$$
\begin{aligned}
\left\|\frac{\partial^2\dot{\mathcal{H}}_{i,k}\left(\{u_{i,j}\}_{j=0}^p\right)}{\partial u_i \partial u_i'}\right\|_F &\leq \frac{\bar{\mathfrak{f}}'}{T}\sum_{t=1}^T \left\|\dot{v}_{t,0,k}^{(1)}\dot{\varpi}_{it}\dot{\varpi}_{it}'\right\|_F \\
&\leq c\bar{\mathfrak{f}}'\left[\max_{t\in[T]}\left\|\dot{v}_{t,0}^{(1)}\right\|_2\right]\frac{1}{T}\sum_{t=1}^T\|\dot{\varpi}_{it}\|_2^2 \\
&\leq c\bar{\mathfrak{f}}'\left[\max_{t\in[T]}\left\|\dot{v}_{t,0}^{(1)}\right\|_2\right]\left[\max_{t\in[T]}(\left\|\dot{v}_{t,0}^{(1)}\right\|_2^2+\left\|\dot{v}_{t,1}^{(1)}\right\|_2^2)\right]\left(1+\frac{1}{T}\sum_{t\in[T]}X_{1,it}^2\right) = O_p(1),
\end{aligned}
$$

where we use the fact that $\max_{t\in[T]}\left\|\dot{v}_{t,0}^{(1)}\right\|_2 = O_p(1)$ by Theorem 3.2(ii) and Lemma B.13(i).

For $k\in[K_0+1,\cdots,K_0+K_1]$, we have

$$
\frac{\partial\dot{\mathcal{H}}_{i,k}\left(\{u_{i,j}\}_{j=0}^p\right)}{\partial u_i} = \frac{1}{T}\sum_{t=1}^T -\mathfrak{f}_{it}\left(u_{i,0}'\dot{v}_{t,0}^{(1)} - u_{i,0}^{0\prime}v_{t,0}^0 + X_{1,it}u_{i,1}'\dot{v}_{t,1}^{(1)} - X_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right)\dot{v}_{t,1,k-K_0}^{(1)}X_{1,it}\dot{\varpi}_{it} \quad \text{and}
$$

$$
\frac{\partial^2\dot{\mathcal{H}}_{i,k}\left(\{u_{i,j}\}_{j=0}^p\right)}{\partial u_i\partial u_i'} = \frac{1}{T}\sum_{t=1}^T -\mathfrak{f}_{it}'\left(u_{i,0}'\dot{v}_{t,0}^{(1)} - u_{i,0}^{0\prime}v_{t,0}^0 + X_{1,it}u_{i,1}'\dot{v}_{t,1}^{(1)} - X_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right)\dot{v}_{t,1,k-K_0}^{(1)}X_{1,it}\dot{\varpi}_{it}\dot{\varpi}_{it}'.
$$

Therefore, we have

$$
\begin{aligned}
\left\|\frac{\partial^2\dot{\mathcal{H}}_{i,k}\left(\{u_{i,j}\}_{j=0}^p\right)}{\partial u_i \partial u_i'}\right\|_F &\leq c\bar{\mathfrak{f}}'\left[\max_{t\in[T]}\left\|\dot{v}_{t,0}^{(1)}\right\|_2\right]\frac{1}{T}\sum_{t=1}^T|X_{1,it}|\,\|\dot{\varpi}_{it}\|_2^2 \\
&\leq c\bar{\mathfrak{f}}'\left[\max_{t\in[T]}\left\|\dot{v}_{t,0}^{(1)}\right\|_2\right]\left[\max_{t\in[T]}(\left\|\dot{v}_{t,0}^{(1)}\right\|_2^2+\left\|\dot{v}_{t,1}^{(1)}\right\|_2^2)\right]\left(1+\frac{1}{T}\sum_{t\in T}X_{1,it}^3\right) = O_p(1).
\end{aligned}
$$

■

Recall that

$$
\begin{aligned}
&\dot{\varpi}_{it} = \left(\dot{v}_{t,0}^{(1)\prime},\dot{v}_{t,1}^{(1)\prime}X_{1,it},\cdots,\dot{v}_{t,p}^{(1)\prime}X_{p,it}\right)', \\
&\varpi_{it}^0 = \left(\left(O_0^{(1)}v_{t,0}^0\right)',\left(O_1^{(1)}v_{t,1}^0\right)'X_{1,it},\cdots,\left(O_p^{(1)}v_{t,p}^0\right)'X_{p,it}\right)', \\
&\dot{D}_i^I := \frac{1}{T}\sum_{t=1}^T\mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right)\dot{\varpi}_{it}\dot{\varpi}_{it}', \quad \dot{D}_i^{II} := \frac{1}{T}\sum_{t=1}^T\left[\tau-\mathbf{1}\left\{\epsilon_{it}\leq\dot{\Delta}_{t,v}'\Psi_{it}^0\right\}\right]\dot{\varpi}_{it}, \\
&D_i^I = \frac{1}{T}\sum_{t=1}^T\mathfrak{f}_{it}(0)\varpi_{it}^0\varpi_{it}^{0\prime}, \quad D_i^{II} = \frac{1}{T}\sum_{t=1}^T\left[\tau-\mathbf{1}\{\epsilon_{it}\leq 0\}\right]\varpi_{it}^0.
\end{aligned}
$$

Lemma B.18 *Under Assumptions 1–5, we have*

(i) $\max_{i\in I_3}\left\|D_i^{II}\right\|_F = O_p\left(\sqrt{\frac{\log(N\vee T)}{T}}\xi_N\right)$,

(ii) $\max_{i\in\in I_3}\left\|\dot{D}_i^I - D_i^I\right\|_F = O_p(\eta_N)$,

(iii) $\max_{i\in\in I_3}\left\|\dot{D}_i^{II} - D_i^{II} - \frac{1}{T}\sum_{t=1}^T\left[\mathbf{1}\{\epsilon_{it}\leq 0\} - \mathbf{1}\left\{\epsilon_{it}\leq\dot{\Delta}_{t,v}'\Psi_{it}^0\right\}\right]\varpi_{it}^0\right\|_2 = o_p\left((N\vee T)^{-\frac{1}{2}}\right)$.

Proof. Throughout the proof, we assume there is only one regressor $p = 1$ for notation simplicity.

(i) We notice that $\mathbb{E}\left(D_i^{II}\middle|\mathscr{D}\right) = 0$. By conditional Bernstein's inequality, for a positive constant c_{16}, we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_3}\left\|\sum_{t\in[T]}\left[\tau - \mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right]O_1^{(1)}v_{t,1}^0 X_{1,it}\right\|_2 > c_{16}\sqrt{T\log(N\vee T)}\xi_N\middle|\mathscr{D}\right)\\
&\leq \sum_{i\in I_3}\exp\left(-\frac{c_9 c_{16}^2 T\xi_N^2\log(N\vee T)}{\frac{4M^2}{c_\sigma^2}T\xi_N^2 + \frac{2Mc_{16}}{c_\sigma}\xi_N^2\sqrt{T\log(N\vee T)}\log T\log\log T}\right) = o(1),
\end{aligned}$$

where the inequality follows from Lemma B.12(i), Assumption 1(ii), Assumption 1(v), and the fact that $\max_{i\in I_3,t\in[T]}\left\|\left[\tau - \mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right]O_1^{(1)}v_{t,1}^0 X_{1,it}\right\|_2 \leq \frac{2M}{c_\sigma}\xi_N$ a.s. Similar arguments hold for the upper block of D_i^{II}. This concludes the proof of (i).

(ii) Notice that

$$\dot{D}_i^I - D_i^I = \frac{1}{T}\sum_{t\in[T]}\begin{bmatrix}\mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right)\dot{v}_{t,0}^{(1)}\dot{v}_{t,0}^{(1)} - \mathfrak{f}_{it}(0)v_{t,0}^0 v_{t,0}^{0\prime} & \mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right)\dot{v}_{t,0}^{(1)}\dot{v}_{t,1}^{(1)}X_{1,it} - \mathfrak{f}_{it}(0)v_{t,0}^0 v_{t,1}^{0\prime}X_{1,it}\\ \mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right)\dot{v}_{t,1}^{(1)}\dot{v}_{t,0}^{(1)}X_{1,it} - \mathfrak{f}_{it}(0)v_{t,1}^0 v_{t,0}^{0\prime}X_{1,it} & \mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right)\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)}X_{1,it}^2 - \mathfrak{f}_{it}(0)v_{t,1}^0 v_{t,1}^{0\prime}X_{1,it}^2\end{bmatrix}.$$

To show the upper bound of $\dot{D}_i^I - D_i^I$, we take the lower block for instance and all other three blocks follow the same pattern. Noted that

$$\begin{aligned}
&\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t\in[T]}\left[\mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right)\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime}X_{1,it}^2 - \mathfrak{f}_{it}(0)v_{t,1}^0 v_{t,1}^{0\prime}X_{1,it}^2\right]\right\|_F\\
&\leq \max_{i\in I_3}\left\|\frac{1}{T}\sum_{t\in[T]}\left[\mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right) - \mathfrak{f}_{it}(0)\right]\left[\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime} - v_{t,1}^0 v_{t,1}^{0\prime}\right]X_{1,it}^2\right\|_F\\
&+ \max_{i\in I_3}\left\|\frac{1}{T}\sum_{t\in[T]}\left[\mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right) - \mathfrak{f}_{it}(0)\right]v_{t,1}^0 v_{t,1}^{0\prime}X_{1,it}^2\right\|_F\\
&+ \max_{i\in I_3}\left\|\frac{1}{T}\sum_{t\in[T]}\mathfrak{f}_{it}(0)\left[\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime} - v_{t,1}^0 v_{t,1}^{0\prime}\right]X_{1,it}^2\right\|_F\\
&= O_p(\eta_N), \qquad \text{(B.47)}
\end{aligned}$$

where the equality is by the fact that

$$\begin{aligned}
&\max_{i\in I_3,t\in[T]}\left|\mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right) - \mathfrak{f}_{it}(0)\right| \lesssim \max_{i\in I_3,t\in[T]}\left\|\Psi_{it}^0\right\|_2^2\left\|\dot{\Delta}_{t,v}\right\|_2, \quad \max_{i\in I_3}\left\|\dot{v}_{t,1}^{(1)}\right\|_2 + \left\|v_{t,1}^0\right\|_2 = O_p(1),\\
&\max_{i\in I_3}\frac{1}{T}\sum_{t\in[T]}X_{1,it}^2 \leq C, \quad \max_{t\in[T]}\left\|\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime} - v_{t,1}^0 v_{t,1}^{0\prime}\right\|_F \leq \max_{i\in I_3}\left(\left\|\dot{v}_{t,1}^{(1)}\right\|_2 + \left\|v_{t,1}^0\right\|_2\right)\left\|\dot{v}_{t,1}^{(1)} - v_{t,1}^0\right\|_2,
\end{aligned}$$

and

$$\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t\in[T]}\left[\mathfrak{f}_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right) - \mathfrak{f}_{it}(0)\right]\left[\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime} - v_{t,1}^0 v_{t,1}^{0\prime}\right]X_{1,it}^2\right\|_F$$

$$\leq \max_{i\in I_3, t\in[T]} \bar{\mathfrak{f}}' \left\| \dot{\Delta}_{t,v} \right\|_2 \left\| \dot{v}_{t,1}^{(1)} \dot{v}_{t,1}^{(1)\prime} - v_{t,1}^0 v_{t,1}^{0\prime} \right\|_F \max_{i\in I_3} \frac{1}{T} \sum_{t\in[T]} \left\| \Psi_{it}^0 \right\|_2 |X_{1,it}|^2 = O_p\left(\eta_N^2\right).$$

(iii) Note that

$$\begin{aligned}
&\dot{D}_i^{II} - D_i^{II} - \frac{1}{T}\sum_{t=1}^T \left[\mathbf{1}\{\epsilon_{it} \leq 0\} - \mathbf{1}\left\{\epsilon_{it} \leq \dot{\Delta}_{t,v}' \Psi_{it}^0 \right\} \right] \varpi_{it}^0 \\
&= \frac{1}{T}\sum_{t=1}^T \left[\mathbf{1}\{\epsilon_{it} \leq 0\} - \mathbf{1}\left\{\epsilon_{it} \leq \dot{\Delta}_{t,v}' \Psi_{it}^0 \right\} \right] \left(\dot{\varpi}_{it} - \varpi_{it}^0\right) + \frac{1}{T}\sum_{t=1}^T \left[\tau - \mathbf{1}\{\epsilon_{it} \leq 0\}\right] \left(\dot{\varpi}_{it} - \varpi_{it}^0\right) \\
&= \frac{1}{T}\sum_{t=1}^T \mathbb{E}\left\{ \left[\mathbf{1}\{\epsilon_{it} \leq 0\} - \mathbf{1}\left\{\epsilon_{it} \leq \dot{\Delta}_{t,v}' \Psi_{it}^0 \right\} \right] \left(\dot{\varpi}_{it} - \varpi_{it}^0\right) \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right\} \\
&+ \frac{1}{T}\sum_{t=1}^T \left\{ \left[\mathbf{1}\{\epsilon_{it} \leq 0\} - \mathbf{1}\left\{\epsilon_{it} \leq \dot{\Delta}_{t,v}' \Psi_{it}^0 \right\} \right] \left(\dot{\varpi}_{it} - \varpi_{it}^0\right) - \mathbb{E}\left[\left[\mathbf{1}\{\epsilon_{it} \leq 0\} - \mathbf{1}\left\{\epsilon_{it} \leq \dot{\Delta}_{t,v}' \Psi_{it}^0 \right\} \right] \left(\dot{\varpi}_{it} - \varpi_{it}^0\right) \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right] \right\} \\
&+ \frac{1}{T}\sum_{t=1}^T \left[\tau - \mathbf{1}\{\epsilon_{it} \leq 0\}\right] \left(\dot{\varpi}_{it} - \varpi_{it}^0\right) \\
&:= S_{1,i} + S_{2,i} + S_{3,i},
\end{aligned}$$

where

$$\max_{i\in I_3} \|S_{1,i}\|_2 \leq \max_{i\in I_3} \frac{1}{T}\sum_{t\in[T]} \left\|\dot{\varpi}_{it} - \varpi_{it}^0\right\|_2 \left| \mathfrak{F}_{it}\left(\dot{\Delta}_{t,v}' \Psi_{it}^0\right) - \mathfrak{F}_{it}(0) \right| \lesssim \max_{i\in I_3} \frac{1}{T} \sum_{t\in[T], j\in[p]} \left|X_{j,it}^2\right| \max_{t\in[T]} \left\| \dot{\Delta}_{t,v} \right\|_2^2 = O_p\left(\eta_N^2\right).$$

As for $S_{2,i}$ and $S_{3,i}$, we first recall, for any $e > 0$, there exists a sufficiently large constant M such that for

$$\mathscr{A}_7(M) = \left\{ \max_{i\in I_3} \left\| O_j^{(1)\prime} \dot{u}_{i,j}^{(1)} - u_{i,j}^0 \right\|_2 \leq M\eta_N, \max_{t\in[T]} \left\| O_j^{(1)\prime} \dot{v}_{t,j}^{(1)} - v_{t,j}^0 \right\|_2 \leq M\eta_N, \forall j \in [p]\cup\{0\} \right\}$$

we have $\mathbb{P}(\mathscr{A}_7^c(M)) \leq e$. In addition, let

$$\mathscr{A}_{7,i}(M) = \left\{ \left\| O_j^{(1)\prime} \dot{u}_{i,j}^{(1)} - u_{i,j}^0 \right\|_2 \leq M\eta_N, \max_{t\in[T]} \left\| O_j^{(1)\prime} \dot{v}_{t,j}^{(1)} - v_{t,j}^0 \right\|_2 \leq M\eta_N, \forall j \in [p]\cup\{0\} \right\}.$$

Then, we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_3} \|S_{2,i}\|_2 > c_{17}\eta_N^2 \right) \leq \mathbb{P}\left(\max_{i\in I_3} \|S_{2,i}\|_2 > c_{17}\eta_N^2, \mathscr{A}_7(M) \right) + e \\
&\leq \sum_{i\in I_3} \mathbb{P}\left(\|S_{2,i}\|_2 > c_{17}\eta_N^2, \mathscr{A}_7(M) \right) + e \\
&\leq \sum_{i\in I_3} \mathbb{P}\left(\|S_{2,i}\|_2 > c_{17}\eta_N^2, \mathscr{A}_{7,i}(M) \right) + e \\
&= \sum_{i\in I_3} \mathbb{E}\mathbb{P}\left(\|S_{2,i}\|_2 > c_{17}\eta_N^2 \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right) 1\{\mathscr{A}_{7,i}(m)\} + e \\
&\leq \sum_{i\in I_2} \exp\left(-\frac{c_9 c_{17}^2 T^2 \eta_N^4}{c_{18} T \eta_N^2 \xi_N^2 + c_{17} c_{18}^{1/2} T \eta_N^3 \xi_N \log T \log\log T} \right) + e = o(1) + e
\end{aligned}$$

with a positive constant c_{17} and the inequality above is by Lemma B.12(i) with the fact that, under $\mathscr{A}_{6,i}(M)$,

$$\max_{i\in I_3, t\in[T]} \left\| \left[\mathbf{1}\{\epsilon_{it} \leq 0\} - \mathbf{1}\left\{\epsilon_{it} \leq \dot{\Delta}_{t,v}' \Psi_{it}^0 \right\} \right] \left(\dot{\varpi}_{it} - \varpi_{it}^0\right) - \mathbb{E}\left[\left[\mathbf{1}\{\epsilon_{it} \leq 0\} - \mathbf{1}\left\{\epsilon_{it} \leq \dot{\Delta}_{t,v}' \Psi_{it}^0 \right\} \right] \left(\dot{\varpi}_{it} - \varpi_{it}^0\right) \middle| \mathscr{D}_{e_i}^{I_1\cup I_2} \right] \right\|_2^2$$

$\leq c_{18}\eta_N\xi_N.$

As e is arbitrary, we have $\max_{i\in I_3}\|S_{2,i}\|_2 = O_p(\eta_N^2)$. Following a similar argument, we have $\max_{i\in I_3}\|S_{3,i}\|_2 = O_p(\eta_N^2)$. By Assumption 1(ix), we note that $O_p(\eta_N^2) = o_p\left((N\vee T)^{-1/2}\right)$. ■

B.3 Lemmas for the Proof Theorem 3.3

Lemma B.19 *Define* $H_{x,j}^l = \left(\frac{\hat{L}_j' L_j^0}{N}\right)^{-1}$ *and* $H_{x,j}^w = W_j^{0\prime}\hat{W}_j\left(\hat{W}_j'\hat{W}_j\right)^{-1}$. *Under Assumptions 6-8, we have*

(i) $H_{x,j}^l = H_{x,j}^w + O_p\left(\frac{1}{N\wedge T}\right)$,

(ii) $\frac{1}{N}\left\|\hat{L}_j - L_j^0 H_{x,j}^w\right\|_F^2 = O_p\left(\frac{1}{N\wedge T}\right)$,

(iii) $\frac{1}{T}\left\|\hat{W}_j - W_j\left(H_{x,j}^{l\prime}\right)^{-1}\right\|_F^2 = O_p\left(\frac{1}{N\wedge T}\right)$.

Proof. The proof can be found in Bai and Ng (2020, Lemma 3 and Proposition 1). ■

Lemma B.20 *Under Assumptions 6-8, we have*

(i) $\frac{1}{N}L_j^{0\prime}\left(\hat{L}_j - L_j^0 H_{x,j}^w\right) = O_p\left(\frac{1}{N\wedge T}\right)$,

(ii) $\frac{1}{T}W_j^{0\prime}\left(\hat{W}_j - W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}\right) = O_p\left(\frac{1}{N\wedge T}\right)$,

(iii) $\max_{t\in[T]}\frac{1}{N}\left(\hat{L}_j - L_j^0 H_{x,j}^w\right)' e_{j,t} = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right)$,

(iv) $\max_{i\in[N]}\frac{1}{T}\left(\hat{W}_j - W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}\right)' e_{j,i} = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right)$.

Proof. Statements (i) and (ii) are the same as Lemma 4(i) and (ii) in Bai and Ng (2020). Statements (iii) and (iv) are the uniform version of Lemma 4(iii) and (iv) in Bai and Ng (2020). Below we focus on part (iv) as the proof of part (iii) follows analogously.

Noting that $\hat{W}_j' - \left(H_{x,j}^l\right)^{-1} W_j^{0\prime} = \frac{1}{N}\hat{L}_j' X_j - \frac{\hat{L}_j' L_j^0}{N} W_j^{0\prime} = \frac{1}{N}\hat{L}_j' E_j$, we have

$$\frac{1}{T}\left(\hat{W}_j - W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}\right)' e_{j,i} = \frac{1}{NT}e_{j,i}' E_j' \hat{L}_j = \frac{1}{NT}e_{j,i}' E_j' L_j^0 H_{x,j}^w + \frac{1}{NT}e_{j,i}' E_j'\left(\hat{L}_j - L_j^0 H_{x,j}^w\right).$$

For the first term on the right side,

$$\max_{i\in[N]}\frac{1}{NT}e_{j,i}' E_j' L_j^0 H_{x,j}^w \lesssim \max_{i\in[N]}\frac{1}{NT}\left\|e_{j,i}' E_j' L_j^0\right\|_2 = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right),$$

by Assumption 8(i). For the second term on the right side, we have

$$\begin{aligned}\max_{i\in[N]}\frac{1}{NT}\left\|e_{j,i}' E_j'\left(\hat{L}_j - L_j^0 H_{x,j}^w\right)\right\|_2 &\leq \max_{i\in[N]}\frac{1}{\sqrt{N}T}\left\|e_{j,i}' E_j'\right\|_2\frac{\left\|\hat{L}_j - L_j^0 H_{x,j}^w\right\|_F}{\sqrt{N}} \\ &= O_p\left(\frac{\log(N\vee T)}{\sqrt{N\wedge T}}\right) O_p\left(\frac{1}{\sqrt{N\wedge T}}\right),\end{aligned}$$

by Assumption 6(iv) and Lemma B.19(ii). Combining the above results completes the proof of part (iv). ■

Lemma B.21 *Under Assumption 1 and Assumptions 6-8, we have* $\forall j \in [p]$,

(i) $\hat{w}_{j,t} - \left(H_{x,j}^{l}\right)^{-1} w_{j,t}^{0} = H_{x,j}^{l\prime} \frac{1}{N} \sum_{i\in[N]} l_{j,i}^{0} e_{j,it} + \mathcal{R}_{w,t}$,

(ii) $\hat{l}_{j,i} - H_{x,j}^{w\prime} l_{j,i}^{0} = \left(\frac{W_j^{0\prime} W_j^0}{T}\right)^{-1} \left(H_{x,j}^{l}\right)^{-1} \frac{1}{T} \sum_{t=1}^{T} w_{j,t}^{0} e_{j,it} + \mathcal{R}_{l,i}$,

(iii) $\hat{\mu}_{j,it} - \mu_{j,it} = e_{j,it} - \hat{e}_{j,it} = w_{j,t}^{0\prime} \left(\frac{W_j^{0\prime} W_j^0}{T}\right)^{-1} \frac{1}{T} \sum_{t=1}^{T} w_{j,t}^{0} e_{j,it} + l_{j,i}^{0\prime} \frac{1}{N} \sum_{i\in[N]} l_{j,i}^{0} e_{j,it} + \mathcal{R}_{j,it}$,

such that

$$\max_{t\in[T]} |\mathcal{R}_{w,t}| = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right), \quad \max_{i\in[N]} |\mathcal{R}_{l,i}| = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right), \quad \max_{i\in[N],t\in[T],j\in[p]} |\mathcal{R}_{j,it}| = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right),$$
$$\max_{t\in[T]} \left\| \hat{w}_{j,t} - \left(H_{x,j}^{l}\right)^{-1} w_{j,t}^{0} \right\|_F = O_p(\eta_N), \quad \max_{i\in[N]} \left\| \hat{l}_{j,i} - H_{x,j}^{w\prime} l_{j,i}^{0} \right\|_F = O_p(\eta_N),$$
$$\max_{j\in[p],i\in[N],t\in[T]} |e_{j,it} - \hat{e}_{j,it}| = \max_{j\in[p],i\in[N],t\in[T]} |\hat{\mu}_{j,it} - \mu_{j,it}| = O_p(\eta_N),$$

with $\eta_N = \frac{\sqrt{\log(N\vee T)}}{\sqrt{N\wedge T}} \xi_N^2$.

Proof. Recall that $X_{j,it} = \mu_{j,it} + e_{j,it} = l_{j,i}^{0\prime} w_{j,t}^{0} + e_{j,it}$ and $X_j = L_j^0 W_j^{0\prime} + E_j$ in matrix form, where $L_j^0 \in \mathbb{R}^{N\times r_j}$ is the factor loading and $W_j^0 \in \mathbb{R}^{T\times r_j}$ is the factor matrix. Following Bai and Ng (2002), Bai (2003) and Bai and Ng (2020), if we impose the normalization restrictions that

$$\frac{L_j^{0\prime} L_j^0}{N} = I_{r_j} \quad \text{and} \quad \frac{W_j^{0\prime} W_j^0}{T} \quad \text{is a diagonal matrix with descending diagonal elements,}$$

we have the principal components estimators:

$$\hat{L}_j = X_j \hat{W}_j \left(\hat{W}_j^{\prime} \hat{W}_j\right)^{-1} \text{ and } \hat{W}_j^{\prime} = \left(\hat{L}_j^{\prime} \hat{L}_j\right)^{-1} \hat{L}_j^{\prime} X_j = \frac{1}{N} \hat{L}_j^{\prime} X_j. \tag{B.48}$$

Let $H_{x,j}^{l} = \left(\frac{1}{N} \hat{L}_j^{\prime} L_j^0\right)^{-1}$. Premultiplying $\frac{1}{N}\hat{L}_j^{\prime}$ on both sides of $X_j = L_j^0 W_j^{0\prime} + E_j$ yields

$$\frac{1}{N} \hat{L}_j^{\prime} X_j = \frac{1}{N} \hat{L}_j^{\prime} L_j^0 W_j^{0\prime} + \frac{1}{N} \hat{L}_j^{\prime} E_j.$$

It follows that

$$\begin{aligned} \hat{W}_j^{\prime} &= \left(H_{x,j}^{l}\right)^{-1} W_j^{0\prime} + \frac{1}{N} \hat{L}_j^{0\prime} E_j \\ &= \left(H_{x,j}^{l}\right)^{-1} W_j^{0\prime} + \frac{1}{N} \left(H_{x,j}^{l}\right)^{\prime} L_j^{0\prime} E_j + \frac{1}{N} \left[\hat{L} - L_j^0 H_{x,j}^{l}\right]^{\prime} E_j. \end{aligned}$$

We then show the expansion for each factor, i.e.,

$$\hat{w}_{j,t} - \left(H_{x,j}^{l}\right)^{-1} w_{j,t}^{0} = H_{x,j}^{l\prime} \frac{1}{N} \sum_{i\in[N]} l_{j,i}^{0} e_{j,it} + \frac{1}{N} \left[\hat{L}_j - L_j^0 H_{x,j}^{l}\right]^{\prime} e_{j,t}. \tag{B.49}$$

For equation (B.49), we have the uniform bound for the second term, i.e.,

$$\max_{t\in[T]} \frac{1}{N} \left[\hat{L}_j - L_j^0 H_{x,j}^{l}\right]^{\prime} e_{j,t} = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right)$$

by Lemma B.20(iii). With c_{19} being a positive constant and $\max_{i\in[N],t\in[T]} \left\| l_{j,i}^0 e_{j,it} \right\|_2 \leq c_{19}\xi_N$ a.s., we show that

$$\mathbb{P}\left(\max_{t\in[T]}\left\|\frac{1}{N}\sum_{i\in[N]} l_{j,i}^0 e_{j,it}\right\|_2 > c_{20}\sqrt{\frac{\log(N\vee T)}{N}}\xi_N\right) \leq \max_{t\in[T]} 2\exp\left(-\frac{c_{20}^2 N\xi_N^2\log(N\vee T)}{N4c_{19}^2\xi_N^2}\right) = o(1) \quad \text{(B.50)}$$

for a positive constant c_{20} by Hoeffding's inequality. It follows that $\hat{w}_{j,t}-\left(H_{x,j}^l\right)^{-1} w_{j,t}^0 = H_{x,j}^{l\prime}\frac{1}{N}\sum_{i\in[N]} l_{j,i}^0 e_{j,it}+\mathcal{R}_{w,t}$, such that

$$\max_{t\in[T]}\left\|\hat{w}_{j,t}-\left(H_{x,j}^l\right)^{-1} w_{j,t}^0\right\|_F = O_p\left(\sqrt{\frac{\log(N\vee T)}{T}}\xi_N\right), \quad \text{(B.51)}$$

and $\max_{t\in[T]}\left|\mathcal{R}_{w,t}\right| = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right)$.

Similarly, if we premultiply $\hat{W}_j\left(\hat{W}_j'\hat{W}_j\right)^{-1}$ to both sides of $X_j = L_j^0 W_j^{0\prime}+E_j$, it yields

$$X_j\hat{W}_j\left(\hat{W}_j'\hat{W}_j\right)^{-1} = L_j^0 W_j^{0\prime}\hat{W}_j\left(\hat{W}_j'\hat{W}_j\right)^{-1}+E_j\hat{W}_j\left(\hat{W}_j'\hat{W}_j\right)^{-1}.$$

It follows that

$$\begin{aligned}\hat{L}_j &= L_j^0 H_{x,j}^w + E_j\hat{W}_j\left(\hat{W}_j'\hat{W}_j\right)^{-1} \\ &= L_j^0 H_{x,j}^w + E_j W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}\left(\hat{W}_j'\hat{W}_j\right)^{-1} + E_j\left(\hat{W}_j - W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}\right)\left(\hat{W}_j'\hat{W}_j\right)^{-1},\end{aligned}$$

where the fist line is due to (B.48) and the definition that $H_{x,l} = W_j^{0\prime}\hat{W}_j\left(\hat{W}_j'\hat{W}_j\right)^{-1}$. Then we obtain the expansion for the factor loading

$$\hat{l}_{j,i} - H_{x,j}^{w\prime} l_{j,i}^0 = \left(\frac{\hat{W}_j'\hat{W}_j}{T}\right)^{-1}\left(H_{x,j}^l\right)^{-1}\frac{1}{T}\sum_{t=1}^T w_{j,t}^0 e_{j,it} + \left(\frac{\hat{W}_j'\hat{W}_j}{T}\right)^{-1}\frac{1}{T}\left[\left(\hat{W}-W\left(H_{x,j}^{l\prime}\right)^{-1}\right)' e_{j,i}\right].$$

Note that

$$\begin{aligned}\frac{\hat{W}_j'\hat{W}_j}{T} &= \frac{\left(\hat{W}_j - W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}\right)'\hat{W}_j}{T} + \frac{\left(H_{x,j}^l\right)^{-1} W_j^{0\prime}\hat{W}_j}{T} \\ &= \frac{\left(\hat{W}_j - W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}\right)'\left(\hat{W}_j - W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}\right)}{T} + \frac{\left(\hat{W}_j - W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}\right)' W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}}{T} \\ &\quad + \frac{\left(H_{x,j}^l\right)^{-1} W_j^{0\prime}\left(\hat{W}_j - W_j\left(H_{x,j}^{l\prime}\right)^{-1}\right)}{T} + \frac{\left(H_{x,j}^l\right)' W_j^{0\prime}W_j^0\left(H_{x,j}^{l\prime}\right)^{-1}}{T} \\ &= \frac{\left(H_{x,j}^l\right)' W_j^{0\prime}W_j\left(H_{x,j}^{l\prime}\right)^{-1}}{T} + O_p\left(\frac{1}{N\wedge T}\right),\end{aligned} \quad \text{(B.52)}$$

where the last equality holds by Lemma B.19(iii). Note that $\max_{i\in[N]}\frac{1}{T}\left(\hat{W}-W\left(H_{x,j}^{l\prime}\right)^{-1}\right)' e_i = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right)$ by Lemma B.20(iv), and we can show that

$$\max_{i\in[N]}\left\|\frac{1}{T}\sum_{t=1}^T w_{j,t}^0 e_{j,it}\right\|_2 = O_p\left(\sqrt{\frac{\log(N\vee T)}{T}}\xi_N\right) \quad \text{(B.53)}$$

as in (B.50) by conditional Bernstein's inequality in Lemma B.12(i) given $\left\{W_j^0\right\}_{j\in[p]}$. It follows that $\hat{l}_{j,i} - H_{x,j}^{w\prime} l_{j,i}^0 = \left(\frac{W_j^{0\prime}W_j^0}{T}\right)^{-1}\left(H_{x,j}^l\right)^{-1}\frac{1}{T}\sum_{t=1}^T w_{j,t}^0 e_{j,it} + \mathcal{R}_{l,i}$ such that

$$\max_{i\in[N]}\left\|\hat{l}_{j,i} - H_{x,j}^{w\prime} l_{j,i}^0\right\|_F = O_p(\eta_N) \text{ and } \max_{i\in[N]}|\mathcal{R}_{l,i}| = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right).$$

Then, it's natural to obtain that

$$\begin{aligned}
\hat{\mu}_{j,it} - \mu_{j,it} &= \hat{l}_{j,i}'\hat{w}_t - l_{j,i}^{0\prime}w_{j,t}^0 \\
&= \left(\hat{l}_{j,i} - \left(H_{x,j}^l\right)' l_{j,i}^0\right)'\left(\hat{w}_{j,t} - \left(H_{x,j}^l\right)^{-1} w_{j,t}^0\right) + \left(\hat{l}_{j,i} - \left(H_{x,j}^l\right)' l_{j,i}^0\right)'\left(H_{x,j}^l\right)^{-1} w_{j,t}^0 \\
&+ l_{j,i}^{0\prime}H_{x,j}^l\left(\hat{w}_{j,t} - \left(H_{x,j}^l\right)^{-1} w_{j,t}^0\right) \\
&= \left(\hat{l}_{j,i} - \left(H_{x,j}^l\right)' l_{j,i}^0\right)'\left(H_{x,j}^l\right)^{-1} w_{j,t}^0 + l_{j,i}^{0\prime}H_{x,j}^l\left(\hat{w}_{j,t} - \left(H_{x,j}^l\right)^{-1} w_{j,t}^0\right) + O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right) \\
&= w_{j,t}^0\left(\frac{W_j^{0\prime}W_j^0}{T}\right)^{-1}\frac{1}{T}\sum_{t=1}^T w_{j,t}^0 e_{j,it} + l_{j,i}^{0\prime}\frac{1}{N}\sum_{i\in[N]} l_{j,i}^0 e_{j,it} + \mathcal{R}_{j,it}
\end{aligned}$$

where the second equality holds by statements (i) and (ii), and the last equality holds with $\max_{i\in[N],t\in[T]}|\mathcal{R}_{j,it}| = O_p\left(\frac{\log(N\vee T)}{N\wedge T}\right)$. By Assumption 1(iv), $\max_{t\in[T]}\left\|w_{j,t}^0\right\|_2 \leq M$ a.s. and $\max_{i\in[N]}\left\|l_{j,i}^0\right\|_2 \leq M$ a.s., which leads to $\max_{j\in[p],i\in[N],t\in[T]}|\hat{\mu}_{j,it} - \mu_{j,it}| = O_p\left(\sqrt{\frac{\log(N\vee T)}{N\wedge T}}\xi_N\right) = O_p(\eta_N)$. ■

Lemma B.22 *Under Assumptions 1-9, for matrices* $\dot{D}_i^F$ *and* D_i^F *defined in the proof of Theorem 3.3, we have* $\max_{i\in I_3}\left\|\dot{D}_i^F - D_i^F\right\|_F = O_p(\eta_N)$ *with* $\eta_N = \frac{\sqrt{\log(N\vee T)}\xi_N^2}{\sqrt{N\wedge T}}$.

Proof. Recall that

$$\begin{aligned}
\dot{D}_i^F &= \frac{1}{T}\sum_{t=1}^T f_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]\begin{bmatrix}\dot{v}_{t,0}^{(1)}\dot{v}_{t,0}^{(1)\prime} & \hat{e}_{1,it}\dot{v}_{t,0}^{(1)}\dot{v}_{t,1}^{(1)\prime} \\ \hat{e}_{1,it}\dot{v}_{t,1}^{(1)}\dot{v}_{t,0}^{(1)\prime} & \hat{e}_{1,it}^2\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime}\end{bmatrix} \text{ and} \\
D_i^F &= \frac{1}{T}\sum_{t=1}^T f_{it}(0)\begin{bmatrix}O_0^{(1)}v_{t,0}^0 v_{t,0}^{0\prime}O_0^{(1)\prime} & 0 \\ 0 & e_{1,it}^2 O_1^{(1)}v_{t,1}^0 v_{t,1}^{0\prime}O_1^{(1)\prime}\end{bmatrix},
\end{aligned}$$

with $\max_{i\in I_2}\left\|D_i^F\right\|_F = O(1)$ a.s..

Let i_{it} denote $\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)$ for short. We have

$$\begin{aligned}
&\dot{D}_i^F - D_i^F \\
&= \frac{1}{T}\sum_{t=1}^T f_{it}(0)\begin{bmatrix}\dot{v}_{t,0}^{(1)}\dot{v}_{t,0}^{(1)\prime} - O_0^{(1)}v_{t,0}^0 v_{t,0}^{0\prime}O_0^{(1)\prime} & \dot{v}_{t,0}^{(1)}\dot{v}_{t,1}^{(1)\prime}\hat{e}_{1,it} - O_0^{(1)}v_{t,0}^0 v_{t,1}^{0\prime}O_1^{(1)\prime}e_{1,it} \\ \dot{v}_{t,1}^{(1)}\dot{v}_{t,0}^{(1)\prime}\hat{e}_{1,it} - O_1^{(1)}v_{t,1}^0 v_{t,0}^{0\prime}O_0^{(1)\prime}e_{1,it} & \hat{e}_{1,it}^2\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime} - e_{1,it}^2 O_1^{(1)}v_{t,1}^0 v_{t,1}^{0\prime}O_1^{(1)\prime}\end{bmatrix} \\
&+ \frac{1}{T}\sum_{t=1}^T f_{it}(0)\begin{bmatrix}0 & O_0^{(1)}v_{t,0}^0 v_{t,1}^{0\prime}O_1^{(1)\prime}e_{1,it} \\ O_1^{(1)}v_{t,1}^0 v_{t,0}^{0\prime}O_0^{(1)\prime}e_{1,it} & 0\end{bmatrix} \\
&+ \frac{1}{T}\sum_{t=1}^T\left(f_{it}\left(i_{it}\right) - f_{it}(0)\right)\begin{bmatrix}\dot{v}_{t,0}^{(1)}\dot{v}_{t,0}^{(1)\prime} & \dot{v}_{t,0}^{(1)}\dot{v}_{t,1}^{(1)\prime}\hat{e}_{1,it} \\ \dot{v}_{t,1}^{(1)}\dot{v}_{t,0}^{(1)\prime}\hat{e}_{1,it} & \hat{e}_{1,it}^2\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime}\end{bmatrix} \\
&:= F_{1,i} + F_{2,i} + F_{3,i}.
\end{aligned}$$

We define $F_{1,i}^m$ for $m = \{1,2,3,4\}$ as fourth clockwise blocks in $F_{1,i} := \begin{bmatrix} F_{1,i}^1 & F_{1,i}^2 \\ F_{1,i}^4 & F_{1,i}^3 \end{bmatrix}$. Define $F_{2,i}^m$ and $F_{3,i}^m$ similarly. We aim to show the uniform bound for each block.

First, we observe that

$$
\begin{aligned}
F_{1,i}^1 &= \frac{1}{T}\sum_{t=1}^T f_{it}(0)\left(\dot{v}_{t,0}^{(1)}\dot{v}_{t,0}^{(1)\prime} - O_0^{(1)}v_{t,0}^0 v_{t,0}^{0\prime} O_0^{(1)\prime}\right) \\
&= \frac{1}{T}\sum_{t=1}^T f_{it}(0)\left[\left(\dot{v}_{t,0}^{(1)} - O_0^{(1)}v_{t,0}^0\right)\left(\dot{v}_{t,0}^{(1)} - O_0^{(1)}v_{t,0}^0\right)' + O_0^{(1)}v_{t,0}^0\left(\dot{v}_{t,0}^{(1)} - O_0^{(1)}v_{t,0}^0\right)' + \left(\dot{v}_{t,0}^{(1)} - O_0^{(1)}v_{t,0}^0\right)\left(O_0^{(1)}v_{t,0}^0\right)'\right] \\
&= O_p\left(\eta_N^2\right) + O_p(\eta_N) = O_p(\eta_N) \qquad \text{(B.54)}
\end{aligned}
$$

uniformly over $i \in I_3$. For $F_{1,i}^2$, we have by Theorem 3.2 and Lemma B.21,

$$
\begin{aligned}
F_{1,i}^2 &= \frac{1}{T}\sum_{t=1}^T f_{it}(0)\left(\dot{v}_{t,0}^{(1)}\dot{v}_{t,1}^{(1)\prime}\hat{e}_{1,it} - O_0^{(1)}v_{t,0}^0 v_{t,1}^{0\prime} O_1^{(1)\prime} e_{1,it}\right) \\
&= \frac{1}{T}\sum_{t=1}^T f_{it}(0)\left\{\dot{v}_{t,0}^{(1)}\dot{v}_{t,1}^{(1)\prime}\left(\hat{e}_{1,it} - e_{1,it}\right) + \left(\dot{v}_{t,0}^{(1)}\dot{v}_{t,1}^{(1)\prime} - O_0^{(1)}v_{t,0}^0 v_{t,1}^{0\prime} O_1^{(1)\prime}\right)e_{1,it}\right\} \\
&= O_p(\eta_N) \text{ uniformly in } i \in I_3. \qquad \text{(B.55)}
\end{aligned}
$$

The same order holds for $\max_{i\in I_3}\left\|F_{1,i}^3\right\|_F$ as $F_{1,i}^3 = F_{1,i}^{2\prime}$. Next, we study $F_{1,i}^4$. Noting that

$$
|\hat{e}_{1,it}^2 - e_{1,it}^2| = |\left(\hat{e}_{1,it} + e_{1,it}\right)\left(\hat{e}_{1,it} - e_{1,it}\right)| \le \left(2|e_{1,it}| + \max_{i\in I_3, t\in[T]}|\hat{e}_{1,it} - e_{1,it}|\right)\left(\max_{i\in I_3, t\in[T]}|\hat{e}_{1,it} - e_{1,it}|\right),
$$

we have

$$
\begin{aligned}
\max_{i\in I_3}||F_{1,i}^4||_F &= \left\|\frac{1}{T}\sum_{t=1}^T f_{it}(0)\left(\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime}\hat{e}_{1,it}^2 - O_1^{(1)}v_{t,1}^0 v_{t,1}^{0\prime}O_1^{(1)\prime}e_{1,it}^2\right)\right\|_F \\
&\le \max_{i\in I_3}\frac{1}{T}\sum_{t=1}^T f_{it}(0)\left\|\dot{v}_{t,1}^{(1)}\right\|_2^2\left(|\hat{e}_{1,it}^2 - e_{1,it}^2|\right) + \max_{i\in I_3}\frac{1}{T}\sum_{t=1}^T f_{it}(0)\left\|\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime} - O_1^{(1)}v_{t,1}^0 v_{t,1}^{0\prime}O_1^{(1)\prime}\right\|_F e_{1,it}^2 \\
&= \max_{t\in[T]}\left\|\dot{v}_{t,1}^{(1)}\right\|_2^2\left(\max_{i\in I_3, t\in[T]}|\hat{e}_{1,it} - e_{1,it}|\right)\max_{i\in I_3}\left(\frac{2}{T}\sum_{t=1}^T f_{it}(0)|e_{1,it}|\right) \\
&+ \max_{t\in[T]}\left\|\dot{v}_{t,1}^{(1)}\right\|_2^2\left(\max_{i\in I_3, t\in[T]}|\hat{e}_{1,it} - e_{1,it}|^2\right)\max_{i\in I_3}\left(\frac{2}{T}\sum_{t=1}^T f_{it}(0)\right) \\
&+ \max_{t\in[T]}\left\|\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime} - O_1^{(1)}v_{t,1}^0 v_{t,1}^{0\prime}O_1^{(1)\prime}\right\|_F \max_{i\in I_3}\left(\frac{1}{T}\sum_{t=1}^T f_{it}(0)e_{1,it}^2\right) = O_p(\eta_N). \qquad \text{(B.56)}
\end{aligned}
$$

Combining (B.54)-(B.56), we conclude $\max_{i\in I_3}\left\|F_{1,i}\right\|_2 = O_p(\eta_N)$.

For $F_{2,i}^2$, we note that $\mathbb{E}\left(f_{it}(0)v_{t,0}^0 v_{t,1}^{0\prime}e_{1,it}\middle|\mathscr{D}\right) = 0$ by Assumption 9 and $\max_{i\in I_3, t\in[T]}\left\|f_{it}(0)v_{t,0}^0 v_{t,1}^{0\prime}e_{1,it}\right\|_F \le c_{22}\xi_N$ *a.s.* by Assumption 1(iv) and Lemma B.13(i). Then, by conditional Bernstein's inequality in Lemma B.12(i) and Assumption 9(iii), we can show that, with positive constants c_{21} and c_{22},

$$
\mathbb{P}\left(\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t\in[T]} f_{it}(0)v_{t,0}^0 v_{t,1}^{0\prime}e_{1,it}\right\|_F > c_{21}\sqrt{\frac{\log(N\vee T)}{T}}\xi_N\middle|\mathscr{D}\right)
$$

$$\leq \sum_{i\in I_3} \exp\left(-\frac{c_{12}c_{21}^2 T\xi_N^2 \log(N\vee T)}{c_{22}^2 T\xi_N^2 + c_{21}c_{22}\sqrt{T\log(N\vee T)}\xi_N^2 \log T \log\log T}\right) = o(1), \tag{B.57}$$

which yields $\max_{i\in I_3} \|F_{2,i}\|_F = O_p(\eta_N)$.

As for $F_{3,i}$, we show the bound for the first block and all other three blocks follow the same argument. Recall that

$$\begin{aligned}
\imath_{it} &= u_{i,0}^{0\prime} O_0^{(1)} \dot{v}_{t,0}^{(1)} - u_{i,0}^{0\prime} v_{t,0}^0 + \hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)} - \mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0 + \hat{e}_{1,it}u_{i,1}^{0\prime}O_1^{(1)}\dot{v}_{t,1}^{(1)} - e_{1,it}u_{i,1}^{0\prime}v_{t,1}^0 \\
&= \left[\left(O_0^{(1)}u_{i,0}^0\right)'\left(\dot{v}_{t,0}^{(1)} - O_0^{(1)}v_{t,0}^0\right)\right] + \left(\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)} - \mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right) + \left(\hat{e}_{1,it}u_{i,1}^{0\prime}O_1^{(1)}\dot{v}_{t,1}^{(1)} - e_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right) \\
&:= \imath_{it}^I + \imath_{it}^{II} + \imath_{it}^{III},
\end{aligned} \tag{B.58}$$

with the fact that $\left|\imath_{it}^I\right| \leq \left\|O_0^{(1)}u_{i,0}^0\right\|_2 \left\|\dot{v}_{t,0}^{(1)} - O_0^{(1)}v_{t,0}^0\right\|_2 := R_{1,it}^I$ such that $\max_{i\in I_3, t\in[T]} R_{1,it}^I = O_p(\eta_N)$. In addition, we have

$$\begin{aligned}
\max_{i\in I_3, t\in[T]} \left|\imath_{it}^{II}\right| &= \max_{i\in I_3, t\in[T]} \left|\left(\hat{\mu}_{1,it} - \mu_{1,it}\right)\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)} + \mu_{1,it}\left(\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)} - u_{i,1}^{0\prime}v_{t,1}^0\right)\right| \\
&\leq R_{1,it}^{II} + R_{2,it}^{II}|\mu_{1,it}|,
\end{aligned} \tag{B.59}$$

where $\max_{i\in I_3, t\in[T]} |R_{1,it}^{II}| = O_p(\eta_N)$ and $\max_{i\in I_3, t\in[T]} |R_{2,it}^{II}| = O_p(\eta_N)$. Similarly, we have

$$\begin{aligned}
\left|\imath_{it}^{III}\right| &= \left|\hat{e}_{1,it}u_{i,1}^{0\prime}O_1^{(1)}\dot{v}_{t,1}^{(1)} - e_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right| \\
&\leq \left|\left(\hat{e}_{1,it} - e_{1,it}\right)\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right) + \left(\hat{e}_{1,it} - e_{1,it}\right)\left(O_1^{(1)}u_{i,1}^0\right)' O_1^{(1)}v_{t,1}^0\right| \\
&\quad + \left|e_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right)\right| \\
&= R_{1,it}^{III} + R_{2,it}^{III}|e_{1,it}|,
\end{aligned} \tag{B.60}$$

where $\max_{i\in I_3, t\in[T]} |R_{1,it}^{III}| = O_p(\eta_N)$ and $\max_{i\in I_3, t\in[T]} |R_{2,it}^{III}| = O_p(\eta_N)$. Therefore, we have

$$\begin{aligned}
\max_{i\in I_3} ||F_{3,i}^1||_F &\leq \frac{1}{T}\sum_{t=1}^T \left|f_{it}\left(\imath_{it}\right) - f_{it}(0)\right| \left\|\dot{v}_{t,0}^{(1)}\right\|_2^2 \\
&\lesssim \max_{t\in[T]} \left\|\dot{v}_{t,0}^{(1)}\right\|_2 \max_{i\in I_3, t\in[T]} \left(R_{1,it}^I + R_{1,it}^{II} + R_{2,it}^{II} + R_{1,it}^{III} + R_{2,it}^{III}\right)\frac{1}{T}\sum_{t\in[T]}\left(1 + |e_{1,it}| + |\mu_{1,it}|\right) \\
&= O_p(\eta_N),
\end{aligned} \tag{B.61}$$

Combining all results above, we obtain that $\max_{i\in I_3} ||\hat{D}_i^F - D_i^F||_F = O_p(\eta_N)$. ■

Lemma B.23 *Under Assumptions 1-9, for matrices $\dot{D}_i^J$ and D_i^J defined in the proof of Theorem 3.3, we have*

$$\max_{i\in I_3} \left\|\dot{D}_i^J - D_i^J\right\|_F = \left\|\begin{bmatrix} O_p\left(\eta_N^2\right) & O_p(\eta_N) \\ O_p\left(\eta_N^2\right) & O_p\left(\eta_N^2\right)\end{bmatrix}\right\|_F.$$

Proof. Recall that

$$D_i^J = \frac{1}{T}\sum_{t=1}^T f_{it}(0)\begin{bmatrix} \dot{v}_{t,0}^{(1)}\left(O_0^{(1)}v_{t,0}^0 - \dot{v}_{t,0}^{(1)}\right)' & 0 \\ 0 & e_{1,it}^2 O_1^{(1)}v_{t,1}^0\left(O_1^{(1)}v_{t,1}^0 - \dot{v}_{t,1}^{(1)}\right)'\end{bmatrix}.$$

We have

$$
\begin{aligned}
&\hat{D}_i^J - D_i^J \\
&= \frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\begin{bmatrix} 0 & \dot{v}_{t,0}^{(1)} v_{t,1}^{0\prime} O_1^{(1)\prime}\left(e_{1,it}-\hat{e}_{1,it}\right) + \dot{v}_{t,0}^{(1)}\left(O_1^{(1)} v_{t,1}^0 - \dot{v}_{t,1}^{(1)}\right)^{\prime}\left(\hat{e}_{1,it}-e_{1,it}\right) \\ \left(\hat{e}_{1,it}-e_{1,it}\right)\dot{v}_{t,1}^{(1)}\left(O_0^{(1)} v_{t,0}^0 - \dot{v}_{t,0}^{(1)}\right)^{\prime} & \dot{v}_{t,1}^{(1)} v_{t,1}^{0\prime} O_1^{(1)\prime} e_{1,it}\left(\hat{e}_{1,it}-e_{1,it}\right) - \dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime}\left(\hat{e}_{1,it}^2 - e_{1,it}^2\right)\end{bmatrix} \\
&+ \frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\begin{bmatrix} 0 & \dot{v}_{t,0}^{(1)}\left(O_1^{(1)} v_{t,1}^0 - \dot{v}_{t,1}^{(1)}\right)^{\prime} e_{1,it} \\ \dot{v}_{t,1}^{(1)}\left(O_0^{(1)} v_{t,0}^0 - \dot{v}_{t,0}^{(1)}\right)^{\prime} e_{1,it} & \left(\dot{v}_{t,1}^{(1)} - O_1^{(1)} v_{t,1}^0\right)\left(O_1^{(1)} v_{t,1}^0 - \dot{v}_{t,1}^{(1)}\right)^{\prime} e_{1,it}^2\end{bmatrix} \\
&+ \frac{1}{T}\sum_{t=1}^{T}\left[f_{it}(\tilde{\iota}_{it}) - f_{it}(0)\right]\begin{bmatrix} \dot{v}_{t,0}^{(1)}\left(O_0^{(1)} v_{t,0}^0 - \dot{v}_{t,0}^{(1)}\right)^{\prime} & \dot{v}_{t,0}^{(1)}\left(e_{1,it} O_1^{(1)} v_{t,1}^0 - \hat{e}_{1,it}\dot{v}_{t,1}^{(1)}\right)^{\prime} \\ \hat{e}_{1,it}\dot{v}_{t,1}^{(1)}\left(O_0^{(1)} v_{t,0}^0 - \dot{v}_{t,0}^{(1)}\right)^{\prime} & \hat{e}_{1,it}\dot{v}_{t,1}^{(1)}\left(e_{1,it} O_1^{(1)} v_{t,1}^0 - \hat{e}_{1,it}\dot{v}_{t,1}^{(1)}\right)^{\prime}\end{bmatrix} \\
&:= J_{1,i} + J_{2,i} + J_{3,i}.
\end{aligned}
$$

As in the proof of the last lemma, we define $J_{1,i}^m$, $J_{2,i}^m$ and $J_{3,i}^m$ for $m = 1,2,3,4$ as four clockwise blocks in $J_{1,i}$, $J_{2,i}$ and $J_{3,i}$, respectively.

First, we study $J_{1,i}$. For $J_{1,i}^2$, we notice that

$$
\begin{aligned}
J_{1,i}^2 &= \frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\dot{v}_{t,0}^{(1)} v_{t,1}^{0\prime} O_1^{(1)\prime}\left(e_{1,it}-\hat{e}_{1,it}\right) + \frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\dot{v}_{t,0}^{(1)}\left(O_1^{(1)} v_{t,1}^0 - \dot{v}_{t,1}^{(1)}\right)^{\prime}\left(\hat{e}_{1,it}-e_{1,it}\right) \\
&= \frac{1}{T}\sum_{t=1}^{T} f_{it}(0) O_0^{(1)} v_{t,0}^0 v_{t,1}^{0\prime} O_1^{(1)\prime}\left(e_{1,it}-\hat{e}_{1,it}\right) \\
&+ \frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\left(\dot{v}_{t,0}^{(1)} - O_0^{(1)} v_{t,0}^0\right) v_{t,1}^{0\prime} O_1^{(1)\prime}\left(e_{1,it}-\hat{e}_{1,it}\right) + O_p\left(\eta_N^2\right) \\
&= \frac{1}{T}\sum_{t=1}^{T} f_{it}(0) O_0^{(1)} v_{t,0}^0 v_{t,1}^{0\prime} O_1^{(1)\prime}\left(e_{1,it}-\hat{e}_{1,it}\right) + O_p\left(\eta_N^2\right) + O_p\left(\eta_N^2\right) \\
&= O_p(\eta_N) \quad \text{uniformly over } i \in I_3.
\end{aligned} \tag{B.62}
$$

Noted that the leading term in $J_{1,i}^2$ is $\frac{1}{T}\sum_{t=1}^{T} f_{it}(0) O_0^{(1)} v_{t,0}^0 v_{t,1}^{0\prime} O_1^{(1)\prime}\left(e_{1,it}-\hat{e}_{1,it}\right)$, which will remain as the bias term of $\hat{u}_{i,0}^{(3,1)}$.

Furthermore, it is clear that

$$
J_{1,i}^3 = \frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\left(\hat{e}_{1,it}-e_{1,it}\right)\dot{v}_{t,1}^{(1)}\left(O_0^{(1)} v_{t,0}^0 - \dot{v}_{t,0}^{(1)}\right)^{\prime} = O_p\left(\eta_N^2\right) \quad \text{uniformly over } i \in I_3.
$$

Next, we obtain that

$$
\begin{aligned}
J_{1,i}^4 &= \frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\dot{v}_{t,1}^{(1)} v_{t,1}^{0\prime} O_1^{(1)\prime} e_{1,it}\left(\hat{e}_{1,it}-e_{1,it}\right) - \frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime}\left(\hat{e}_{1,it}^2 - e_{1,it}^2\right) \\
&= O_1^{(1)}\left[\frac{1}{T}\sum_{t=1}^{T} f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} e_{1,it}\left(\hat{e}_{1,it}-e_{1,it}\right)\right] O_1^{(1)\prime} - \frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime}\left(\hat{e}_{1,it}-e_{1,it}\right)^2 \\
&- \frac{2}{T}\sum_{t=1}^{T} f_{it}(0)\dot{v}_{t,1}^{(1)}\dot{v}_{t,1}^{(1)\prime} e_{1,it}\left(\hat{e}_{1,it}-e_{1,it}\right) + O_p(\eta_N^2)
\end{aligned}
$$

$$= -O_1^{(1)} \left[\frac{1}{T} \sum_{t=1}^{T} f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} e_{1,it} \left(\hat{e}_{1,it} - e_{1,it} \right) \right] O_1^{(1)\prime} + O_p(\eta_N^2), \tag{B.63}$$

where the first and second equalities hold by Theorem 3.2(ii) and Lemma B.21. We deal with the first term in the second equality of (B.63) by inserting the linear expansion of $\hat{e}_{1,it} - e_{1,it}$ in Lemma B.21(iii), i.e.,

$$\begin{aligned}
&\frac{1}{T} \sum_{t=1}^{T} f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} e_{1,it} \left(\hat{e}_{1,it} - e_{1,it} \right) \\
&= -\frac{1}{T} \sum_{t=1}^{T} \left\{ f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} e_{1,it} \left[w_{1,t}^{0\prime} \left(\frac{W_1^{0\prime} W_1^0}{T} \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} w_{1,t}^0 e_{1,it} \right] \right\} \\
&\quad - \frac{1}{T} \sum_{t=1}^{T} \left\{ f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} \left[e_{1,it} l_{1,i}^{0\prime} \frac{1}{N} \sum_{i^* \in [N]} l_{1,i^*}^0 e_{1,i^*t} \right] \right\} + O_p\left(\frac{\log(N \vee T)}{N \wedge T} \right)
\end{aligned} \tag{B.64}$$

uniformly over $i \in I_3$. For the first term in the right side of (B.64), we notice that

$$\begin{aligned}
&\max_{i \in I_3} \left\| \frac{1}{T} \sum_{t=1}^{T} \left\{ f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} e_{1,it} \left[w_{1,t}^{0\prime} \left(\frac{W_1^{0\prime} W_1^0}{T} \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} w_{1,t}^0 e_{1,it} \right] \right\} \right\|_F \\
&= \max_{i \in I_3} \left\| \frac{1}{T} \sum_{t=1}^{T} \left\{ f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} e_{1,it} \left[\sum_{k \in [r_1]} w_{1,t,k}^0 \left[\left(\frac{W_1^{0\prime} W_1^0}{T} \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} w_{1,t}^0 e_{1,it} \right]_k \right] \right\} \right\|_F \\
&\leq \sum_{k \in [r_1]} \max_{i \in I_3} \left\| \left[\frac{1}{T} \sum_{t=1}^{T} f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} w_{1,t,k}^0 e_{1,it} \right] \right\|_F \max_{k \in [r_1], i \in I_3} \left| \left[\left(\frac{W_1^{0\prime} W_1^0}{T} \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} w_{1,t}^0 e_{1,it} \right]_k \right| \\
&= \sum_{k \in [r_1]} \max_{i \in I_3} \left\| \left[\frac{1}{T} \sum_{t=1}^{T} f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} w_{1,t,k}^0 e_{1,it} \right] \right\|_F O_p(\eta_N) \\
&= O_p(\eta_N^2),
\end{aligned}$$

where $w_{1,t}$ is the fact of $X_{1,it}$, r_1 is the dimension of $w_{1,t}$, and $w_{1,t,k}$ is the kth element of $w_{1,it}$, the second equality holds by the results in (B.52) and (B.53), and the last equality is by the fact that

$$\mathbb{P}\left(\max_{i \in I_3} \left\| \left[\frac{1}{T} \sum_{t=1}^{T} f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} w_{1,t,k}^0 e_{1,it} \right] \right\|_F > c_{23} \sqrt{\frac{\log(N \vee T)}{T}} \xi_N \middle| \mathscr{D} \right) = o(1)$$

following analogous analysis as (B.57) with a positive constant c_{23}.

For the second term on the RHS of equation (B.64), we notice that

$$\begin{aligned}
&\frac{1}{T} \sum_{t=1}^{T} \left\{ f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} \left[e_{1,it} l_{1,i}^{0\prime} \frac{1}{N} \sum_{i^* \in N} l_{1,i^*}^0 e_{1,i^*t} \right] \right\} \\
&= \frac{1}{NT} \sum_{i^* \in [N]} \sum_{t \in [T]} f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} l_{1,i}^{0\prime} l_{1,i^*}^0 e_{1,it} e_{1,i^*t} \\
&= \frac{1}{NT} \sum_{t \in [T]} f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} l_{1,i}^{0\prime} l_{1,i}^0 e_{1,it}^2 + \frac{1}{NT} \sum_{i^* \neq i} \sum_{t \in [T]} f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} l_{1,i}^{0\prime} l_{1,i^*}^0 e_{1,it} e_{1,i^*t}
\end{aligned} \tag{B.65}$$

with

$$\max_{i \in I_3} \left\| \frac{1}{NT} \sum_{t \in [T]} f_{it}(0) v_{t,1}^0 v_{t,1}^{0\prime} l_{1,i}^{0\prime} l_{1,i}^0 e_{1,it}^2 \right\| \leq \frac{1}{N} \max_{i \in I_3} \left\| l_{1,i}^0 \right\|_2^2 \max_{t \in [T]} \left(\left\| v_{t,1}^0 \right\|_2 \left\| v_{t,1}^0 \right\|_2 \right) \frac{1}{T} \sum_{t \in [T]} e_{1,it}^2$$

$$= O_p\left(\frac{\xi_N^2}{N}\right),$$

where the equality holds by Assumption 1(iv), Lemma B.13(i) and Theorem 3.2(ii). For the second term in (B.65), by Assumption 1(iii), $e_{j,it}$ is strong mixing across t and independent given fixed effect. We define $E_i^{vec} = (\tilde{e}_{1,1}', \cdots, \tilde{e}_{1,i-1}', \tilde{e}_{1,i+1}', \cdots, \tilde{e}_{1,N}')'$ with $\tilde{e}_{1,i^*} = (e_{1,i^*1}e_{1,i1}, \cdots, e_{1,i^*T}e_{1,iT})$ for $i^* \neq i$. We can see E_i^{vec} will also be a strong mixing sequence, conditional on $\mathscr{D}$, which implies

$$\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_3}\left\|\frac{1}{NT}\sum_{i^*\neq i}\sum_{t\in[T]} f_{it}(0)\dot{v}_{t,1}^{(1)}v_{t,1}^{0\prime}O_1^{(1)\prime}l_{1,i}^{0\prime}l_{1,i^*}^0 e_{1,it}e_{1,i^*t}\right\|_F > c_{23}\eta_N^2\right)\\
&\leq \sum_{i\in I_3}\mathbb{E}\mathbb{P}\left(\left\|\frac{1}{NT}\sum_{i^*\neq i}\sum_{t\in[T]} f_{it}(0)\dot{v}_{t,1}^{(1)}v_{t,1}^{0\prime}O_1^{(1)\prime}l_{1,i}^{0\prime}l_{1,i^*}^0 e_{1,it}e_{1,i^*t}\right\|_F > c_{23}\eta_N^2\middle|\mathscr{D}\right) = o(1),
\end{aligned}$$

where the equality holds by Lemma B.12(i). This implies

$$\begin{aligned}
\max_{i\in I_3}\left\|O_1^{(1)}\left[\frac{1}{T}\sum_{t=1}^T f_{it}(0)v_{t,1}^0 v_{t,1}^{0\prime}e_{1,it}\left(\hat{e}_{1,it}-e_{1,it}\right)\right]O_1^{(1)\prime}\right\|_F &= \max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T f_{it}(0)v_{t,1}^0 v_{t,1}^{0\prime}e_{1,it}\left(\hat{e}_{1,it}-e_{1,it}\right)\right\|_F\\
&= O_p(\eta_N^2).
\end{aligned}$$

Therefore, we conclude that $\max_{i\in I_3}\left\|J_{1,i}^4\right\|_F = O_p\left(\eta_N^2\right)$. Then

$$\max_{i\in I_3}\|J_{1,i}\|_F = \left\|\begin{bmatrix} 0 & O_p(\eta_N) \\ O_p\left(\eta_N^2\right) & O_p\left(\eta_N^2\right)\end{bmatrix}\right\|_F.$$

Next, for $J_{2,i}^2$ and $J_{2,i}^3$, conditioning on $\mathscr{D}^{I_1\cup I_2}$ and following Bernstein's inequality in Lemma B.12(i), we have

$$\begin{aligned}
\max_{i\in I_3}\left\|J_{2,i}^2\right\|_F &= \max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T f_{it}(0)\dot{v}_{t,0}^{(1)}\left(O_1^{(1)}v_{t,1}^0 - \dot{v}_{t,1}^{(1)}\right)' e_{1,it}\right\|_F = O_p\left(\eta_N^2\right),\\
\max_{i\in I_3}\left\|J_{2,i}^3\right\|_F &= \max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T f_{it}(0)\dot{v}_{t,1}^{(1)}\left(O_0^{(1)}v_{t,0}^0 - \dot{v}_{t,0}^{(1)}\right)' e_{1,it}\right\|_F = O_p\left(\eta_N^2\right),
\end{aligned}$$

which can be obtained by the similar arguments as Lemma B.24(i). These results, in conjunction with the fact that

$$\begin{aligned}
\max_{i\in I_3}\left\|J_{2,i}^4\right\|_F &= \max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T f_{it}(0)\left(\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right)\left(O_1^{(1)}v_{t,1}^0 - \dot{v}_{t,1}^{(1)}\right)' e_{1,it}^2\right\|_F\\
&\leq \max_{i\in I_3, t\in[T]}\left(\frac{1}{T}\sum_{t=1}^T f_{it}(0)e_{1,it}^2\right)\max_{t\in[T]}\left\|\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right\|_2^2 = O_p\left(\eta_N^2\right),
\end{aligned}$$

imply that $\max_{i\in I_3}\|J_{2,i}\|_F = O_p(\eta_N^2)$.

For $J_{3,i}$, we have

$$\max_{i\in I_3}\left\|J_{3,i}^1\right\|_F = \max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T \left[f_{it}\left(\tilde{\iota}_{it}\right) - f_{it}(0)\right]\dot{v}_{t,0}^{(1)}\left(O_0'v_{t,0}^0 - \dot{v}_{t,0}^{(1)}\right)'\right\|_F$$

$$
\begin{aligned}
&\leq \max_{t\in[T]}\left\|\dot{v}_{t,0}^{(1)}\right\|_2\left\|\dot{v}_{t,0}^{(1)}\left(O_0'v_{t,0}^0-\dot{v}_{t,0}^{(1)}\right)'\right\|_2\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left|f_{it}\left(\tilde{\iota}_{it}\right)-f_{it}(0)\right|\\
&\lesssim O_p(\eta_N)\max_{i\in I_3}\frac{1}{T}\sum_{t\in[T]}\left|\tilde{\iota}_{it}\right|\\
&=O_p\left(\eta_N^2\right), \qquad\qquad \text{(B.66)}
\end{aligned}
$$

where the third line is by Lipschitz continuity of the density function, Theorem 3.2(ii), Assumption 2, and the fact that $|\tilde{\iota}_{it}|$ lies between 0 and $|\iota_{it}|$, and the last line is by the fact that $\max_{i\in I_3}\frac{1}{T}\sum_{t\in[T]}|\tilde{\iota}_{it}|=O_p\left(\eta_N\right)$ by (B.59) and (B.60). The bounds for other three blocks in $J_{3,i}$ can be established in the same manner. Hence, we have $\max_{i\in I_3}||J_{3,i}||_F=O_p(\eta_N^2)$.

Combining all results above yields the desired result. ■

Lemma B.24 *Under Assumptions 1-9, we have*

(i) $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_F=O_p\left(\eta_N^2\right)$,

(ii) $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{v}_{t,1}^{(1)}-O_0^{(1)}v_{t,1}^0\right)\left(\hat{e}_{1,it}-e_{1,it}\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_F=O_p\left(\eta_N^2\right)$,

(iii) $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}e_{1,it}\left(\dot{v}_{t,1}^{(1)}-O_0^{(1)}v_{t,1}^0\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_F=O_p\left(\eta_N^2\right)$,

(iv) $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}O_1^{(1)}v_{t,1}^0\left(\hat{e}_{1,it}-e_{1,it}\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_F=O_p\left(\eta_N^2\right)$.

Proof. (i) We notice that $\mathbb{E}\left[\left(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\middle|\mathscr{D}^{I_1\cup I_2}\right]=0$ by Assumption 1(ii). Define event $\mathscr{A}_{10}(M)=\left\{\max_{t\in[T]}\left\|\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right\|_2\leq M\xi_N\right\}$ with $\mathbb{P}\left(\mathscr{A}_{10}(M)^c\right)\leq e$ for any $e>0$ by Theorem 3.2(ii). With some positive constant c_{24}, it follows that

$$
\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_F>c_{24}\eta_N^2\right)\\
&\leq\mathbb{P}\left(\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_F>c_{24}\eta_N^2,\mathscr{A}_{10}(M)\right)+e\\
&\leq\sum_{i\in I_3}\mathbb{P}\left(\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_F>c_{24}\eta_N^2,\mathscr{A}_{10}(M)\right)+e\\
&\leq\sum_{i\in I_3}\mathbb{E}\mathbb{P}\left(\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_F>c_{24}\eta_N^2\middle|\mathscr{D}^{I_1\cup I_2}\right)\mathbf{1}\left\{\mathscr{A}_{10}(M)\right\}+e\\
&\leq\sum_{i\in I_3}\exp\left\{-\frac{c_{12}c_{24}^2T^2\eta_N^4}{M^2T\eta_N^2+c_{24}MT\eta_N^2\log T\log\log T}\right\}+e\\
&=o(1)+e. \qquad\qquad \text{(B.67)}
\end{aligned}
$$

Since e can be made arbitrarily small, this completes the proof of statement (i).

(ii) It's clear that

$$
\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\left(\dot{v}_{t,1}^{(1)}-O_0^{(1)}v_{t,1}^0\right)\left(\hat{e}_{1,it}-e_{1,it}\right)\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)\right\|_F
$$

$$\leq \max_{i\in I_3, t\in[T]} |\hat{e}_{1,it} - e_{1,it}| \max_{t\in[T]} \left\| \dot{v}_{t,1}^{(1)} - O_0^{(1)} v_{t,1}^0 \right\|_2 = O_p(\eta_N^2),$$

where the equality holds by Lemma B.21(iii) and Theorem 3.2(ii).

(iii) Noting that

$$\mathbb{E}\left[e_{1,it} \left(\dot{v}_{t,1}^{(1)} - O_0^{(1)} v_{t,1}^0 \right) (\tau - \mathbf{1}\{\epsilon_{it} \leq 0\}) \middle| \mathscr{D}^{I_1\cup I_2} \right] = 0$$

by law of iterated expectation and Assumption 1(ii), we can obtain the desired result as in (B.67).

(iv) Note that

$$\begin{aligned}
&\frac{1}{T}\sum_{t\in[T]} v_{t,1}^0 \left(\hat{e}_{1,it} - e_{1,it} \right) (\tau - \mathbf{1}\{\epsilon_{it} \leq 0\}) \\
&= \frac{1}{T^2}\sum_{s\in[T]}\sum_{t\in[T]} v_{t,1}^0 (\tau - \mathbf{1}\{\epsilon_{it} \leq 0\}) w_{j,t}^{0\prime} \left(\frac{W_j^{0\prime} W_j^0}{T} \right)^{-1} w_{j,s}^0 e_{j,is} \\
&+ \frac{1}{NT}\sum_{m\in[N]}\sum_{t\in[T]} v_{t,1}^0 (\tau - \mathbf{1}\{\epsilon_{it} \leq 0\}) l_{j,i}^{0\prime} l_{j,m}^0 e_{j,ms} + \mathcal{R}_{j,it}.
\end{aligned} \tag{B.68}$$

For the first term on the RHS of (B.68), we have

$$\begin{aligned}
&\max_{i\in I_3} \left\| \frac{1}{T^2}\sum_{s\in[T]}\sum_{t\in[T]} v_{t,1}^0 (\tau - \mathbf{1}\{\epsilon_{it} \leq 0\}) w_{j,t}^{0\prime} \left(\frac{W_j^{0\prime} W_j^0}{T} \right)^{-1} w_{j,s}^0 e_{j,is} \right\|_2 \\
&\leq \max_{i\in I_3} \left\| \frac{1}{T}\sum_{t\in[T]} v_{t,1}^0 (\tau - \mathbf{1}\{\epsilon_{it} \leq 0\}) w_{j,t}^{0\prime} \right\|_F \left\| \frac{1}{T}\sum_{s\in[T]} \left(\frac{W_j^{0\prime} W_j^0}{T} \right)^{-1} w_{j,s}^0 e_{j,is} \right\|_2 \\
&= O_p\left(\eta_N^2 \right).
\end{aligned}$$

For the second term on RHS of (B.68), by Assumption 9(ii), Bernstein's inequality in Lemma B.12(i) conditional on factors, we have

$$\max_{i\in I_3} \left\| \frac{1}{NT}\sum_{m\in[N]}\sum_{t\in[T]} v_{t,1}^0 (\tau - \mathbf{1}\{\epsilon_{it} \leq 0\}) l_{j,i}^{0\prime} l_{j,m}^0 e_{j,ms} \right\|_2 = O_p\left(\eta_N^2 \right).$$

Then statement (iv) follows. ■

Lemma B.25 *Under Assumptions 1-9, we have*

(i) $\max_{i\in I_3} \left\| \frac{1}{T}\sum_{t=1}^T e_{1,it} O_1^{(1)} v_{t,1}^0 f_{it}(0) \left(\mu_{1,it} u_{i,1}^{0\prime} v_{t,1}^0 - \hat{\mu}_{1,it} \dot{u}_{i,1}^{(1)\prime} \dot{v}_{t,1}^{(1)} \right) \right\|_2 = O_p\left(\eta_N^2 \right),$

(ii) $\max_{i\in I_3} \left\| \frac{1}{T}\sum_{t=1}^T O_0^{(1)} v_{t,0}^0 f_{it}(0) \left(\mu_{1,it} u_{i,1}^{0\prime} v_{t,1}^0 - \hat{\mu}_{1,it} \dot{u}_{i,1}^{(1)\prime} \dot{v}_{t,1}^{(1)} \right) \right\|_2 = O_p(\eta_N).$

Proof. (i) Recall from (B.59), we have

$$\begin{aligned}
\hat{\mu}_{1,it} \dot{u}_{i,1}^{(1)\prime} \dot{v}_{t,1}^{(1)} - \mu_{1,it} u_{i,1}^{0\prime} v_{t,1}^0 &= \left(\hat{\mu}_{1,it} - \mu_{1,it} \right) u_{i,j}^{0\prime} v_{t,j}^0 + \mu_{1,it} \left(\dot{u}_{i,1}^{(1)} - O_1^{(1)} u_{i,1}^0 \right) O_1^{(1)} v_{t,1}^0 \\
&+ \mu_{1,it} \left(O_1^{(1)} u_{i,1}^0 \right)' \left(\dot{v}_{t,1}^{(1)} - O_1^{(1)} v_{t,1}^0 \right) + O_p\left(\eta_N^2 \right) \quad \text{uniformly over } i \in I_3, t \in [T].
\end{aligned}$$

Then

$$\frac{1}{T}\sum_{t=1}^{T} e_{1,it}O_1^{(1)}v_{t,1}^0 f_{it}(0)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0 - \hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)$$
$$= \frac{1}{T}\sum_{t=1}^{T} e_{1,it}O_1^{(1)}v_{t,1}^0 f_{it}(0)u_{i,j}^{0\prime}v_{t,j}^0\left(\hat{\mu}_{1,it}-\mu_{1,it}\right)$$
$$+\frac{1}{T}\sum_{t=1}^{T} e_{1,it}O_1^{(1)}v_{t,1}^0 f_{it}(0)\mu_{1,it}\left(\dot{u}_{i,1}^{(1)}-O_1^{(1)}u_{i,1}^0\right)' O_1^{(1)}v_{t,1}^0$$
$$+\frac{1}{T}\sum_{t=1}^{T} e_{1,it}O_1^{(1)}v_{t,1}^0 f_{it}(0)\mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)}-O_1^{(1)}v_{t,1}^0\right)+O_p\left(\eta_N^2\right). \tag{B.69}$$

First, note that $\mathbb{E}\left(e_{1,it}v_{t,1}^0 f_{it}(0)\mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)}-O_1^{(1)}v_{t,1}^0\right)\middle|\mathscr{D}^{I_1\cup I_2}\right)=0$ by Assumption 9(i). Following similar arguments used in the proof of Lemma B.21(i), we can show that the third term on the RHS of (B.69) is $O_p\left(\eta_N^2\right)$. By analogous arguments in (B.64) with the fact that $\hat{\mu}_{1,it}-\mu_{1,it}=e_{1,it}-\hat{e}_{1,it}$, we obtain that the first term on the RHS of (B.69) is $O_p\left(\eta_N^2\right)$. In addition

$$\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T} e_{1,it}v_{t,1}^0 f_{it}(0)\mu_{1,it}\left(\dot{u}_{i,1}^{(1)}-O_1^{(1)}u_{i,1}^0\right)' O_1^{(1)}v_{t,1}^0\right\|_2$$
$$\leq \max_{i\in I_3}||\dot{u}_{i,1}^{(1)}-O_1^{(1)}u_{i,1}^0||_2 \max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T} f_{it}(0)e_{1,it}\mu_{1,it}v_{t,1}^0 v_{t,1}^{0\prime}\right\|_F = O_p\left(\eta_N^2\right), \tag{B.70}$$

where the equality holds by the fact that

$$\mathbb{P}\left(\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T} f_{it}(0)e_{1,it}\mu_{1,it}v_{t,1}^0 v_{t,1}^{0\prime}\right\|_F > c_{25}\sqrt{\frac{\log(N\vee T)}{T}}\xi_N\middle|\mathscr{D}\right)=o(1)$$

as in (B.57). Noted that Assumption 1(iv) implies factor and factor loading of $X_{j,it}$ is uniformly bounded for $\forall j\in[p]$, which indicates that $\mu_{j,it}$ is also uniformly bounded a.s.. This completes the proof statement (i).

(ii). Note that uniformly over $i\in I_3$, we have

$$\frac{1}{T}\sum_{t=1}^{T} O_0^{(1)}v_{t,0}^0 f_{it}(0)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0-\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}\right)$$
$$=\frac{1}{T}\sum_{t=1}^{T} O_0^{(1)}v_{t,0}^0 f_{it}(0)\left(\hat{\mu}_{1,it}-\mu_{1,it}\right)u_{i,j}^{0\prime}v_{t,j}^0+\frac{1}{T}\sum_{t=1}^{T} O_0^{(1)}v_{t,0}^0 f_{it}\mu_{1,it}\left(\dot{u}_{i,1}^{(1)}-O_1^{(1)}u_{i,1}^0\right)O_1^{(1)}v_{t,1}^0$$
$$+\frac{1}{T}\sum_{t=1}^{T} O_0^{(1)}v_{t,0}^0 f_{it}\mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)}-O_1^{(1)}v_{t,1}^0\right)+O_p\left(\eta_N^2\right)$$
$$=O_p(\eta_N). \tag{B.71}$$

This term will remain as the bias term in the linear expansion of $\hat{u}_{i,0}^{(3,1)}$. ■

Lemma B.26 *Under Assumptions 1-9, we have*

(i) $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T} e_{1,it}O_1^{(1)}v_{t,1}^0\left\{\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{\epsilon_{it}\leq \iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right\}\right]\right.\right.$
$\left.\left.-\left(F_{it}(0)-F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]\right)\right\}\right\|_2 = o_p\left((N\vee T)^{-\frac{1}{2}}\right),$

(ii) $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T} O_0^{(1)}v_{t,0}^0\left\{\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{\epsilon_{it}\leq \iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right\}\right]\right.\right.$
$\left.\left.-\left(F_{it}(0)-F_{it}\left[\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right)\right]\right)\right\}\right\|_2 = o_p\left((N\vee T)^{-\frac{1}{2}}\right).$

Proof. (i) We still assume $K_1 = 1$ for notation simplicity. Let $O^{(1)} = diag\left(O_0^{(1)}, O_1^{(1)}\right)$. Recall from Theorem 3.2(iii) that with

$$D_i^I = O^{(1)}\frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\begin{bmatrix} v_{t,0}^0 v_{t,0}' & v_{t,0}^0 v_{t,1}^{0\prime}X_{1,it} \\ v_{t,1}^0 v_{t,0}^{0\prime}X_{1,it} & v_{t,1}^0 v_{t,1}^{0\prime}X_{1,it}^2\end{bmatrix}O^{(1)\prime}, \quad D_i^{II} = O^{(1)}\frac{1}{T}\sum_{t=1}^{T}\left[\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right]\begin{bmatrix} v_{t,0}^0 \\ v_{t,1}^0 X_{1,it}\end{bmatrix}$$

$$\mathbb{J}_i\left(\left\{\dot{\Delta}_{t,v}\right\}_{t\in[T]}\right) := O^{(1)}\frac{1}{T}\sum_{t=1}^{T}\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{\epsilon_{it}\leq \dot{\Delta}_{t,v}'\Psi_{it}^0\right\}\right]\begin{bmatrix} v_{t,0}^0 \\ v_{t,1}^0 X_{1,it}\end{bmatrix},$$

uniformly over $i \in I_3$, we have

$$\begin{aligned}
\dot{\Delta}_{i,u} &= \left[D_i^I\right]^{-1}\left[D_i^{II}+\mathbb{J}_i\left(\left\{\dot{\Delta}_{t,v}\right\}_{t\in[T]}\right)\right]+o_p\left((N\vee T)^{-\frac{1}{2}}\right) \\
&= \left(O^{(1)\prime}\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^{T} f_{it}(0)\begin{bmatrix} v_{t,0}^0 v_{t,0}' & v_{t,0}^0 v_{t,1}^{0\prime}X_{1,it} \\ v_{t,1}^0 v_{t,0}^{0\prime}X_{1,it} & v_{t,1}^0 v_{t,1}^{0\prime}X_{1,it}^2\end{bmatrix}\right)^{-1}\left[\frac{1}{T}\sum_{t=1}^{T}\left[\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right]\begin{bmatrix} v_{t,0}^0 \\ v_{t,1}^0 X_{1,it}\end{bmatrix}\right. \\
&\quad\left.+\frac{1}{T}\sum_{t=1}^{T}\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{\epsilon_{it}\leq \dot{\Delta}_{t,v}'\Psi_{it}^0\right\}\right]\begin{bmatrix} v_{t,0}^0 \\ v_{t,1}^0 X_{1,it}\end{bmatrix}\right]+o_p\left((N\vee T)^{-\frac{1}{2}}\right) \\
&:= h_i + o_p\left((N\vee T)^{-\frac{1}{2}}\right).
\end{aligned}$$

Let $\mathfrak{l} = \left(\mathbf{0}_{K_0}', \mathbf{1}_{K_1}'\right)'$ with $\mathbf{0}_{K_0}$ being a $K_0\times 1$ vector of zeros and $\mathbf{1}_{K_1}$ a $K_1 \times 1$ vector of ones. Let

$$h_i^I = \mathfrak{l}' h_i. \tag{B.72}$$

Then we have

$$O_1^{(1)\prime}\dot{u}_{i,1}^{(1)} - u_{i,1}^0 = h_i^I + o_p\left((N\vee T)^{-\frac{1}{2}}\right) \quad \text{uniformly over } i\in I_3,$$

and $\max_{i\in I_3}||h_i^I||_2 = O_p(\eta_N)$.

Combining (B.58)-(B.60), uniformly in $i \in I_3$ and $t \in [T]$ we have

$$\begin{aligned}
&\iota_{it}\left(O_0^{(1)}u_{i,0}^0, O_1^{(1)}u_{i,1}^0, \dot{u}_{i,1}^{(1)}\right) \\
&= \left[\left(O_0^{(1)}u_{i,0}^0\right)'\left(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right)\right]+\left(\hat{\mu}_{1,it}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t,1}^{(1)}-\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right)+\left(\hat{e}_{1,it}u_{i,1}^{0\prime}O_1^{(1)}\dot{v}_{t,1}^{(1)}-e_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\right) \\
&= \left(O_0^{(1)}u_{i,0}^0\right)'\left(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right)+\left(\hat{\mu}_{1,it}^{(1)}-\mu_{1,it}\right)\left(O_1^{(1)}u_{i,1}^0\right)'O_1^{(1)}v_{t,1}^0 \\
&+\mu_{1,it}\left(\dot{u}_{i,1}^{(1)}-O_1^{(1)}u_{i,1}^0\right)'O_1^{(1)}v_{t,1}^0+\mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)}-O_1^{(1)}v_{t,1}^0\right)+\left(\hat{e}_{1,it}-e_{1,it}\right)\left(O_1^{(1)}u_{i,1}^0\right)'O_1^{(1)}v_{t,1}^0
\end{aligned}$$

$$
\begin{aligned}
&+ e_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right) + o_p\left((N\vee T)^{-\frac{1}{2}}\right)\\
&= \mu_{1,it}\left(\dot{u}_{i,1}^{(1)} - O_1^{(1)}u_{i,1}^0\right)'O_1^{(1)}v_{t,1}^0 + \left(O_0^{(1)}u_{i,0}^0\right)'\left(\dot{v}_{t,0}^{(1)} - O_0^{(1)}v_{t,0}^0\right) + \left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right)X_{1,it}\\
&+ o_p\left((N\vee T)^{-\frac{1}{2}}\right)\\
&= \mu_{1,it}v_{t,1}^{0\prime}h_i^I + \left\{\left(O_0^{(1)}u_{i,0}^0\right)'\left(\dot{v}_{t,0}^{(1)} - O_0^{(1)}v_{t,0}^0\right) + \left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right)X_{1,it}\right\} + \mathcal{R}_{\iota,it}\\
&:= \mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right) + \mathcal{R}_{\iota,it} \qquad \text{(B.73)}\\
&:= R_{\iota,it}^1\left(|\mu_{1,it}| + |e_{1,it}|\right) + R_{\iota,it}^2
\end{aligned}
$$

such that $\max_{i\in I_3,t\in[T]}|\mathcal{R}_{\iota,it}| = o_p\left((N\vee T)^{-\frac{1}{2}}\right)$, $\max_{i\in I_3,t\in[T]}\left|R_{\iota,it}^1\right| = O_p(\eta_N)$ and $\max_{i\in I_3,t\in[T]}\left|R_{\iota,it}^2\right| = O_p(\eta_N)$, where we use the fact that $\hat{\mu}_{1,it} - \mu_{1,it} + \hat{e}_{1,it} - e_{1,it} = 0$, and $h_{it}^{II}\left(\dot{\Delta}_{t,v}\right) = \Psi_{it}^{0\prime}\dot{\Delta}_{t,v}$ with $\Psi_{it} = \left((O_0^{(1)}u_{i,0}^0)', (O_1^{(1)}u_{i,1}^0X_{1,it})'\right)'$. Let

$$
\hat{\mathbb{I}}_{3,it}^I = e_{1,it}v_{t,1}^0\left\{\left[\mathbf{1}\{\epsilon_{it}\le 0\} - \mathbf{1}\left\{\epsilon_{it}\le \mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right\}\right] - \left(F_{it}(0) - F_{it}\left[\mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right]\right)\right\},
$$

$$
\begin{aligned}
\hat{\mathbb{I}}_{3,it}^{II} = e_{1,it}v_{t,1}^0\Bigg\{&\left[\mathbf{1}\left\{\epsilon_{it}\le \mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right) + \mathcal{R}_{\iota,it}\right\} - \mathbf{1}\left\{\epsilon_{it}\le \mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right\}\right]\\
&- \left(F_{it}\left[\mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right) + \mathcal{R}_{\iota,it}\right] - F_{it}\left[\mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right]\right)\Bigg\}.
\end{aligned}
$$

We first show that $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T\hat{\mathbb{I}}_{3,it}^I\right\|_2 = O_p\left(\eta_N^2\right)$.

For some sufficiently large constant M, define event $\mathscr{A}_{11}(M) = \left\{\max_{i\in I_3}\left\|h_i^I\right\|_2 \le M\eta_N\right\}$. We have $\mathbb{P}\left(\mathscr{A}_{11}^c(M)\right)\le e$ for any $e>0$. For some positive large enough constant c_{26}, we have

$$
\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T\hat{\mathbb{I}}_{3,it}^I\right\|_2 > c_{26}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}\right)\\
&\le \mathbb{P}\left(\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T\hat{\mathbb{I}}_{3,it}^I\right\|_2 > c_{26}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}, \mathscr{A}_{11}(M)\right) + e\\
&\le \mathbb{P}\left(\max_{i\in I_3}\sup_{\xi\in\Xi}\left\|\frac{1}{T}\sum_{t=1}^T\hat{\mathbb{I}}_{3,it}^I(\xi)\right\|_2 > c_{26}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}\right) + e. \qquad \text{(B.74)}
\end{aligned}
$$

with $\Xi^1 := \left\{\xi\in\mathbb{R}^{(K_0+K_1)\times 1} : \|\xi\|_2\le M\eta_N\right\}$ and

$$
\hat{\mathbb{I}}_{3,it}^I(\xi) = e_{1,it}v_{t,1}^0\left\{\left[\mathbf{1}\{\epsilon_{it}\le 0\} - \mathbf{1}\left\{\epsilon_{it}\le \mu_{1,it}v_{t,1}^{0\prime}\xi + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right\}\right] - \left[F_{it}(0) - F_{it}\left(\mu_{1,it}v_{t,1}^{0\prime}\xi + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right)\right]\right\}.
$$

Divide Ξ^1 into sub classes Ξ_s^1 for $s = 1,...,n_{\Xi^1}$ such that $\left\|\xi - \tilde{\xi}\right\|_2 < \frac{\varepsilon}{T}$ for $\forall\xi,\ \bar{\xi}\in\Xi_s^1$ and $n_{\Xi^1}\asymp T^{K_0+K_1}$. With analogous arguments from (A.14)-(A.17), for $\forall\xi_s\in\Xi_s^1$, we have

$$
\left\|\frac{1}{T}\sum_{t\in[T]}\hat{\mathbb{I}}_{3,it}^I(\xi)\right\|_2 \le \left\|\frac{1}{T}\sum_{t\in[T]}\hat{\mathbb{I}}_{3,it}^I(\xi_s)\right\|_2 + \left\|\frac{1}{T}\sum_{t\in[T]}\left[\hat{\mathbb{I}}_{3,it}^I(\xi) - \hat{\mathbb{I}}_{3,it}^I(\xi_s)\right]\right\|_2
$$

with

$$
\begin{aligned}
&\max_{i\in I_3, s\in[n_{\Xi 1}]}\sup_{\xi\in\Xi_s}\left\|\frac{1}{T}\sum_{t\in[T]}\left[\hat{\mathbb{I}}^{I}_{3,it}(\xi)-\hat{\mathbb{I}}^{I}_{3,it}(\xi_s)\right]\right\|_2 \\
&\le \max_{i\in I_3}\frac{1}{T}\sum_{t\in[T]}\mathbb{E}\left[\left\|e_{1,it}v^{0}_{t,1}\right\|_2 \mathbf{1}\left\{\left|\epsilon_{it}-h^{II}_{it}\left(\dot{\Delta}_{t,v}\right)\right|\le|\mu_{1,it}|\left\|v^{0\prime}_{t,1}\right\|_2\frac{\varepsilon}{T}\right\}\middle|\mathscr{D}_e^{I_1\cup I_2}\right] \\
&+\max_{i\in I_3}\left\|\hat{\mathbb{I}}^{III}_{3,i}\right\|_2+\max_{i\in I_3}\frac{1}{T}\sum_{t\in[T]}\left\|e_{1,it}v^{0}_{t,1}\right\|_2|\mu_{1,it}|\left\|v^{0\prime}_{t,1}\right\|_2\frac{\varepsilon}{T} \\
&\le \max_{i\in I_3}\left\|\hat{\mathbb{I}}^{III}_{3,i}\right\|_2+O\left(\frac{\varepsilon}{T}\right)\ a.s. \qquad \text{(B.75)}
\end{aligned}
$$

where $\hat{\mathbb{I}}^{III}_{3,i}=\frac{1}{T}\sum_{t\in[T]}\hat{\mathbb{I}}^{III}_{3,it}$ and

$$
\begin{aligned}
\hat{\mathbb{I}}^{III}_{3,it}=\left\|e_{1,it}v^{0}_{t,1}\right\|_2\Bigg\{&\mathbf{1}\left\{\left|\epsilon_{it}-h^{II}_{it}\left(\dot{\Delta}_{t,v}\right)\right|\le|\mu_{1,it}|\left\|v^{0\prime}_{t,1}\right\|_2\frac{\varepsilon}{T}\right\} \\
&-\mathbb{E}\left[\mathbf{1}\left\{\left|\epsilon_{it}-h^{II}_{it}\left(\dot{\Delta}_{t,v}\right)\right|\le|\mu_{1,it}|\left\|v^{0\prime}_{t,1}\right\|_2\frac{\varepsilon}{T}\right\}\middle|\mathscr{D}_e^{I_1\cup I_2}\right]\Bigg\}.
\end{aligned}
$$

Similarly to (A.17), we can show that

$$
\mathbb{P}\left(\max_{i\in I_3}\left\|\hat{\mathbb{I}}^{III}_{3,it}\right\|_2>c_{26}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}\right)=o(1),\quad\text{and} \qquad \text{(B.76)}
$$

$$
\mathbb{P}\left(\max_{i\in I_3}\max_{s\in[n_{\Xi 1}]}\left\|\hat{\mathbb{I}}^{I}_{3,it}\left(\xi_s\right)\right\|_2>c_{26}\xi_N^{\frac{5+\vartheta}{2+\vartheta}}\left(\frac{\log(N\vee T)}{N\wedge T}\right)^{\frac{1}{4+2\vartheta}}\sqrt{\frac{\log(N\vee T)}{T}}\right)=o(1). \qquad \text{(B.77)}
$$

Combining (B.74)-(B.76), we have

$$
\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{\mathbb{I}}^{I}_{3,it}\right\|_2=o_p\left((N\vee T)^{-\frac{1}{2}}\right).
$$

Next, we notice that

$$
\begin{aligned}
&\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{\mathbb{I}}^{II}_{3,it}\right\|_2 \\
&\le \max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|e_{1,it}v^{0}_{t,1}\right\|_2\mathbf{1}\left\{\left|\mu_{1,it}v^{0\prime}_{t,1}h^{I}_{i}+h^{II}_{it}\left(\dot{\Delta}_{t,v}\right)\right|\le|\epsilon_{it}|\le\left|\mu_{1,it}v^{0\prime}_{t,1}h^{I}_{i}+h^{II}_{it}\left(\dot{\Delta}_{t,v}\right)\right|+|\mathcal{R}_{\iota,it}|\right\} \\
&+\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|e_{1,it}v^{0}_{t,1}\right\|_2\left|F_{it}\left[\mu_{1,it}v^{0\prime}_{t,1}h^{I}_{i}+h^{II}_{it}\left(\dot{\Delta}_{t,v}\right)+\mathcal{R}_{\iota,it}\right]-F_{it}\left[\mu_{1,it}v^{0\prime}_{t,1}h^{I}_{i}+h^{II}_{it}\left(\dot{\Delta}_{t,v}\right)\right]\right| \\
&\le \max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|e_{1,it}v^{0}_{t,1}\right\|_2\mathbf{1}\left\{\left|\mu_{1,it}v^{0\prime}_{t,1}h^{I}_{i}+h^{II}_{it}\left(\dot{\Delta}_{t,v}\right)\right|\le|\epsilon_{it}|\le\left|\mu_{1,it}v^{0\prime}_{t,1}h^{I}_{i}+h^{II}_{it}\left(\dot{\Delta}_{t,v}\right)\right|+|\mathcal{R}_{\iota,it}|\right\} \\
&+\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|e_{1,it}v^{0}_{t,1}\right\|_2|\mathcal{R}_{\iota,it}| \\
&=\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\hat{\mathbb{I}}^{IV}_{3,it}+o_p\left((N\vee T)^{-\frac{1}{2}}\right), \qquad \text{(B.78)}
\end{aligned}
$$

where the first inequality is by triangle inequality, the second inequality is by mean value theorem and the last line is by the uniform convergence rate for $\mathcal{R}_{\iota,it}$ with

$$\hat{\mathbb{I}}_{3,it}^{IV} = \left\|e_{1,it}v_{t,1}^0\right\|_2 \mathbf{1}\left\{\left|\mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right| \leq |\epsilon_{it}| \leq \left|\mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right| + |\mathcal{R}_{\iota,it}|\right\}.$$

Define the event $\mathscr{A}_{12}(M) := \left\{\max_{i\in I_3,t\in[T]} |\mathcal{R}_{\iota,it}| \leq M\eta_N^2\right\}$ with $\mathbb{P}\left\{\mathscr{A}_{12}^c(M)\right\} \leq e$ for any $e > 0$. Then for a large enough constant c_{26}, similarly as (A.37), we have

$$\begin{aligned}
&\mathbb{P}\left(\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{\mathbb{I}}_{3,it}^{IV}\right\|_2 > c_{26}\eta_N^2\right) \leq \mathbb{P}\left(\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{\mathbb{I}}_{3,it}^{IV}\right\|_2 > c_{26}\eta_N^2, \mathscr{A}_{12}(M)\right) + e \\
&\leq \mathbb{P}\left(\max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\tilde{\mathbf{1}}_{it}(h_i^I) > c_{26}\eta_N^2\right) + e \\
&\leq \mathbb{P}\left(\sup_{i\in I_3,\xi\in\Xi^1}\frac{1}{T}\sum_{t=1}^{T}\tilde{\mathbf{1}}_{it}(\xi) > c_{26}\eta_N^2\right) + 2e \\
&\leq \mathbb{P}\left(\max_{i\in I_3}\sup_{\xi\in\Xi^1}\left|\frac{1}{T}\sum_{t=1}^{T}\tilde{\mathbf{1}}_{it}(\xi) - \bar{\tilde{\mathbf{1}}}_{it}(\xi)\right| > \frac{c_{26}\eta_N^2}{2}\right) + \mathbb{P}\left(\max_{i\in I_3}\sup_{\xi\in\Xi^1}\left|\frac{1}{T}\sum_{t=1}^{T}\bar{\tilde{\mathbf{1}}}_{it}(\xi)\right| > \frac{c_{26}\eta_N^2}{2}\right) + 2e \\
&= \mathbb{E}\left[\mathbb{P}\left(\max_{i\in I_3}\sup_{\xi\in\Xi^1}\left|\frac{1}{T}\sum_{t=1}^{T}\tilde{\mathbf{1}}_{it}(\xi) - \bar{\tilde{\mathbf{1}}}_{it}(\xi)\right| > \frac{c_{26}\eta_N^2}{2}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\right] + \mathbb{P}\left(\max_{i\in I_3}\sup_{\xi\in\Xi^1}\left|\frac{1}{T}\sum_{t=1}^{T}\bar{\tilde{\mathbf{1}}}_{it}(\xi)\right| > \frac{c_{26}\eta_N^2}{2}\right) + 2e
\end{aligned} \tag{B.79}$$

where $\tilde{\mathbf{1}}_{it}(h_i^I) := \left\|e_{1,it}v_{t,1}^0\right\|_2 \mathbf{1}\left\{\left|\mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right| \leq |\epsilon_{it}| \leq \left|\mu_{1,it}v_{t,1}^{0\prime}h_i^I + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right| + M\eta_N^2\right\}$ and $\bar{\tilde{\mathbf{1}}}_{it}(h_i^I) = \mathbb{E}\left[\tilde{\mathbf{1}}_{it}(h_i^I)\middle|\mathscr{D}_e^{I_1\cup I_2}\right]$. By analogous arguments for the first term on the RHS of inequality (B.74), we can show that

$$\mathbb{E}\left[\mathbb{P}\left(\max_{i\in I_3}\sup_{\xi\in\Xi^1}\left|\frac{1}{T}\sum_{t=1}^{T}\tilde{\mathbf{1}}_{it}(\xi) - \bar{\tilde{\mathbf{1}}}_{it}(\xi)\right| > \frac{c_{26}\eta_N^2}{2}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\right] = o(1).$$

Besides, we observe that

$$\begin{aligned}
&\max_{i\in I_3}\sup_{\xi\in\Xi^1}\left|\frac{1}{T}\sum_{t=1}^{T}\bar{\tilde{\mathbf{1}}}_{it}(\xi)\right| \\
&\leq \max_{i\in I_3}\sup_{\xi\in\Xi^1}\frac{1}{T}\sum_{t=1}^{T}\left\|e_{1,it}v_{t,1}^0\right\|_2\left|F_{it}\left[\left|\mu_{1,it}v_{t,1}^{0\prime}\xi + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right| + M\eta_N^2\right] - F_{it}\left[\left|\mu_{1,it}v_{t,1}^{0\prime}\xi + h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)\right|\right]\right| \\
&\leq \max_{i\in I_3}\frac{1}{T}\sum_{t=1}^{T}\left\|e_{1,it}v_{t,1}^0\right\|_2 M\eta_N^2,
\end{aligned}$$

which yields

$$\mathbb{P}\left(\max_{i\in I_3}\sup_{\xi\in\Xi^1}\left|\frac{1}{T}\sum_{t=1}^{T}\bar{\tilde{\mathbf{1}}}_{it}(\xi)\right| > \frac{c_{26}\eta_N^2}{2}\right) = o(1). \tag{B.80}$$

Combining (B.78)-(B.80) yields $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T}\hat{\mathbb{I}}_{3,it}^{II}\right\|_2 = O_p\left(\eta_N^2\right) = o_p\left((N\vee T)^{-\frac{1}{2}}\right)$ by Assumption 1(ix). This concludes the proof of statement (i).

(ii) The proof is analogous to that of part (i) and thus omitted. ■

Lemma B.27 *Under Assumptions 1-9, uniformly in $i \in I_3$, we can show that*

(i) $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T} e_{1,it}O_1^{(1)}v_{t,1}^0\left\{\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le \iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right\}-\left(F_{it}(0)-F_{it}\left[\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right]\right)\right\}\right\|_2$
$= o_p\left((N\vee T)^{-\frac{1}{2}}\right)$,

(ii) $\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^{T} v_{t,0}^0\left\{\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le \iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right\}-\left(F_{it}(0)-F_{it}\left[\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\right]\right)\right\}\right\|_2$
$= o_p\left((N\vee T)^{-\frac{1}{2}}\right)$.

Proof. As in (B.73), we can show that

$$
\begin{aligned}
&\iota_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\dot{u}_{i,1}^{(1)}\right)\\
&=\mu_{1,it}v_{t,1}^{0\prime}h_i^I+\left[\left(\hat{u}_{i,0}^{(3,1)}-O_0^{(1)}u_{i,0}^0\right)^{\prime}O_0^{(1)}v_{t,0}^0+e_{1,it}\left(\hat{u}_{i,1}^{(3,1)}-O_1^{(1)}u_{i,1}^0\right)^{\prime}O_1^{(1)}v_{t,1}^0\right]\\
&+\left[\left(O_0^{(0)}u_{i,0}^0\right)^{\prime}\left(\dot{v}_{t,0}^{(1)}-O_0^{(1)}v_{t,0}^0\right)+X_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)^{\prime}\left(\dot{v}_{t,1}^{(1)}-O_1^{(1)}v_{t,1}^0\right)\right]+\mathcal{R}_{\hat{\iota},it}\\
:&=\mu_{1,it}v_{t,1}^{0\prime}h_i^I+\left(\hat{u}_{i,0}^{(3,1)}-O_0^{(1)}u_{i,0}^0\right)^{\prime}O_0^{(1)}v_{t,0}^0+e_{1,it}\left(\hat{u}_{i,1}^{(3,1)}-O_1^{(1)}u_{i,1}^0\right)^{\prime}O_1^{(1)}v_{t,1}^0+h_{it}^{II}\left(\dot{\Delta}_{t,v}\right)+\mathcal{R}_{\hat{\iota},it},
\end{aligned}
\tag{B.81}
$$

where $\max_{i\in I_3,t\in[T]}\left|\mathcal{R}_{\hat{\iota},it}\right|=o_p\left((N\vee T)^{-\frac{1}{2}}\right)$. As in the proof of Theorem 3.2, we can show that $\max_{i\in I_3}\left\|\hat{u}_{i,j}^{(3,1)}-O_j^{(1)}u_{i,j}^0\right\|_2$ $=O_p\left(\eta_N\right)$ for $\forall j\in[p]\cup\{0\}$. Then by changing the event set $\mathscr{A}_{11}(M)$ to

$$
\left\{\max_{i\in I_3}\left\|h_i^I\right\|_2\le M\eta_N,\max_{i\in I_3}\left\|\hat{u}_{i,j}-O_j^{(1)}u_{i,j}^0\right\|_2\le M\eta_N\right\},
$$

we can repeat the analysis in Lemma B.26 and obtain the desired results for statement (i). With some obvious adjustment, statement (ii) can be proved. ∎

Lemma B.28 *Under Assumptions 1-9, for block matrices $\hat{D}_t^F$, D_t^F, $\hat{D}_t^J$ and D_t^J defined in Appendix A, we have*

$$
\max_{t\in[T]}||\hat{D}_t^F-D_t^F||_F=O_p(\eta_N),\ \text{and}\ \max_{t\in[T]}||\hat{D}_t^J-D_t^J||_F=\left\|\begin{bmatrix}O_p\left(\eta_N^2\right) & O_p(\eta_N)\\ O_p\left(\eta_N^2\right) & O_p\left(\eta_N^2\right)\end{bmatrix}\right\|_F.
$$

Proof. By analogous arguments as used in the proofs of Lemmas B.22 and B.23, we can prove the lemma. ∎

Lemma B.29 *Under Assumptions 1-9, we have*

(i) $\max_{t\in[T]}\left\|\frac{1}{N_3}\sum_{i\in I_3}O_{u,0}^{(1)}u_{i,0}^0f_{it}(0)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0-\hat{\mu}_{1,it}\hat{u}_{i,1}^{(3,1)\prime}\dot{v}_{t,1}^{(1)}\right)\right\|_2=O_p(\eta_N)$,

(ii) $\max_{t\in[T]}\left\|\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}O_{u,1}^{(1)}u_{i,1}^0f_{it}(0)\left(\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0-\hat{\mu}_{1,it}\hat{u}_{i,1}^{(3,1)\prime}\dot{v}_{t,1}^{(1)}\right)\right\|_2=o_p\left((N\vee T)^{-\frac{1}{2}}\right)$.

Proof. (i) Note that

$$
\begin{aligned}
&\hat{\mu}_{1,it}\hat{u}_{i,1}^{(3,1)\prime}\dot{v}_{t,1}^{(1)}-\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}\\
&=\left(\hat{\mu}_{1,it}-\mu_{1,it}\right)\left(\hat{u}_{i,1}^{(3,1)}-O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}\left(\dot{v}_{t,1}^{(1)}-O_{1}^{(1)}v_{t,1}^{0}\right)+\left(\hat{\mu}_{1,it}-\mu_{1,it}\right)\left(O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}\left(\dot{v}_{t,1}^{(1)}-O_{1}^{(1)}v_{t,1}^{0}\right)\\
&+\left(\hat{\mu}_{1,it}-\mu_{1,it}\right)\left(\hat{u}_{i,1}^{(3,1)}-O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}O_{1}^{(1)}v_{t,1}^{0}+\left(\hat{\mu}_{1,it}-\mu_{1,it}\right)\left(O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}O_{1}^{(1)}v_{t,1}^{0}\\
&+\mu_{1,it}\left(\hat{u}_{i,1}^{(3,1)}-O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}\left(\dot{v}_{t,1}^{(1)}-O_{1}^{(1)}v_{t,1}^{0}\right)+\mu_{1,it}\left(\hat{u}_{i,1}^{(3,1)}-O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}O_{1}^{(1)}v_{t,1}^{0}\\
&+\mu_{1,it}\left(O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}\left(\dot{v}_{t,1}^{(1)}-O_{1}^{(1)}v_{t,1}^{0}\right)+\mu_{1,it}\left(O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}O_{1}^{(1)}v_{t,1}^{0}-\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}\\
&=\left(\hat{\mu}_{1,it}-\mu_{1,it}\right)\left(O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}O_{1}^{(1)}v_{t,1}^{0}+\mu_{1,it}\left(\hat{u}_{i,1}^{(3,1)}-O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}O_{1}^{(1)}v_{t,1}^{0}+\mu_{1,it}\left(O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}\left(\dot{v}_{t,1}^{(1)}-O_{1}^{(1)}v_{t,1}^{0}\right)\\
&+\mu_{1,it}v_{t,1}^{0\prime}O_{1}^{(1)\prime}\left(O_{u,1}^{(1)}-O_{1}^{(1)}\right)u_{i,1}^{0}+O_{p}\left(\eta_{N}^{2}\right)\\
&=\left(\hat{\mu}_{1,it}-\mu_{1,it}\right)u_{i,1}^{0\prime}v_{t,1}^{0}+\mu_{1,it}\left(\hat{u}_{i,1}^{(3,1)}-O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}O_{1}^{(1)}v_{t,1}^{0}+\mu_{1,it}\left(O_{1}^{(1)}u_{i,1}^{0}\right)^{\prime}\left(\dot{v}_{t,1}^{(1)}-O_{1}^{(1)}v_{t,1}^{0}\right)\\
&+\mu_{1,it}v_{t,1}^{0\prime}O_{1}^{(1)\prime}\left(O_{u,1}^{(1)}-O_{1}^{(1)}\right)u_{i,1}^{0}+O_{p}\left(\eta_{N}^{2}\right)=O_{p}(\eta_{N}),
\end{aligned}\tag{B.82}
$$

uniformly over $i\in I_3$ and $t\in[T]$, where the last equality holds by the fact that $\left\|O_{u,1}^{(1)}-O_{1}^{(1)}\right\|_F=O_p(\eta_N)$. It follows that

$$
\max_{t\in[T]}\left\|\frac{1}{N_3}\sum_{i\in I_3}O_{u,0}^{(1)}u_{i,0}^{0}f_{it}(0)\left(\hat{\mu}_{1,it}\hat{u}_{i,1}^{(3,1)\prime}\dot{v}_{t,1}^{(1)}-\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}\right)\right\|_2=O_p(\eta_N).
$$

(ii) Observe that

$$
\begin{aligned}
&\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}O_{u,1}^{(1)}u_{i,1}^{0}f_{it}(0)\left(\hat{\mu}_{1,it}^{(1)}\hat{u}_{i,1}^{(3,1)\prime}\dot{v}_{t,1}^{(1)}-\mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^{0}\right)\\
&=O_{u,1}^{(1)}\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}u_{i,1}^{0}f_{it}(0)\left(\hat{\mu}_{1,it}^{(1)}-\mu_{1,it}\right)u_{i,1}^{0\prime}v_{t,1}^{0}\\
&+O_{u,1}^{(1)}\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}u_{i,1}^{0}f_{it}(0)\mu_{1,it}\left(\hat{u}_{i,1}^{(3,1)}-O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}O_{1}^{(1)}v_{t,1}^{0}\\
&+O_{u,1}^{(1)}\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}u_{i,1}^{0}f_{it}(0)\mu_{1,it}\left(O_{1}^{(1)}u_{i,1}^{0}\right)^{\prime}\left(\dot{v}_{t,1}^{(1)}-O_{1}^{(1)}v_{t,1}^{0}\right)\\
&+O_{u,1}^{(1)}\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}u_{i,1}^{0}f_{it}(0)\mu_{1,it}v_{t,1}^{(1)\prime}O_{1}^{(1)\prime}\left(O_{u,1}^{(1)}-O_{1}^{(1)}\right)u_{i,1}^{0}+O_{p}\left(\eta_{N}^{2}\right).
\end{aligned}\tag{B.83}
$$

By similar arguments as used in (B.64), we can show that the first term on the RHS of (B.83) is $o_p\left((N\vee T)^{-\frac{1}{2}}\right)$ uniformly over t. For the second term, by inserting the linear expansion for $\hat{u}_{i,1}^{(3,1)}-O_{u,1}^{(1)}u_{i,1}^{0}$ in (A.41), we notice that

$$
\begin{aligned}
&\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}u_{i,1}^{0}f_{it}(0)\mu_{1,it}\left(\hat{u}_{i,1}^{(3,1)}-O_{u,1}^{(1)}u_{i,1}^{0}\right)^{\prime}O_{1}^{(1)}v_{t,1}^{0}\\
&=\frac{1}{N_3T}\sum_{i\in I_3}\sum_{t\in[T]}e_{1,it}f_{it}\left(0\right)\mu_{1,it}u_{i,1}^{0}v_{t^*,1}^{0\prime}\hat{V}_{u_1}^{-1}e_{1,it^*}v_{t^*,1}^{0}\left(\tau-\mathbf{1}\left\{\epsilon_{it^*}\leq 0\right\}\right)+o_p\left((N\vee T)^{-\frac{1}{2}}\right).
\end{aligned}
$$

By arguments as used in (B.65), we can show the leading term in the last equality is $O_p\left(\sqrt{\frac{\log(N\vee T)}{NT}}\xi_N^2\right)$. Then the second term on the RHS of (B.83) is $o_p\left((N\vee T)^{-\frac{1}{2}}\right)$. For the third and the fourth terms on the RHS of (B.83), by conditional on $\mathscr{D}^{I_1\cup I_2}$, we notice that $\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)\mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)}-O_1^{(1)}v_{t,1}^0\right)$ and $\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)\mu_{1,it}v_{t,1}^{0\prime}O_1^{(1)\prime}\left(O_{u,1}^{(1)}-O_1^{(1)}\right)u_{i,1}^0$ are both mean zero and the randomness depends only on $\{e_{1,it}f_{it}(0)\}$. By conditional Hoeffding's inequality, we can show that both these two terms are $o_p\left((N\vee T)^{-\frac{1}{2}}\right)$. ■

Define

$$
\begin{aligned}
&\mathbb{J}_{it}\left(\dot{\Delta}_{t,v}\right)=\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{\epsilon_{it}\leq\dot{\Delta}_{t,v}'\Psi_{it}^0\right\}\right]\begin{bmatrix}v_{t,0}^0\\ v_{t,1}^0X_{1,it}\end{bmatrix},\\
&h_i^{I,1}=\frac{1}{T}\sum_{t=1}^T\mathbb{E}\left[f_{it}(0)\begin{bmatrix}v_{t,0}^0v_{t,0}' & v_{t,0}^0v_{t,1}^{0\prime}X_{1,it}\\ v_{t,1}^0v_{t,0}^{0\prime}X_{1,it} & v_{t,1}^0v_{t,1}^{0\prime}X_{1,it}^2\end{bmatrix}\middle|\mathscr{D}\right],\\
&h_i^{I,2}=\frac{1}{T}\sum_{t=1}^T\mathbb{E}\left(\mathbb{J}_i\left(\dot{\Delta}_{t,v}\right)\middle|\mathscr{D}^{I_1\cup I_2}\right),\\
&h_{it}^{III}=v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^T v_{t^*,0}^0\left(\tau-1\left\{\epsilon_{it^*}\leq 0\right\}\right)+X_{1,it}v_{t,1}^{0\prime}\hat{V}_{u_1}^{-1}\frac{1}{T}\sum_{t^*=1}^T e_{1,it^*}v_{t^*,1}^0\left(\tau-1\left\{\epsilon_{it^*}\leq 0\right\}\right)\\
&-\left(v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^T v_{t^*,0}^0 f_{it^*}(0)\mu_{1,it^*}v_{t^*,1}^{0\prime}\right)\mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(h_i^{I,1}\right)^{-1}\frac{1}{T}\sum_{t^*=1}^T\left[\tau-\mathbf{1}\left\{\epsilon_{it^*}\leq 0\right\}\right]\begin{bmatrix}v_{t^*,0}^0\\ v_{t^*,1}^0X_{1,it^*}\end{bmatrix}\\
&-\left(v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^T v_{t^*,0}^0\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\mu_{1,it^*}v_{t^*,1}^{0\prime}\right)\mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(h_i^{I,1}\right)^{-1}h_i^{I,2}\\
&=\mathcal{R}_{h,it}^1+X_{1,it}\mathcal{R}_{h,it}^2, \qquad\qquad \text{(B.84)}
\end{aligned}
$$

with $\max_{i\in I_3,t\in[T]}\left|\mathcal{R}_{h,it}^a\right|=O_p(\eta_N)$ for $a\in\{1,2\}$. Note that $\mathbb{E}\left[f_{it}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]=\mathbb{E}\left[f_{it}(0)\middle|\mathscr{D}\right]$.

Assumption 13 *Let $\mathcal{F}_{it}(\cdot)$ and $f_{it|h_{it}^{III}}(\cdot)$ be the conditional CDF and PDF of ϵ_{it} given $\mathscr{D}_e$ and h_{it}^{III}.*

(i) The derivative of the density $f_{it|h_{it}^{III}}$ is uniformly bounded in absolute value.

(ii) $\max_{i\in[N],t\in[T]}\left|f_{it|h_{it}^{III}}(0)-f_{it|h_{it}^{III}=0}(0)\right|\leq C\left|h_{it}^{III}\right|$ for some Lipschitz constant $C>0$.

Lemma B.30 *Under Assumptions 1-9 and Assumption 13, we have*

(i) $$\max_{t\in[T]}\left\|\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}O_{u,1}^{(1)}u_{i,1}^0\left\{\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{\epsilon_{it}\leq\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0,\left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right\}\right]\right.\right.$$
$$\left.\left.-\left(F_{it}(0)-F_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0,\left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right]\right)\right\}\right\|_2=o_p\left((N\vee T)^{-\frac{1}{2}}\right),$$

(ii) $$\max_{t\in[T]}\left\|\frac{1}{N_3}\sum_{i\in I_3}O_{u,0}^{(1)}u_{i,0}^0\left\{\left[\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{\epsilon_{it}\leq\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0,\left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right\}\right]\right.\right.$$
$$\left.\left.-\left(F_{it}(0)-F_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0,\left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right]\right)\right\}\right\|_2=O_p(\eta_N).$$

Proof. (i) Recall from (B.72) that

$$h_i^I = \mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(\frac{1}{T}\sum_{t=1}^T f_{it}(0)\begin{bmatrix} v_{t,0}^0 v_{t,0}' & v_{t,0}^0 v_{t,1}^{0\prime} X_{1,it} \\ v_{t,1}^0 v_{t,0}^{0\prime} X_{1,it} & v_{t,1}^0 v_{t,1}^{0\prime} X_{1,it}^2 \end{bmatrix}\right)^{-1}\left[\frac{1}{T}\sum_{t=1}^T [\tau - \mathbf{1}\{\epsilon_{it} \le 0\}]\begin{bmatrix} v_{t,0}^0 \\ v_{t,1}^0 X_{1,it}\end{bmatrix}\right.$$
$$\left. + \frac{1}{T}\sum_{t=1}^T \left[\mathbf{1}\{\epsilon_{it} \le 0\} - \mathbf{1}\left\{\epsilon_{it} \le \dot{\Delta}_{t,v}' \Psi_{it}^0\right\}\right]\begin{bmatrix} v_{t,0}^0 \\ v_{t,1}^0 X_{1,it}\end{bmatrix}\right].$$

By Bernstein's inequality in Lemma B.12, we can show that

$$\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T\left\{f_{it}(0)\begin{bmatrix} v_{t,0}^0 v_{t,0}' & v_{t,0}^0 v_{t,1}^{0\prime} X_{1,it} \\ v_{t,1}^0 v_{t,0}^{0\prime} X_{1,it} & v_{t,1}^0 v_{t,1}^{0\prime} X_{1,it}^2 \end{bmatrix} - \mathbb{E}\left(f_{it}(0)\begin{bmatrix} v_{t,0}^0 v_{t,0}' & v_{t,0}^0 v_{t,1}^{0\prime} X_{1,it} \\ v_{t,1}^0 v_{t,0}^{0\prime} X_{1,it} & v_{t,1}^0 v_{t,1}^{0\prime} X_{1,it}^2 \end{bmatrix}\middle|\mathscr{D}\right)\right\}\right\|_F$$
$$= O_p\left(\sqrt{\frac{\log(N\vee T)}{T}}\xi_N^2\right), \text{ and}$$
$$\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T [\tau - \mathbf{1}\{\epsilon_{it} \le 0\}]\begin{bmatrix} v_{t,0}^0 \\ v_{t,1}^0 X_{1,it}\end{bmatrix}\right\|_2 = O_p\left(\sqrt{\frac{\log(N\vee T)}{T}}\xi_N\right).$$

Besides, we observe that

$$\begin{aligned}
&\max_{i\in I_3, t\in[T]}\left\|Var\left(\mathbb{J}_{it}\left(\dot{\Delta}_{t,v}\right)\middle|\mathscr{D}^{I_1\cup I_2}\right)\right\|_F \\
&\le \max_{i\in I_3, t\in[T]}\left\|\mathbb{E}\left(\left[\mathbf{1}\{\epsilon_{it}\le 0\} - \mathbf{1}\left\{\epsilon_{it} \le \dot{\Delta}_{t,v}'\Psi_{it}^0\right\}\right]\begin{bmatrix} v_{t,0}^0 v_{t,0}' & v_{t,0}^0 v_{t,1}^{0\prime} X_{1,it} \\ v_{t,1}^0 v_{t,0}^{0\prime} X_{1,it} & v_{t,1}^0 v_{t,1}^{0\prime} X_{1,it}^2 \end{bmatrix}\middle|\mathscr{D}^{I_1\cup I_2}\right)\right\|_F \\
&= \max_{i\in I_3, t\in[T]}\left\|\mathbb{E}\left(\left[F_{it}(0) - F_{it}\left(\dot{\Delta}_{t,v}'\Psi_{it}^0\right)\right]\begin{bmatrix} v_{t,0}^0 v_{t,0}' & v_{t,0}^0 v_{t,1}^{0\prime} X_{1,it} \\ v_{t,1}^0 v_{t,0}^{0\prime} X_{1,it} & v_{t,1}^0 v_{t,1}^{0\prime} X_{1,it}^2 \end{bmatrix}\middle|\mathscr{D}^{I_1\cup I_2}\right)\right\|_F \\
&= \max_{t\in[T]}\left\|\dot{\Delta}_{t,v}\right\|_2 \max_{i\in I_3, t\in[T]}\left\|\mathbb{E}\left(f_{it}(\tilde{s}_{it})\Psi_{it}^0\begin{bmatrix} v_{t,0}^0 v_{t,0}' & v_{t,0}^0 v_{t,1}^{0\prime} X_{1,it} \\ v_{t,1}^0 v_{t,0}^{0\prime} X_{1,it} & v_{t,1}^0 v_{t,1}^{0\prime} X_{1,it}^2 \end{bmatrix}\middle|\mathscr{D}^{I_1\cup I_2}\right)\right\|_F \\
&= O_p(\eta_N),
\end{aligned} \tag{B.85}$$

where the first equality holds by law of iterated expectations, the second equality is by mean-value theorem for $|\tilde{s}_{it}|$ lies between 0 and $\left|\dot{\Delta}_{t,v}'\Psi_{it}^0\right|$ and the last line is by Theorem 3.2(ii) and Assumption 1(iv). Similarly, for any $\vartheta > 0$, we also have

$$\max_{i\in I_3, t\in[T]}\sum_{s=t+1}^T\left\|Cov\left(\mathbb{J}_{it}\left(\dot{\Delta}_{t,v}\right), \mathbb{J}_{is}\left(\dot{\Delta}_{t,v}\right)'\middle|\mathscr{D}^{I_1\cup I_2}\right)\right\|_F = O_p\left(\eta_N^{\frac{2}{2+\vartheta}}\right).$$

By similar arguments as used in (A.12) and (A.13), we have

$$\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t=1}^T\left\{\mathbb{J}_{it}\left(\dot{\Delta}_{t,v}\right) - \mathbb{E}\left(\mathbb{J}_{it}\left(\dot{\Delta}_{t,v}\right)\middle|\mathscr{D}^{I_1\cup I_2}\right)\right\}\right\|_2 = o_p\left((N\vee T)^{-1/2}\right). \tag{B.86}$$

Together with the fact that $\max_{i\in I_3}\left\|\mathbb{J}_i\left(\dot{\Delta}_{t,v}\right)\right\|_2 = O_p(\eta_N)$, uniformly over $t\in[T]$, we have

$$h_i^I = \mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(h_i^{I,1}\right)^{-1}\frac{1}{T}\sum_{t=1}^T[\tau - \mathbf{1}\{\epsilon_{it}\le 0\}]\begin{bmatrix} v_{t,0}^0 \\ v_{t,1}^0 X_{1,it}\end{bmatrix} + \mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(h_i^{I,1}\right)^{-1}h_i^{I,2} + o_p\left((N\vee T)^{-1/2}\right).$$

By Assumption 1(iv) and Lemma B.13, we have $\max_{i\in I_3}\left\|h_i^{I,1}\right\|_F = O(1)\ a.s.$. Like (B.85), we can show that $\max_{i\in I_3}\left\|h_i^{I,2}\right\|_2 = O_p(\eta_N)$.

Similarly as (B.82), uniformly in $i \in I_3$ and $t \in [T]$ we have

$$
\begin{aligned}
&\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)^{\prime -1} v_{t,0}^0, \left(O_{u,1}^{(1)}\right)^{\prime -1} v_{t,1}^0\right)\\
&= \hat{u}_{i,0}^{(3,1)\prime}\left(O_{u,0}^{(1)}\right)^{\prime -1} v_{t,0}^0 - u_{i,0}^{0\prime}v_{t,0}^0 + \hat{\mu}_{1,it}\hat{u}_{i,1}^{(3,1)\prime}\dot{v}_{t,1}^{(1)} - \mu_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\\
&+ \hat{e}_{1,it}\hat{u}_{i,1}^{(3,1)\prime}\left(O_{u,1}^{(1)}\right)^{\prime -1} v_{t,1}^0 - e_{1,it}u_{i,1}^{0\prime}v_{t,1}^0\\
&= v_{t,0}^{0\prime}O_{u,0}^{-1}\left(\hat{u}_{i,0}^{(3,1)} - O_{u,0}^{(1)}u_{i,0}^0\right) + X_{1,it}v_{t,1}^{0\prime}O_{u,1}^{(1)-1}\left(\hat{u}_{i,1}^{(3,1)} - O_{u,1}^{(1)}u_{i,1}^0\right) + \mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)^{\prime}\left(\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right)\\
&+ \mu_{1,it}u_{i,1}^{0\prime}O_{u,1}^{(1)\prime}\left(O_{u,1}^{(1)} - O_1^{(1)}\right)v_{t,1}^0 + o_p\left((N\vee T)^{-1/2}\right)\\
&= v_{t,0}^{0\prime}\left(O_0^{(1)}\right)^{-1}\left(\hat{u}_{i,0}^{(3,1)} - O_{u,0}^{(1)}u_{i,0}^0\right) + X_{1,it}v_{t,1}^{0\prime}\left(O_1^{(1)}\right)^{-1}\left(\hat{u}_{i,1}^{(3,1)} - O_{u,1}^{(1)}u_{i,1}^0\right) + \mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)^{\prime}\left(\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right)\\
&+ \mu_{1,it}u_{i,1}^{0\prime}O_1^{(1)\prime}\left(O_{u,1}^{(1)} - O_1^{(1)}\right)v_{t,1}^0 + o_p\left((N\vee T)^{-1/2}\right),
\end{aligned}
\tag{B.87}
$$

where the last line is by the fact that $\left\|O_{u,0}^{(1)} - O_0^{(1)}\right\|_F = O_p(\eta_N)$ and $\left\|O_{u,1}^{(1)} - O_1^{(1)}\right\|_F = O_p(\eta_N)$. Combining (A.41), (A.42) and Theorem 3.2(iii), we have

$$
\begin{aligned}
&\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)}, \hat{u}_{i,1}^{(3,1)}, \left(O_{u,0}^{(1)}\right)^{\prime -1} v_{t,0}^0, \left(O_{u,1}^{(1)}\right)^{\prime -1} v_{t,1}^0\right)\\
&= v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\left(\tau - 1\left\{\epsilon_{it^*}\le 0\right\}\right) + v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0 f_{it^*}(0)\left(\mu_{1,it^*}u_{i,1}^{0\prime}v_{t^*,1}^0 - \hat{\mu}_{1,it^*}^{(1)}\dot{u}_{i,1}^{(1)\prime}\dot{v}_{t^*,1}^{(1)}\right)\\
&+ v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t=1}^{T} f_{it^*}(0)v_{t^*,0}^0 v_{t^*,1}^{0\prime}u_{i,1}^0\left(e_{1,it^*} - \hat{e}_{1,it^*}\right) + X_{1,it}v_{t,1}^{0\prime}\hat{V}_{u_1}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} e_{1,it^*}v_{t^*,1}^0\left(\tau - 1\left\{\epsilon_{it^*}\le 0\right\}\right)\\
&+ \mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)^{\prime}\left(\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right) + \mu_{1,it}u_{i,1}^{0\prime}O_1^{(1)\prime}\left(O_{u,1}^{(1)} - O_1^{(1)}\right)v_{t,1}^0 + o_p\left((N\vee T)^{-1/2}\right)\\
&= v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\left(\tau - 1\left\{\epsilon_{it^*}\le 0\right\}\right) + X_{1,it}v_{t,1}^{0\prime}\hat{V}_{u_1}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} e_{1,it^*}v_{t^*,1}^0\left(\tau - 1\left\{\epsilon_{it^*}\le 0\right\}\right)\\
&- v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0 f_{it^*}(0)\mu_{1,it^*}\left(\dot{u}_{i,1}^{(1)} - O_1^{(1)}u_{i,1}^0\right)^{\prime}O_1^{(1)}v_{t^*,1}^0\\
&- v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0 f_{it^*}(0)\left(\hat{\mu}_{1,it^*} - \mu_{1,it^*}\right)u_{i,1}^{0\prime}v_{t^*,1}^0 - v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0 f_{it^*}(0)\mu_{1,it^*}\left(O_1^{(1)}u_{i,1}^0\right)^{\prime}\left(\dot{v}_{t^*,1}^{(1)} - O_1^{(1)}v_{t^*,1}^0\right)\\
&+ v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} f_{it^*}(0)v_{t^*,0}^0 v_{t^*,1}^{0\prime}u_{i,1}^0\left(e_{1,it^*} - \hat{e}_{1,it^*}\right) + \mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)^{\prime}\left(\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right)\\
&+ \mu_{1,it}u_{i,1}^{0\prime}O_1^{(1)\prime}\left(O_{u,1}^{(1)} - O_1^{(1)}\right)v_{t,1}^0 + o_p\left((N\vee T)^{-1/2}\right)\\
&= v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\left(\tau - 1\left\{\epsilon_{it^*}\le 0\right\}\right) + X_{1,it}v_{t,1}^{0\prime}\hat{V}_{u_1}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} e_{1,it^*}v_{t^*,1}^0\left(\tau - 1\left\{\epsilon_{it^*}\le 0\right\}\right)\\
&- v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\mu_{1,it^*}v_{t^*,1}^{0\prime}h_i^I + h_{it}^{IV} + o_p\left((N\vee T)^{-1/2}\right)
\end{aligned}
$$

$$
\begin{aligned}
&- v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\left(f_{it^*}(0)-\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\right)\mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t^*,1}^{(1)}-O_1^{(1)}v_{t^*,1}^0\right)\\
&- v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\left(f_{it^*}(0)-\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\right)\mu_{1,it^*}v_{t^*,1}^{0\prime}h_i^I\\
&:= h_{it}^{III}+h_{it}^{IV}+\mathcal{R}_{\varrho,it} \qquad \text{(B.88)}
\end{aligned}
$$

such that $\max_{i\in I_3,t\in[T]}|\mathcal{R}_{\varrho,it}| = O_p(\eta_N^2)$, where the first equality is by inserting the linear expansion for $\hat{u}_{i,0}^{(3,1)} - O_{u,0}^{(1)}u_{i,0}^0$ in (A.41) and $\hat{u}_{i,1}^{(3,1)} - O_{u,1}^{(1)}u_{i,1}^0$ in (A.42), the second equality is by inserting the linear expansion for $\dot{u}_{i,0}^{(1)} - O_{u,0}^{(1)}u_{i,0}^0$, the third equality is by the fact that $\hat{\mu}_{1,it}-\mu_{1,it} = e_{1,it}-\hat{e}_{1,it}$ which leads to the cancelling of the fourth and sixth terms in the second equality, the last equality is by the definition in (B.84) and

$$
\begin{aligned}
h_{it}^{IV} &= -v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\mu_{1,it^*}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t^*,1}^{(1)}-O_1^{(1)}v_{t^*,1}^0\right)\\
&\quad+\mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)}-O_1^{(1)}v_{t,1}^0\right)+\mu_{1,it}u_{i,1}^{0\prime}O_1^{(1)\prime}\left(O_{u,1}^{(1)}-O_1^{(1)}\right)v_{t,1}^0\\
&= \mathcal{R}_{h,it}^3+\mu_{1,it}\mathcal{R}_{h,it}^4
\end{aligned}
$$

with $\max_{i\in I_3,t\in[T]}\left|\mathcal{R}_{h,it}^a\right| = O_p(\eta_N)$ for $a\in\{3,4\}$, and the last equality holds by the fact that

$$
\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\left(f_{it^*}(0)-\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\right)\mu_{1,it^*}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t^*,1}^{(1)}-O_1^{(1)}v_{tv,1}^0\right)\right\|_2 = o_p\left((N\vee T)^{-1/2}\right)
$$

and

$$
\begin{aligned}
&\max_{i\in I_3}\left\|\frac{1}{T}\sum_{t^*=1}^{T} v_{t,0}^0\left(f_{it^*}(0)-\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\right)\mu_{1,it^*}v_{t^*,1}^{0\prime}h_i^I\right\|_2\\
&\le \max_{i\in I_3}\left|\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\left(f_{it^*}(0)-\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\right)\mu_{1,it^*}v_{t^*,1}^{0\prime}\right|\max_{i\in I_3}||h_i^I||_2 = o_p\left((N\vee T)^{-1/2}\right)
\end{aligned}
$$

by conditional Bernstein's inequality given $\mathscr{D}^{I_1\cup I_2}$ and similar arguments as used in (B.86).

We notice that

$$
\begin{aligned}
&\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0\Bigg\{\left[\mathbf{1}\{\epsilon_{it}\le 0\}-\mathbf{1}\left\{\epsilon_{it}\le\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0,\left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right\}\right]\\
&-\left(F_{it}(0)-F_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\left(O_{u,0}^{(1)}\right)'^{-1}v_{t,0}^0,\left(O_{u,1}^{(1)}\right)'^{-1}v_{t,1}^0\right)\right]\right)\Bigg\}\\
&=\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0\left\{\left[\mathbf{1}\{\epsilon_{it}\le 0\}-\mathbf{1}\left\{\epsilon_{it}\le h_{it}^{III}+h_{it}^{IV}\right\}\right]-\left(F_{it}(0)-F_{it}\left[h_{it}^{III}+h_{it}^{IV}\right]\right)\right\}\\
&+\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le h_{it}^{III}+h_{it}^{IV}\right\}-\mathbf{1}\left\{\epsilon_{it}\le h_{it}^{III}+h_{it}^{IV}+\mathcal{R}_{\varrho,it}\right\}\right]\\
&-\left[F_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)-F_{it}\left(h_{it}^{III}+h_{it}^{IV}+\mathcal{R}_{\varrho,it}\right)\right]\Bigg\}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \left\{ \left[\mathbf{1}\left\{ \epsilon_{it} \leq 0 \right\} - \mathbf{1}\left\{ \epsilon_{it} \leq h_{it}^{III} + h_{it}^{IV} \right\} \right] - \left[\mathcal{F}_{it}(0) - \mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) \right] \right\} \\
&+ \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \Bigg\{ \left[\mathbf{1}\left\{ \epsilon_{it} \leq h_{it}^{III} + h_{it}^{IV} \right\} - \mathbf{1}\left\{ \epsilon_{it} \leq h_{it}^{III} + h_{it}^{IV} + \mathcal{R}_{\varrho,it} \right\} \right] \\
&- \left[\mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) - \mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV} + \mathcal{R}_{\varrho,it} \right) \right] \Bigg\} \\
&+ \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \left\{ \left[\mathcal{F}_{it}(0) - \mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) \right] - \left[F_{it}(0) - F_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) \right] \right\} \\
&+ \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \left[\mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) - \mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV} + \mathcal{R}_{\varrho,it} \right) \right] \\
&- \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \left[F_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) - F_{it}\left(h_{it}^{III} + h_{it}^{IV} + \mathcal{R}_{\varrho,it} \right) \right] \\
&:= \mathbb{I}_{6,t}^{I} + \mathbb{I}_{6,t}^{II} - \mathbb{I}_{6,t}^{III} + \mathbb{I}_{6,t}^{IV} + o_p\left((N \vee T)^{-\frac{1}{2}} \right) \qquad \text{(B.89)}
\end{aligned}
$$

such that $\mathbb{I}_{6,t}^{I} := \frac{1}{N_3} \sum_{i \in I_3} \mathbb{I}_{6,it}^{I}$, $\mathbb{I}_{6,t}^{II} := \frac{1}{N_3} \sum_{i \in I_3} \mathbb{I}_{6,it}^{II}$, $\mathbb{I}_{6,t}^{III} := \frac{1}{N_3} \sum_{i \in I_3} \mathbb{I}_{6,it}^{III}$, $\mathbb{I}_{6,t}^{IV} := \frac{1}{N_3} \sum_{i \in I_3} \mathbb{I}_{6,it}^{IV}$, with

$$
\begin{aligned}
\mathbb{I}_{6,it}^{I} &= e_{1,it} u_{i,1}^0 \left\{ \left[\mathbf{1}\left\{ \epsilon_{it} \leq 0 \right\} - \mathbf{1}\left\{ \epsilon_{it} \leq h_{it}^{III} + h_{it}^{IV} \right\} \right] - \left[\mathcal{F}_{it}(0) - \mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) \right] \right\}, \\
\mathbb{I}_{6,it}^{II} &= e_{1,it} u_{i,1}^0 \Bigg\{ \left[\mathbf{1}\left\{ \epsilon_{it} \leq h_{it}^{III} + h_{it}^{IV} \right\} - \mathbf{1}\left\{ \epsilon_{it} \leq h_{it}^{III} + h_{it}^{IV} + \mathcal{R}_{\varrho,it} \right\} \right] \\
&\quad - \left[\mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) - \mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV} + \mathcal{R}_{\varrho,it} \right) \right] \Bigg\}, \qquad \text{(B.90)} \\
\mathbb{I}_{6,it}^{III} &= e_{1,it} u_{i,1}^0 \left[F_{it}(0) - F_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) \right], \\
\mathbb{I}_{6,it}^{IV} &= e_{1,it} u_{i,1}^0 \left[\mathcal{F}_{it}(0) - \mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) \right].
\end{aligned}
$$

The last line in (B.89) is due to the fact that the last two terms in the second equality is $o_p\left((N \vee T)^{-\frac{1}{2}} \right)$ by mean-value theorem, Assumption (1)(iv), and Assumption (1)(viii) and the union bound for $\mathcal{R}_{\varrho,it}$.

For $\mathbb{I}_{6,t}^{I}$ and $\mathbb{I}_{6,t}^{II}$, conditional on $\mathscr{D}_e^{I_1 \cup I_2}$ and h_{it}^{III}, the randomness is from ϵ_{it}, and we observe that $\mathbb{I}_{6,it}^{I}$ and $\mathbb{I}_{6,it}^{II}$ are independent over i by conditioning on $\mathscr{D}_e^{I_1 \cup I_2}$ and h_{it}^{III}. Therefore, we obtain that $\max_{t \in [T]} \left\| \mathbb{I}_{6,t}^{I} \right\|_2 = o_p\left((N \vee T)^{-\frac{1}{2}} \right)$ and $\max_{t \in [T]} \left\| \mathbb{I}_{6,t}^{II} \right\|_2 = o_p\left((N \vee T)^{-\frac{1}{2}} \right)$ by conditional Bernstein's inequality for independent sequence given $\mathscr{D}_e^{I_1 \cup I_2}$ and h_{it}^{III}.

For $\mathbb{I}_{6,t}^{III}$, we notice that

$$
\begin{aligned}
\mathbb{I}_{6,t}^{III} &= \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \left[F_{it}(0) - F_{it}\left(h_{it}^{III} + h_{it}^{IV} \right) \right] \\
&= \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 f_{it}(\dot{s}_{it}) \left(h_{it}^{III} + h_{it}^{IV} \right) \\
&= \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 f_{it}(0) \left(h_{it}^{III} + h_{it}^{IV} \right) + \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 \left[f_{it}(\dot{s}_{it}) - f_{it}(0) \right] \left(h_{it}^{III} + h_{it}^{IV} \right) \\
&= \frac{1}{N_3} \sum_{i \in I_3} e_{1,it} u_{i,1}^0 f_{it}(0) \left(h_{it}^{III} + h_{it}^{IV} \right) + o_p\left((N \vee T)^{-\frac{1}{2}} \right) \quad \text{uniformly,} \qquad \text{(B.91)}
\end{aligned}
$$

where the second line is by mean-value theorem with $|\dot{s}_{it}|$ lies between 0 and $\left|h_{it}^{III}+h_{it}^{IV}\right|$ and the last line is by Assumption 1(viii). By inserting h_{it}^{III} and h_{it}^{IV}, we have

$$
\begin{aligned}
&\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)\left(h_{it}^{III}+h_{it}^{IV}\right)\\
&=\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\left(\tau-1\left\{\epsilon_{it^*}\leq 0\right\}\right)\\
&+\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)X_{1,it}v_{t,1}^{0\prime}\hat{V}_{u_1}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} e_{1,it^*}v_{t^*,1}^0\left(\tau-1\left\{\epsilon_{it^*}\leq 0\right\}\right)\\
&-\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)\left(v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0 f_{it^*}(0)\mu_{1,it^*}v_{t^*,1}^{0\prime}\right)\mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(h_i^{I,1}\right)^{-1}\frac{1}{T}\sum_{t^*=1}^{T}\left[\tau-\mathbf{1}\left\{\epsilon_{it^*}\leq 0\right\}\right]\begin{bmatrix} v_{t^*,0}^0\\ v_{t^*,1}^0X_{1,it^*}\end{bmatrix}\\
&-\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)\left(v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\mu_{1,it^*}v_{t^*,1}^{0\prime}\right)\mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(h_i^{I,1}\right)^{-1}h_i^{I,2}\\
&-\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t^*=1}^{T} v_{t^*,0}^0\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\mu_{1,it^*}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t^*,1}^{(1)}-O_1^{(1)}v_{t^*,1}^0\right)\\
&+\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)\mu_{1,it}\left(O_1^{(1)}u_{i,1}^0\right)'\left(\dot{v}_{t,1}^{(1)}-O_1^{(1)}v_{t,1}^0\right)\\
&+\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)\mu_{1,it}u_{i,1}^{0\prime}O_1^{(1)\prime}\left(O_{u,1}^{(1)}-O_1^{(1)}\right)v_{t,1}^0\\
&:=\sum_{m\in[7]}\mathbb{I}_{6,t}^{III,m}. \qquad \text{(B.92)}
\end{aligned}
$$

$\mathbb{I}_{6,t}^{III,1}$, $\mathbb{I}_{6,t}^{III,2}$, $\mathbb{I}_{6,t}^{III,3}$ can be analyzed in the same manner, and we take $\mathbb{I}_{6,t}^{III,1}$ for instance. Noticed that

$$
\mathbb{I}_{6,t}^{III,1}=\frac{1}{N_3T}\sum_{i\in I_3}\sum_{t^*=1}^{T} e_{1,it}u_{i,1}^0 f_{it}(0)v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}v_{t^*,0}^0\left(\tau-1\left\{\epsilon_{it^*}\leq 0\right\}\right),
$$

it is clear that conditioning on $\mathscr{D}_e$, $\mathbb{I}_{6,t}^{III,1}$ is mean zero by Assumption 1(ii) and the randomness is from ϵ_{it} which is strong mixing. With the similar arguments as the second term in (B.65), we obtain that $\max_{t\in[T]}\left\|\mathbb{I}_{6,t}^{III,1}\right\|_2=O_p\left(\sqrt{\frac{\log(N\vee T)}{NT}}\xi_N\right)$, and by Assumption 1(ix), it follows that

$$
\max_{t\in[T]}\left\|\mathbb{I}_{6,t}^{III,1}\right\|_2=o_p\left((N\vee T)^{-1/2}\right).
$$

We can also analyze $\mathbb{I}_{6,t}^{III,4}$ and $\mathbb{I}_{6,t}^{III,5}$ in the same manner. Take $\mathbb{I}_{6,t}^{III,4}$ for instance. Note that

$$
-\mathbb{I}_{6,t}^{III,4}=\frac{1}{N_3T}\sum_{i\in I_3}\sum_{t^*=1}^{T} e_{1,it}u_{i,1}^0 f_{it}(0)\left(v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}v_{t,0}^0\mathbb{E}\left[f_{it^*}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\mu_{1,it^*}v_{t^*,1}^{0\prime}\right)\mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(h_i^{I,1}\right)^{-1}h_i^{I,2}
$$

which is mean zero by conditioning on $\mathscr{D}^{I_1\cup I_2}$ owing to Assumption 9(i) and the fact that $h_i^{I,1}$ and $h_i^{I,2}$ are fixed given $\mathscr{D}^{I_1\cup I_2}$. Then by similar arguments for $\mathbb{I}_{6,t}^{III,1}$ above, we have

$$
\max_{t\in[T]}\left\|\mathbb{I}_{6,t}^{III,4}\right\|=o_p\left((N\vee T)^{-1/2}\right). \qquad \text{(B.93)}
$$

For $\mathbb{I}_{6,t}^{III,6}$ and $\mathbb{I}_{6,t}^{III,7}$, we have

$$\begin{aligned}
\max_{t\in[T]}\left\|\mathbb{I}_{6,t}^{III,6}\right\|_2 &\leq \max_{t\in[T]}\left\|\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)\mu_{1,it}u_{i,1}^{0\prime}\right\|_2 \max_{t\in[T]}\left\|\dot{v}_{t,1}^{(1)} - O_1^{(1)}v_{t,1}^0\right\|_2 \\
&\leq \max_{t\in[T]}\left\|\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)\mu_{1,it}u_{i,1}^{0\prime}\right\|_2 O_p(\eta_N) \\
&\leq O_p\left(\sqrt{\frac{\log(N\vee T)}{N}}\xi_N\right) O_p(\eta_N) = o_p\left((N\vee T)^{-1/2}\right), \qquad \text{(B.94)} \\
\max_{t\in[T]}\left\|\mathbb{I}_{6,t}^{III,7}\right\|_2 &\leq \max_{t\in[T]}\left\|\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it}(0)\mu_{1,it}u_{i,1}^{0\prime}\right\|_2 \max_{t\in[T]}\left\|v_{t,1}^0\right\|_2 \left\|O_{u,1}^{(1)} - O_1^{(1)}\right\|_F \\
&\leq O_p\left(\sqrt{\frac{\log(N\vee T)}{N}}\xi_N\right) O_p(\eta_N) = o_p\left((N\vee T)^{-1/2}\right), \qquad \text{(B.95)}
\end{aligned}$$

where the second inequality is by Theorem 3.2(ii) and the third inequality is by Hoeffding's inequality conditional on $\mathscr{D}$ and the last line combines Lemma B.13(i) and the fact that $\left\|O_{u,1}^{(1)} - O_1^{(1)}\right\|_F = O_p(\eta_N)$. Combining (B.91)-(B.94), we have $\max_{t\in[T]}\left\|\mathbb{I}_{6,it}^{III}\right\|_2 = o_p\left((N\vee T)^{-1/2}\right)$.

Last, we analyze $\mathbb{I}_{6,t}^{IV}$. Like (B.91), we have

$$\begin{aligned}
\mathbb{I}_{6,t}^{IV} &= \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0\left[\mathcal{F}_{it}(0) - \mathcal{F}_{it}\left(h_{it}^{III} + h_{it}^{IV}\right)\right] \\
&= \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it|h_{it}^{III}}(\dot{s}_{it})\left(h_{it}^{III} + h_{it}^{IV}\right) \\
&= \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it|h_{it}^{III}}(0)\left(h_{it}^{III} + h_{it}^{IV}\right) + \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0\left[f_{it|h_{it}^{III}}(\dot{s}_{it}) - f_{it|h_{it}^{III}}(0)\right]\left(h_{it}^{III} + h_{it}^{IV}\right) \\
&= \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it|h_{it}^{III}}(0)h_{it}^{III} + \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it|h_{it}^{III}}(0)h_{it}^{IV} + o_p\left((N\vee T)^{-\frac{1}{2}}\right) \\
&= \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it|h_{it}^{III}}(0)\left[h_{it}^{III} - \mathbb{E}\left(h_{it}^{III}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\right] + \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it|h_{it}^{III}}(0)\mathbb{E}\left(h_{it}^{III}\middle|\mathscr{D}_e^{I_1\cup I_2}\right) \\
&+ \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 h_{it}^{IV}\left[f_{it|h_{it}^{III}}(0) - \mathbb{E}\left(f_{it|h_{it}^{III}}(0)|\mathscr{D}_e^{I_1\cup I_2}\right)\right] + \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 h_{it}^{IV}\mathbb{E}\left(f_{it|h_{it}^{III}}(0)|\mathscr{D}_e^{I_1\cup I_2}\right) \\
&+ o_p\left((N\vee T)^{-\frac{1}{2}}\right) \\
&= \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0\left[f_{it|h_{it}^{III}}(0) - f_{it|h_{it}^{III}=0}(0)\right]\left[h_{it}^{III} - \mathbb{E}\left(h_{it}^{III}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\right] \\
&+ \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 f_{it|h_{it}^{III}=0}(0)\left[h_{it}^{III} - \mathbb{E}\left(h_{it}^{III}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\right] \\
&+ \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0\mathbb{E}\left(h_{it}^{III}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\left[f_{it|h_{it}^{III}}(0) - \mathbb{E}\left(f_{it|h_{it}^{III}}(0)|\mathscr{D}_e^{I_1\cup I_2}\right)\right] \\
&+ \frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0\mathbb{E}\left(h_{it}^{III}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\mathbb{E}\left(f_{it|h_{it}^{III}}(0)|\mathscr{D}_e^{I_1\cup I_2}\right)
\end{aligned}$$

$$+\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 h_{it}^{IV}\left[f_{it|h_{it}^{III}}(0)-\mathbb{E}\left(f_{it|h_{it}^{III}}(0)|\mathscr{D}_e^{I_1\cup I_2}\right)\right]+\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}u_{i,1}^0 h_{it}^{IV}\mathbb{E}\left(f_{it|h_{it}^{III}}(0)|\mathscr{D}_e^{I_1\cup I_2}\right)$$
$$+o_p\left((N\vee T)^{-\frac{1}{2}}\right)$$
$$:=\sum_{m\in[6]}\mathbb{I}_{6,t}^{IV,m}+o_p\left((N\vee T)^{-\frac{1}{2}}\right), \tag{B.96}$$

where all the $o_p\left((N\vee T)^{-\frac{1}{2}}\right)$ terms hold uniformly over $t\in[T]$ and $o_p\left((N\vee T)^{-\frac{1}{2}}\right)$ term in the fourth equality is by mean-value theorem, Assumption 13(i) and the fact that $|\dot{s}_{it}|$ lies between 0 and $\left|h_{it}^{III}+h_{it}^{IV}\right|$.

For $\mathbb{I}_{6,t}^{IV,1}$, with Assumption 13(ii), (B.84) and the fact that

$$\begin{aligned}
&\max_{i\in I_3,t\in[T]}\left|\mathbb{E}\left(h_{it}^{III}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\right| \\
&=\max_{i\in I_3,t\in[T]}\left|\mathbb{E}\left[\left(v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t=1}^{T}v_{t,0}^0\mathbb{E}\left[f_{it}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\mu_{1,it}v_{t,1}^{0\prime}\right)\mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(h_i^{I,1}\right)^{-1}h_i^{I,2}\middle|\mathscr{D}_e^{I_1\cup I_2}\right]\right| \\
&=\max_{i\in I_3,t\in[T]}\left|\left(v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t=1}^{T}v_{t,0}^0\mathbb{E}\left[f_{it}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\mu_{1,it}v_{t,1}^{0\prime}\right)\mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(h_i^{I,1}\right)^{-1}h_i^{I,2}\right| \\
&=O_p(\eta_N),
\end{aligned}$$

we have

$$\max_{t\in[T]}\left\|\mathbb{I}_{6,t}^{IV,1}\right\|_2\lesssim\max_{t\in[T]}\frac{1}{N_3}\sum_{i\in I_3}\left\|e_{1,it}u_{i,1}^0\right\|_2\left|h_{it}^{III}\right|\left(\left|h_{it}^{III}\right|+\left|\mathbb{E}\left(h_{it}^{III}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)\right|\right)=O_p(\eta_N^2). \tag{B.97}$$

For $\mathbb{I}_{6,t}^{IV,2}$, $\mathbb{I}_{6,t}^{IV,3}$ and $\mathbb{I}_{6,t}^{IV,5}$, conditioning on $\mathscr{D}_e^{I_1\cup I_2}$, the randomness is only from h_{it}^{III}, which is independent across i, and $\mathbb{I}_{6,t}^{IV,2}$, $\mathbb{I}_{6,t}^{IV,3}$ and $\mathbb{I}_{6,t}^{IV,5}$ are zero mean by conditioning on $\mathscr{D}_e^{I_1\cup I_2}$. Similar to the arguments for $\mathbb{I}_{6,t}^{I}$ and $\mathbb{I}_{6,t}^{II}$ in (B.90), we have

$$\max_{t\in[T]}\left\|\mathbb{I}_{6,t}^{IV,m}\right\|_2=o_p\left((N\vee T)^{-\frac{1}{2}}\right), m\in\{2,3,5\}. \tag{B.98}$$

For $\mathbb{I}_{6,t}^{IV,4}$, by inserting $\mathbb{E}\left(h_{it}^{III}\middle|\mathscr{D}_e^{I_1\cup I_2}\right)$ and the fact that $\mathbb{E}\left(f_{it|h_{it}^{III}}(0)|\mathscr{D}_e^{I_1\cup I_2}\right)=f_{it}(0)$, it yields

$$\begin{aligned}
\mathbb{I}_{6,t}^{IV,4}&=-\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}f_{it}(0)u_{i,1}^0\left(v_{t,0}^{0\prime}\hat{V}_{u_0,i}^{-1}\frac{1}{T}\sum_{t=1}^{T}v_{t,0}^0\mathbb{E}\left[f_{it}(0)\middle|\mathscr{D}^{I_1\cup I_2}\right]\mu_{1,it}v_{t,1}^{0\prime}\right)\mathfrak{l}'\left(O_1^{(1)}\right)^{-1}\left(h_i^{I,1}\right)^{-1}h_i^{I,2} \\
&=\mathbb{I}_{6,t}^{III,4}=o_p\left((N\vee T)^{-1/2}\right)\quad\text{uniformly,}
\end{aligned} \tag{B.99}$$

where the last equality is by (B.93).

For $\mathbb{I}_{6,t}^{IV,6}$, we notice that

$$\begin{aligned}
\mathbb{I}_{6,t}^{IV,6}&=\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}u_{i,1}^0h_{it}^{IV}\mathbb{E}\left(f_{it|h_{it}^{III}}(0)|\mathscr{D}_e^{I_1\cup I_2}\right) \\
&=\frac{1}{N_3}\sum_{i\in I_3}e_{1,it}f_{it}(0)u_{i,1}^0h_{it}^{IV}=\sum_{m\in\{5,6,7\}}\mathbb{I}_{6,t}^{III,m}=o_p\left((N\vee T)^{-1/2}\right)\quad\text{uniformly.}
\end{aligned} \tag{B.100}$$

Combining (B.96)-(B.100) yields $\max_{t\in[T]}\left\|\mathbb{I}_{6,it}^{III}\right\|_2 = o_p\left((N\vee T)^{-1/2}\right)$, which leads to the desired result in statement (i).

(ii) As in (B.89), we have

$$
\begin{aligned}
&\frac{1}{N_3}\sum_{i\in I_3} O_{u,0}^{(1)}u_{i,0}^{0}\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le \varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\left(O_{u,0}^{(1)}\right)^{\prime -1}v_{t,0}^{0},\left(O_{u,1}^{(1)}\right)^{\prime -1}v_{t,1}^{0}\right)\right\}\right]\\
&-\left(F_{it}(0)-F_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\left(O_{u,0}^{(1)}\right)^{\prime -1}v_{t,0}^{0},\left(O_{u,1}^{(1)}\right)^{\prime -1}v_{t,1}^{0}\right)\right]\right)\Bigg\}\\
&=\frac{1}{N_3}\sum_{i\in I_3} u_{i,1}^{0}\left\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le h_{it}^{III}+h_{it}^{IV}\right\}\right]-\left[\mathcal{F}_{it}(0)-\mathcal{F}_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)\right]\right\}\\
&+\frac{1}{N_3}\sum_{i\in I_3} u_{i,1}^{0}\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le h_{it}^{III}+h_{it}^{IV}\right\}-\mathbf{1}\left\{\epsilon_{it}\le h_{it}^{III}+h_{it}^{IV}+\mathcal{R}_{\varrho,it}\right\}\right]\\
&-\left[\mathcal{F}_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)-\mathcal{F}_{it}\left(h_{it}^{III}+h_{it}^{IV}+\mathcal{R}_{\varrho,it}\right)\right]\Bigg\}\\
&+\frac{1}{N_3}\sum_{i\in I_3} u_{i,1}^{0}\left\{\left[\mathcal{F}_{it}(0)-\mathcal{F}_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)\right]-\left[F_{it}(0)-F_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)\right]\right\}\\
&+\frac{1}{N_3}\sum_{i\in I_3} u_{i,1}^{0}\left[\mathcal{F}_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)-\mathcal{F}_{it}\left(h_{it}^{III}+h_{it}^{IV}+\mathcal{R}_{\varrho,it}\right)\right]\\
&-\frac{1}{N_3}\sum_{i\in I_3} u_{i,1}^{0}\left[F_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)-F_{it}\left(h_{it}^{III}+h_{it}^{IV}+\mathcal{R}_{\varrho,it}\right)\right]\\
&=\frac{1}{N_3}\sum_{i\in I_3} u_{i,1}^{0}\left\{\left[\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le h_{it}^{III}+h_{it}^{IV}\right\}\right]-\left[\mathcal{F}_{it}(0)-\mathcal{F}_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)\right]\right\}\\
&+\frac{1}{N_3}\sum_{i\in I_3} u_{i,1}^{0}\Bigg\{\left[\mathbf{1}\left\{\epsilon_{it}\le h_{it}^{III}+h_{it}^{IV}\right\}-\mathbf{1}\left\{\epsilon_{it}\le h_{it}^{III}+h_{it}^{IV}+\mathcal{R}_{\varrho,it}\right\}\right]\\
&-\left[\mathcal{F}_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)-\mathcal{F}_{it}\left(h_{it}^{III}+h_{it}^{IV}+\mathcal{R}_{\varrho,it}\right)\right]\Bigg\}\\
&+\frac{1}{N_3}\sum_{i\in I_3} u_{i,1}^{0}\left\{\left[\mathcal{F}_{it}(0)-\mathcal{F}_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)\right]-\left[F_{it}(0)-F_{it}\left(h_{it}^{III}+h_{it}^{IV}\right)\right]\right\}+o_p\left((N\vee T)^{-1/2}\right),
\end{aligned}
$$

where the last equality holds by similar arguments as used in the last line in (B.89). We can show the first and second terms are $o_p\left((N\vee T)^{-1/2}\right)$ by similar arguments for $\mathbb{I}_{6,t}^{I}$ and $\mathbb{I}_{6,t}^{II}$. The third term is $O_p(\eta_N)$ by mean-value theorem and Assumption 1(viii). Compared with $\mathbb{I}_{6,t}^{III}$ and $\mathbb{I}_{6,t}^{IV}$ in the proof of statement (i), the third term here is not mean zero, and converges to zero at the rate η_N. ■

Lemma B.31 *Under Assumptions 1-9 and Assumption 13, we have*

(i) $\max_{t\in[T]}\Bigg\|\frac{1}{N_3}\sum_{i\in I_3} e_{1,it}O_{u,1}^{(1)}u_{i,1}^{0}\Bigg\{\mathbf{1}\left\{\epsilon_{it}\le 0\right\}-\mathbf{1}\left\{\epsilon_{it}\le \varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\hat{v}_{t,0}^{(3,1)},\hat{v}_{t,1}^{(3,1)}\right)\right\}$
$-\left(F_{it}(0)-F_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\hat{v}_{t,0}^{(3,1)},\hat{v}_{t,1}^{(3,1)}\right)\right]\right)\Bigg\}\Bigg\|_2 = o_p\left((N\vee T)^{-\frac{1}{2}}\right),$

(ii) $\max_{t\in[T]}\left\|\frac{1}{N_3}\sum_{i\in I_3} O_{u,0}^{(1)}u_{i,0}^{0}\left\{\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}-\mathbf{1}\left\{\epsilon_{it}\leq \varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\hat{v}_{t,0}^{(3,1)},\hat{v}_{t,1}^{(3,1)}\right)\right\}\right.\right.$
$\left.\left.-\left(F_{it}(0)-F_{it}\left[\varrho_{it}\left(\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)},\hat{v}_{t,0}^{(3,1)},\hat{v}_{t,1}^{(3,1)}\right)\right]\right)\right\}\right\|_2=O_p(\eta_N).$

Proof. To handle the correlation between $\{\epsilon_{it},e_{it}\}$ and $\left\{\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)}\right\}$, we follow similar arguments as used in the proof of Lemma B.27 by putting $\left\{\hat{u}_{i,0}^{(3,1)},\hat{u}_{i,1}^{(3,1)}\right\}$ in a parameter set. Then by similar arguments as used in the proof of Lemma B.30, we can obtain the desired results. ■

Lemma B.32 *Under Assumptions 1-9 and Assumption 13, we have*

$$\max_{t\in[T]}\left\|O_{v_1,t}^{(1)}-\left(O_{u,1}^{(1)\prime}\right)^{-1}\right\|_F=o_p\left((N\vee T)^{-\frac{1}{2}}\right).$$

Proof. Recall that

$$O_{v_1,t}^{(1)}=\left\{I_{K_1}+\left(O_{u,1}^{(1)\prime}\right)^{-1}\left[\hat{V}_{v_1,t}^{I}\right]^{-1}\left[\frac{1}{N_3}\sum_{i\in I_3}f_{it}(0)e_{1,it}^2u_{i,1}^0\left(O_{u,1}^{(1)}u_{i,1}^0-\hat{u}_{i,1}^{(3,1)}\right)'\right]\right\}\left(O_{u,1}^{(1)\prime}\right)^{-1}.$$

Then

$$O_{v_1,t}^{(1)}-\left(O_{u,1}^{(1)\prime}\right)^{-1}=\left(O_{u,1}^{(1)\prime}\right)^{-1}\left[\hat{V}_{v_1,t}^{3}\right]^{-1}\left[\frac{1}{N_3}\sum_{i\in I_3}f_{it}(0)e_{1,it}^2u_{i,1}^0\left(O_{u,1}^{(1)}u_{i,1}^0-\hat{u}_{i,1}^{(3,1)}\right)'\right].$$

Note that

$$\begin{aligned}
&\frac{1}{N_3}\sum_{i\in I_3}f_{it}(0)e_{1,it}^2u_{i,1}^0\left(O_{u,1}^{(1)}u_{i,1}^0-\hat{u}_{i,1}^{(3,1)}\right)'\\
&=\frac{1}{N_3}\sum_{i\in I_3}f_{it}(0)e_{1,it}^2\left\{O_1^{(1)}\hat{V}_{u_1}^{-1}\frac{1}{T}\sum_{t=1}^{T}e_{1,it}v_{t,1}^0\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)+\mathcal{R}_{i,u}^1\right\}u_{i,1}^{0\prime}\\
&=\frac{1}{N_3}\sum_{i\in I_3}f_{it}(0)e_{1,it}^2O_1^{(1)}\hat{V}_{u_1}^{-1}\frac{1}{T}\sum_{t=1}^{T}e_{1,it}v_{t,1}^0\left(\tau-\mathbf{1}\left\{\epsilon_{it}\leq 0\right\}\right)u_{i,1}^{0\prime}+o_p\left((N\vee T)^{-\frac{1}{2}}\right)\\
&=\frac{1}{N_3T}\sum_{i\in I_3}\sum_{t=1'}^{T}f_{it}(0)e_{1,it}^2O_1^{(1)}\hat{V}_{u_1}^{-1}e_{1,it^*}v_{t^*,1}^0\left(\tau-1\left\{\epsilon_{it^*}\leq 0\right\}\right)u_{i,1}^{0\prime}+o_p\left((N\vee T)^{-\frac{1}{2}}\right)\\
&=o_p\left((N\vee T)^{-\frac{1}{2}}\right)\quad\text{uniformly over } t\in[T],
\end{aligned}$$

where the second equality is by uniform convergence rate of $\mathcal{R}_{i,u}^1$ and the last line follows by similar arguments as in (B.65) by Bernstein's inequality conditional on $\mathscr{D}_e$. Then the result follows by noting that $O_{u,1}$ is bounded and $\hat{V}_{v_1,t}^{I}$ is bounded uniformly over $t\in[T]$. ■

B.4 Lemmas for the Consistent Estimation of the Asymptotic Variances

Recall that

$$\hat{\mathbb{V}}_{u_j}=\frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]}k_{h_N}(\hat{\epsilon}_{it})\hat{e}_{j,it}^2\hat{\mathrm{v}}_{t,t,j},\quad \hat{\mathbb{V}}_{v_j}=\frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]}k_{h_N}(\hat{\epsilon}_{it})\hat{e}_{j,it}^2\hat{\mathrm{u}}_{i,i,j},$$

$$
\begin{aligned}
\hat{\Omega}_{u_j} &= \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]}\tau(1-\tau)\hat{e}_{j,it}^2\hat{\mathrm{v}}_{t,t,j} \\
&+ \frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}\left[\tau - K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)\right]\left[\tau - K\left(\frac{\hat{\epsilon}_{is}}{h_N}\right)\right] \\
&+ \frac{1}{NT}\sum_{i\in[N]}\sum_{t=1+T_1}^{T}\sum_{s=t-T_1}^{t-1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}\left[\tau - K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)\right]\left[\tau - K\left(\frac{\hat{\epsilon}_{is}}{h_N}\right)\right], \\
\hat{\Omega}_{v_j} &= \tau(1-\tau)\frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]}\hat{e}_{j,it}^2\hat{\mathrm{u}}_{i,i,j}, \\
\hat{\Sigma}_{u_j} &= \left(\hat{\mathbb{V}}_{u_j}\right)^{-1}\hat{\Omega}_{u_j}\left(\hat{\mathbb{V}}_{u_j}\right)^{-1}, \quad \hat{\Sigma}_{v_j} = \left(\hat{\mathbb{V}}_{v_j}\right)^{-1}\hat{\Omega}_{v_j}\left(\hat{\mathbb{V}}_{v_j}\right)^{-1}, \\
V_{u_j} &= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left[f_{it}(0)e_{j,it}^2 v_{t,j}^0 v_{t,j}^{0\prime}\right], \quad V_{v_j}^{(a)} = \frac{1}{N_a}\sum_{i\in I_a}\mathbb{E}\left[f_{it}(0)e_{j,it}^2\right]u_{i,j}^0 u_{i,j}^{0\prime}, \\
\Omega_{u_j} &= Var\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}e_{j,it}v_{t,j}^0(\tau - \mathbf{1}\{\epsilon_{it}\le 0\})\right], \quad \Omega_{v_j} = \tau(1-\tau)\frac{1}{N}\sum_{i\in I_3}\mathbb{E}\left(e_{j,it}^2 u_{i,j}^0 u_{i,j}^{0\prime}\right), \\
\Sigma_{u_j} &= O_j^{(1)}V_{u_j}^{-1}\Omega_{u_j}V_{u_j}^{-1}O_j^{(1)\prime}, \quad \Sigma_{v_j} = O_j^{(1)}V_{v_j}^{-1}\Omega_{v_j}V_{v_j}^{-1}O_j^{(1)\prime}.
\end{aligned}
$$

Lemma B.33 *Under Assumptions 1-10 and Assumption 13,* $\hat{\Sigma}_{u_j} = \Sigma_{u_j} + o_p(1)$ *and* $\hat{\Sigma}_{v_j} = \Sigma_{v_j} + o_p(1)$.

Proof. First, we show that $\hat{\mathbb{V}}_{u_j} = O_j^{(1)}V_{u_j}O_j^{(1)\prime} + o_p(1)$. Note that

$$
\begin{aligned}
\max_{i\in I_3,t\in[T]}|\hat{\epsilon}_{it} - \epsilon_{it}| &\le \max_{i\in I_3,t\in[T]}\left|\hat{\Theta}_{0,it} - \Theta_{0,it}^0\right| + \max_{i\in I_3,t\in[T]}\sum_{j\in[p]}|X_{j,it}|\left|\hat{\Theta}_{j,it} - \Theta_{j,it}^0\right| \\
&= R_{\epsilon,it}^1 + \max_{i\in I_3,t\in[T],j\in[p]}|X_{j,it}|\,R_{\epsilon,it}^2, \text{ and} \\
\max_{i\in I_3,t\in[T]}|k_{h_N}(\hat{\epsilon}_{it}) - k_{h_N}(\epsilon_{it})| &= \frac{1}{h_N}\max_{i\in I_3,t\in[T]}\left|k(\frac{\hat{\epsilon}_{it}}{h_N}) - k(\frac{\epsilon_{it}}{h_N})\right| \lesssim \frac{1}{h_N^2}\max_{i\in I_3,t\in[T]}|\hat{\epsilon}_{it} - \epsilon_{it}| \\
&= R_{k,it}^1 + \max_{i\in I_3,t\in[T],j\in[p]}|X_{j,it}|\,R_{k,it}^2, \qquad \text{(B.101)}
\end{aligned}
$$

where $\max_{i\in I_3,t\in[T]}\left|R_{\epsilon,it}^1\right| = O_p(\eta_N)$, $\max_{i\in I_3,t\in[T]}\left|R_{\epsilon,it}^2\right| = O_p\left(\frac{\log N\vee T}{N\wedge T}\right)$, $\max_{i\in I_3,t\in[T]}\left|R_{k,it}^1\right| = O_p(\eta_N h_N^{-2})$ and $\max_{i\in I_3,t\in[T]}\left|R_{k,it}^2\right| = O_p(\frac{\log N\vee T}{N\wedge T}h_N^{-2})$ by (A.48) and Assumption 1(iv). Let

$$
\mathrm{v}_{t,t,j}^0 = \frac{1}{6}\sum_{a\in[3]}\sum_{b\in[3]\backslash\{a\}}O_j^{(b)}v_{t,j}^0 v_{t,j}^{0\prime}O_j^{(b)\prime}
$$

and recall that $\hat{\mathrm{v}}_{t,t,j} = \frac{1}{6}\sum_{a\in[3]}\sum_{b\in[3]\backslash\{a\}}\hat{v}_{t,j}^{(a,b)}\hat{v}_{t,j}^{(a,b)\prime}$. With Theorem 3.3, it is clear that

$$
\begin{aligned}
\max_{t\in[T]}\left\|\hat{\mathrm{v}}_{t,t,j} - \mathrm{v}_{t,t,j}^0\right\|_F &= \frac{1}{6}\sum_{a\in[3]}\sum_{b\in[3]\backslash\{a\}}\left(\hat{v}_{t,j}^{(a,b)}\hat{v}_{t,j}^{(a,b)\prime} - O_j^{(b)}v_{t,j}^0 v_{t,j}^{0\prime}O_j^{(b)\prime}\right) \\
&= \frac{1}{6}\sum_{a\in[3]}\sum_{b\in[3]\backslash\{a\}}\left[\left(\hat{v}_{t,j}^{(a,b)} - O_j^{(b)}v_{t,j}^0\right)\left(\hat{v}_{t,j}^{(a,b)} - O_j^{(b)}v_{t,j}^0\right)' + O_j^{(b)}v_{t,j}^0\left(\hat{v}_{t,j}^{(a,b)} - O_j^{(b)}v_{t,j}^0\right)'\right. \\
&\left. + \left(\hat{v}_{t,j}^{(a,b)} - O_j^{(b)}v_{t,j}^0\right)\left(O_j^{(b)}v_{t,j}^0\right)'\right]
\end{aligned}
$$

$$= O_p\left(\sqrt{\frac{\log N \vee T}{N}}\right).$$

Let $\mathrm{v}^0_{t,s,j} = \frac{1}{6}\sum_{a\in[3]}\sum_{b\in[3]\setminus\{a\}} O_j^{(b)} v^0_{t,j} v^{0\prime}_{s,j} O_j^{(b)\prime}$ and recall that $\hat{\mathrm{v}}_{t,s,j} = \frac{1}{6}\sum_{a\in[3]}\sum_{b\in[3]\setminus\{a\}} \hat{v}^{(a,b)}_{t,j}\hat{v}^{(a,b)\prime}_{s,j}$, similarly as above, we have $\max_{t\in[T],s\in[T]}\left\|\hat{\mathrm{v}}_{t,s,j} - \mathrm{v}^0_{t,s,j}\right\|_F = O_p\left(\sqrt{\frac{\log N\vee T}{N}}\right)$. It follows that

$$\begin{aligned}
\hat{\mathbb{V}}_{u_j} &= \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} k_{h_N}(\hat{\epsilon}_{it})\hat{e}^2_{j,it}\hat{\mathrm{v}}_{t,t,j} \\
&= \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} k_{h_N}(\epsilon_{it})e^2_{j,it}\mathrm{v}^0_{t,t,j} + \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} k_{h_N}(\epsilon_{it})e^2_{j,it}\left(\hat{\mathrm{v}}_{t,t,j} - \mathrm{v}^0_{t,t,j}\right) \\
&+ \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} k_{h_N}(\epsilon_{it})\left(\hat{e}^2_{j,it} - e^2_{j,it}\right)\mathrm{v}^0_{t,t,j} + \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} k_{h_N}(\epsilon_{it})\left(\hat{e}^2_{j,it} - e^2_{j,it}\right)\left(\hat{\mathrm{v}}_{t,t,j} - \mathrm{v}^0_{t,t,j}\right) \\
&+ \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} \left[k_{h_N}(\hat{\epsilon}_{it}) - k_{h_N}(\epsilon_{it})\right]e^2_{j,it}\mathrm{v}^0_{t,t,j} + \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} \left[k_{h_N}(\hat{\epsilon}_{it}) - k_{h_N}(\epsilon_{it})\right]e^2_{j,it}\left(\hat{\mathrm{v}}_{t,t,j} - \mathrm{v}^0_{t,t,j}\right) \\
&+ \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} \left[k_{h_N}(\hat{\epsilon}_{it}) - k_{h_N}(\epsilon_{it})\right]\left(\hat{e}^2_{j,it} - e^2_{j,it}\right)\mathrm{v}^0_{t,t,j} \\
&+ \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} \left[k_{h_N}(\hat{\epsilon}_{it}) - k_{h_N}(\epsilon_{it})\right]\left(\hat{e}^2_{j,it} - e^2_{j,it}\right)\left(\hat{\mathrm{v}}_{t,t,j} - \mathrm{v}^0_{t,t,j}\right) \\
&= \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} k_{h_N}(\epsilon_{it})e^2_{j,it}\mathrm{v}^0_{t,t,j} + O_p\left(\eta_N h_N^{-2}\right) \\
&= \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]} k_{h_N}(\epsilon_{it})e^2_{j,it}\mathrm{v}^0_{t,t,j} + o_p(1),
\end{aligned} \tag{B.102}$$

where the last two lines combines (A.47), (B.101), Assumption 10(ii) and facts that $\hat{e}^2_{j,it} - e^2_{j,it} = \left(\hat{e}_{j,it} - e_{j,it}\right)^2 + e_{j,it}\left(\hat{e}_{j,it} - e_{j,it}\right) = O_p(\eta_N^2) + e_{j,it}O_p(\eta_N)$ uniformly by Lemma B.21 and $\max_{i\in I_3,t\in[T]}|k_{h_N}(\epsilon_{it})| = O(h_N^{-1})$. By Bernstein's inequality, we obtain that

$$\begin{aligned}
\left\|\frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]}\left[k_{h_N}(\epsilon_{it})e^2_{j,it} - \mathbb{E}\left(k_{h_N}(\epsilon_{it})e^2_{j,it}\middle|\mathscr{D}\right)\right]v^0_{t,j}v^{0\prime}_{t,j}\right\|_F &= O_p\left(\sqrt{\frac{\log N}{NT}\frac{\xi_N^2}{h_N}}\right), \\
\left\|\frac{1}{T}\sum_{t\in[T]}\mathbb{E}\left[f_{it}(0)e^2_{j,it}\middle|\mathscr{D}\right]v^0_{t,j}v^{0\prime}_{t,j} - \mathbb{E}\left[f_{it}(0)e^2_{j,it}v^0_{t,j}v^{0\prime}_{t,j}\right]\right\|_F &= O_p\left(\sqrt{\frac{\log T}{T}}\xi_N^2\right).
\end{aligned} \tag{B.103}$$

Besides, by Assumption 10(i), we observe that

$$\mathbb{E}\left[k_{h_N}(\epsilon_{it})\middle|\mathscr{D}_e\right] = f_{it}(0) + O(h_N^m), \tag{B.104}$$

together with Assumption 10(v), and it gives

$$\frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]}\mathbb{E}\left(k_{h_N}(\epsilon_{it})e^2_{j,it}\middle|\mathscr{D}\right)v^0_{t,j}v^{0\prime}_{t,j} = \frac{1}{T}\sum_{t\in[T]}\mathbb{E}\left[f_{it}(0)e^2_{j,it}\middle|\mathscr{D}\right]v^0_{t,j}v^{0\prime}_{t,j} + O(h_N^m). \tag{B.105}$$

Combining (B.102)-(B.105) and Assumption 10(ii), we obtain that $\hat{\mathbb{V}}_{u_j} = O_j^{(1)}V_{u_j}O_j^{(1)\prime} + o_p(1)$. By analogous analysis and Assumption 10(v), we can also show that $\hat{\mathbb{V}}^{(1)}_{v_j} = O_j^{(1)}V_{v_j}O_j^{(1)\prime} + o_p(1)$.

Next, we show the consistency of $\hat{\Omega}_{u_j}$. With the restriction for T_1 in Assumption 10(iii), we first note that

$$
\begin{aligned}
\Omega_{u_j} &= Var\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} e_{j,it} v_{t,j}^{0}(\tau - \mathbf{1}\{\epsilon_{it} \le 0\})\right] \\
&= \tau(1-\tau)\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left(e_{j,it}^2 v_{t,j}^{0} v_{t,j}^{0\prime}\right) + \frac{1}{T}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\mathbb{E}\left[e_{j,it}e_{j,is}v_{t,j}^{0}v_{s,j}^{0\prime}(\tau - \mathbf{1}\{\epsilon_{it}\le 0\})(\tau - \mathbf{1}\{\epsilon_{is}\le 0\})\right] \\
&+ \frac{1}{T}\sum_{t=1+T_1}^{T}\sum_{s=t-T_1}^{t-1}\mathbb{E}\left[e_{j,it}e_{j,is}v_{t,j}^{0}v_{s,j}^{0\prime}(\tau - \mathbf{1}\{\epsilon_{it}\le 0\})(\tau - \mathbf{1}\{\epsilon_{is}\le 0\})\right] + O(\alpha^{T_1}) \\
&= \tau(1-\tau)\frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left(e_{j,it}^2 v_{t,j}^{0} v_{t,j}^{0\prime}\right) + \frac{1}{T}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\mathbb{E}\left[e_{j,it}e_{j,is}v_{t,j}^{0}v_{s,j}^{0\prime}\left(F_{i,ts}(0,0)-\tau^2\right)\right] \\
&+ \frac{1}{T}\sum_{t=1+T_1}^{T}\sum_{s=t-T_1}^{t-1}\mathbb{E}\left[e_{j,it}e_{j,is}v_{t,j}^{0}v_{s,j}^{0\prime}\left(F_{i,ts}(0,0)-\tau^2\right)\right] + o(1),
\end{aligned}
$$

where the second equality is by Assumption 1(iii), the third equality is by Assumption 1(vii) and Assumption 10(iii). Compared to $\hat{\Omega}_{u_j}$, what remains to show are

$$
\frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]}\hat{e}_{j,it}^2\hat{\mathrm{v}}_{t,t,j} = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left(e_{j,it}^2\mathrm{v}_{t,t,j}^{0}\right) + o_p(1), \tag{B.106}
$$

$$
\begin{aligned}
&\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}\left[\tau - K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)\right]\left[\tau - K\left(\frac{\hat{\epsilon}_{is}}{h_N}\right)\right] \\
&= \frac{1}{T}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\mathbb{E}\left[e_{j,it}e_{j,is}\mathrm{v}_{t,s,j}^{0}\left(F_{i,ts}(0,0)-\tau^2\right)\right] + o_p(1),
\end{aligned} \tag{B.107}
$$

$$
\begin{aligned}
&\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1+T_1}^{T}\sum_{s=t-T_1}^{t-1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}\left[\tau - K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)\right]\left[\tau - K\left(\frac{\hat{\epsilon}_{is}}{h_N}\right)\right] \\
&= \frac{1}{T}\sum_{t=1+T_1}^{T}\sum_{s=t-T_1}^{t-1}\mathbb{E}\left[e_{j,it}e_{j,is}\mathrm{v}_{t,s,j}^{0}\left(F_{i,ts}(0,0)-\tau^2\right)\right] + o_p(1).
\end{aligned} \tag{B.108}
$$

For (B.106), like (B.102), we notice that

$$
\begin{aligned}
\frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]}\hat{e}_{j,it}^2\hat{\mathrm{v}}_{t,t,j} &= \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]}e_{j,it}^2\mathrm{v}_{t,t,j}^{0} + o_p(1) \\
&= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left(e_{j,it}^2\mathrm{v}_{t,t,j}^{0}\right) + \frac{1}{NT}\sum_{i\in[N]}\sum_{t\in[T]}\left[e_{j,it}^2\mathrm{v}_{t,t,j}^{0} - \mathbb{E}\left(e_{j,it}^2\mathrm{v}_{t,t,j}^{0}\right)\right] + o_p(1) \\
&= \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}\left(e_{j,it}^2\mathrm{v}_{t,t,j}^{0}\right) + o_p(1),
\end{aligned} \tag{B.109}
$$

where the first equality is by Lemma B.21 and (A.47), and the second equality is by Bernstein's inequality and Assumption 10(v).

For (B.107), we observe that

$$
\begin{aligned}
&\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}\left[\tau-K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)\right]\left[\tau-K\left(\frac{\hat{\epsilon}_{is}}{h_N}\right)\right]\\
&=\tau^2\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}-\tau\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)\\
&-\tau\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}K\left(\frac{\hat{\epsilon}_{is}}{h_N}\right)+\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)K\left(\frac{\hat{\epsilon}_{is}}{h_N}\right)
\end{aligned}
\tag{B.110}
$$

such that

$$
\begin{aligned}
&\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}=\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}+O_p\left(T_1\eta_N\right)\\
&=\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\mathbb{E}\left(e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}\right)+\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\left[e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}-\mathbb{E}\left(e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}\right)\right]\\
&+O_p\left(T_1\eta_N\right)\\
&=\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\mathbb{E}\left(e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}\right)+o_p(1),
\end{aligned}
\tag{B.111}
$$

where the first equality is by the similar arguments as (B.102) and the last line combines Assumption 10(iii) and the fact that the second term in the second equality can be shown to be $o_p(1)$ by Bernstein's inequality. Furthermore, with the fact that $\max_{i\in I_3,t\in[T]}\left|K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)-K\left(\frac{\epsilon_{it}}{h_N}\right)\right|\lesssim\frac{1}{h_N}\max_{i\in I_3,t\in[T]}\left|\hat{\epsilon}_{it}-\epsilon_{it}\right|=O_p\left(\eta_N h_N^{-1}\right)$ and by the analogous arguments as above, we can show that

$$
\begin{aligned}
&\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)\\
&=\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}K\left(\frac{\epsilon_{it}}{h_N}\right)+O_p\left(T_1\eta_N h_N^{-1}\right)\\
&=\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}\mathbb{E}\left[K\left(\frac{\epsilon_{it}}{h_N}\right)\Big|\mathscr{D}_e\right]+o_p(1)\\
&=\frac{\tau}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}+O(h_N^m)+o_p(1)\\
&=\frac{\tau}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\mathbb{E}\left[e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}\right]+O(h_N^m)+o_p(1)\\
&=\frac{\tau}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\mathbb{E}\left[e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}\right]+o_p(1),
\end{aligned}
\tag{B.112}
$$

where the first equality is similar as (B.102), the second equality is by addition and subtracting and Assumption 10(iii), the third equality is by the fact that $\mathbb{E}\left[K\left(\frac{\epsilon_{it}}{h_N}\right)\big|\mathscr{D}_e\right]=\tau+O(h_N^m)$ by the calculation of nonparametric kernel estimator which can be found in Galvao and Kato (2016). The last equality is similar as the second equality and combines Assumption 10(i) and Assumption 10(ii).

Moreover, similarly as (B.111), with the fact that $\mathbb{E}\left[K\left(\frac{\epsilon_{it}}{h_N}\right)K\left(\frac{\epsilon_{is}}{h_N}\right)|\mathscr{D}_e\right] = F_{i,ts}(0,0) + O(h_N^m)$, we can show that

$$
\begin{aligned}
&\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\hat{e}_{j,it}\hat{e}_{j,is}\hat{\mathrm{v}}_{t,s,j}K\left(\frac{\hat{\epsilon}_{it}}{h_N}\right)K\left(\frac{\hat{\epsilon}_{is}}{h_N}\right)\\
&=\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}K\left(\frac{\epsilon_{it}}{h_N}\right)K\left(\frac{\epsilon_{is}}{h_N}\right)+O_p\left(\sqrt{\frac{\log(N\vee T)}{N\wedge T}}\frac{T_1\xi_N^2}{h_N^2}\right)\\
&=\frac{1}{NT}\sum_{i\in[N]}\sum_{t=1}^{T-T_1}\sum_{s=t+1}^{t+T_1}\mathbb{E}\left[e_{j,it}e_{j,is}\mathrm{v}^0_{t,s,j}F_{i,ts}(0,0)\right]+o_p(1). \qquad \text{(B.113)}
\end{aligned}
$$

Combining (B.110)-(B.113), we complete the proof for (B.107). By the analogous arguments, we can show the proof for (B.108), which yields $\hat{\Omega}_{u_j} = \Omega_{u_j} + o_p(1)$. ■

C Algorithm for low-rank Estimation

In this section, we provide the algorithm for the case of low-rank estimation with two regressors, the case of more than two regressors is self-evident. To solve the regularized quantile regression, let the optimization problem with two regressors be as

$$
\min_{\Theta_0,\Theta_1,\Theta_2}\quad \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\rho_\tau\left(y_{it}-\Theta_{0,it}-x_{1,it}\Theta_{1,it}-x_{2,it}\Theta_{2,it}\right)+\nu_0\left\|\Theta_0\right\|_*+\nu_1\left\|\Theta_1\right\|_*+\nu_2\left\|\Theta_2\right\|_*.
$$

As in Belloni et al. (2022), the above minimization problem is equivalent to the following one:

$$
\begin{aligned}
&\min_{\Theta_0,\Theta_1,\Theta_2,V,W,Z_{\Theta_0},Z_{\Theta_1},Z_{\Theta_2}}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\rho_\tau\left(V_{it}\right)+\nu_0\left\|\Theta_0\right\|_*+\nu_1\left\|Z_{\Theta_1}\right\|_*+\nu_2\left\|Z_{\Theta_2}\right\|_*\\
&s.t.\ V=W,\ \ W=Y-X_1\odot\Theta_1-X_2\odot\Theta_2-Z_{\Theta_0},\\
&Z_{\Theta_0}-\Theta_0=0,\ \ Z_{\Theta_1}-\Theta_1=0,\ \ Z_{\Theta_2}-\Theta_2=0.
\end{aligned}
$$

As our theoretical results show, ν_0, ν_1 and ν_2 converge to zero at rate $\frac{\sqrt{N}\vee\sqrt{T}}{NT}$. The augmented Lagrangian is

$$
\begin{aligned}
&\mathscr{L}\left(V,W,\Theta_0,Z_{\Theta_0},\Theta_1,Z_{\Theta_1},\Theta_2,Z_{\Theta_2},U_v,U_w,U_{\Theta_0},U_{\Theta_1},U_{\Theta_2}\right)\\
&=\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\rho_\tau(V_{it})+\nu_0\left\|\Theta_0\right\|_*+\nu_1\left\|Z_{\Theta_1}\right\|_*+\nu_2\left\|Z_{\Theta_2}\right\|_*+\frac{\rho}{2NT}\left\|V-W+U_v\right\|_F^2\\
&+\frac{\rho}{2NT}\left\|W-Y+X_1\odot\Theta_1+X_2\odot\Theta_2+Z_{\Theta_0}+U_W\right\|_F^2+\frac{\rho}{2NT}\left\|Z_{\Theta_0}-\Theta_0+U_{\Theta_0}\right\|_F^2\\
&+\frac{\rho}{2NT}\left\|Z_{\Theta_1}-\Theta_1+U_{\Theta_1}\right\|_F^2+\frac{\rho}{2NT}\left\|Z_{\Theta_2}-\Theta_2+U_{\Theta_2}\right\|_F^2,
\end{aligned}
$$

where $\rho>0$ is the penalty parameter.

By ADMM algorithm, similarly as Belloni et al. (2019), updates are as follows:

$$V^{k+1} \leftarrow \arg\min_{V} \left\{ \frac{1}{NT} \sum_{i} \sum_{t} \rho_{\tau} \left(V_{it} \right) + \frac{\rho}{2NT} \left\| V - W^{k} + U_{v}^{k} \right\|_F^2 \right\} \tag{C.1}$$

$$(\Theta_1^{k+1}, \Theta_2^{k+1)} \leftarrow \arg\min_{\Theta_1, \Theta_2} \left\{ \left\| W^{k} - Y + X_1 \odot \Theta_1 + X_2 \odot \Theta_2 + Z_{\Theta_0}^{k} + U_{W}^{k} \right\|_F^2 + \left\| Z_{\Theta_1}^{k} - \Theta_1 + U_{\Theta_1}^{k} \right\|_F^2 \right. \tag{C.2}$$

$$\left. + \left\| Z_{\Theta_2}^{k} - \Theta_2 + U_{\Theta_2}^{k} \right\|_F^2 \right\}$$

$$\Theta_0^{k+1} \leftarrow \arg\min_{\Theta_0} \left\{ \frac{1}{2} \left\| Z_{\Theta_0}^{k} - \Theta_0 + U_{\Theta_0}^{k} \right\|_F^2 + \frac{\nu_0 NT}{\rho} \left\| \Theta_0 \right\|_* \right\} \tag{C.3}$$

$$Z_{\Theta_1}^{k+1} \leftarrow \arg\min_{Z_{\Theta_1}} \left\{ \frac{1}{2} \left\| \Theta_1^{k+1} - U_{\Theta_1}^{k} - Z_{\Theta_1} \right\|_F^2 + \frac{\nu_1 NT}{\rho} \left\| Z_{\Theta_1} \right\|_* \right\}$$

$$Z_{\Theta_2}^{k+1} \leftarrow \arg\min_{Z_{\Theta_2}} \left\{ \frac{1}{2} \left\| \Theta_2^{k+1} - U_{\Theta_2}^{k} - Z_{\Theta_2} \right\|_F^2 + \frac{\nu_2 NT}{\rho} \left\| Z_{\Theta_2} \right\|_* \right\}$$

$$(Z_{\Theta_0}^{k+1}, W^{k+1}) \leftarrow \arg\min_{Z_{\Theta_0}, W} \left\{ \left\| V^{k+1} - W + U_{v}^{k} \right\|_F^2 + \left\| W - Y + X_1 \odot \Theta_1^{k+1} + X_2 \odot \Theta_2^{k+1} + Z_{\Theta_0} + U_{W}^{k} \right\|_F^2 \right.$$

$$\left. + \left\| Z_{\Theta_0} - \Theta_0^{k+1} + U_{\Theta_0}^{k} \right\|_F^2 \right\}$$

$$U_{v}^{k+1} \leftarrow V^{k+1} - W^{k+1} + U_{v}^{k}$$

$$U_{W}^{k+1} \leftarrow W^{k+1} - Y + X_1 \odot \Theta_1^{k+1} + X_2 \odot \Theta_2^{k+1} + Z_{\Theta_0}^{k+1} + U_{W}^{k}$$

$$U_{\Theta_0}^{k+1} \leftarrow Z_{\Theta_0}^{k+1} - \tilde{\Theta_0}^{k+1} + U_{\Theta_0}^{k}$$

$$U_{\Theta_1}^{k+1} \leftarrow Z_{\Theta_1}^{k+1} - \Theta_1^{k+1} + U_{\Theta_1}^{k}$$

$$U_{\Theta_2}^{k+1} \leftarrow Z_{\Theta_2}^{k+1} - \Theta_2^{k+1} + U_{\Theta_2}^{k}$$

For (C.1), by Ali et al. (2016),

$$V^{k+1} \leftarrow P_{+} \left(W^{k} - U_{v}^{k} - \frac{\tau}{\rho} \iota_N \iota_T' \right) + P_{-} \left(W^{k} - U_{v}^{k} - \frac{(1-\tau)}{\rho} \iota_N \iota_T' \right).$$

where ι_N is the $N \times 1$ all-ones vector, and same for ι_T. For (C.2), first order condition gives

$$\Theta_{1,it}^{k+1} = \frac{\left(1 + x_{2,it}^2\right) \left(Z_{\Theta_1,it}^{k} + U_{\Theta_1,it}^{k} - A_{it} x_{1,it}\right) - x_{1,it} x_{2,it} \left(Z_{\Theta_2,it}^{k} + U_{\Theta_2,it}^{k} - A_{it} x_{2,it}\right)}{1 + x_{1,it}^2 + x_{2,it}^2},$$

$$\Theta_{2,it}^{k+1} = \frac{\left(1 + x_{1,it}^2\right) \left(Z_{\Theta_2,it}^{k} + U_{\Theta_2,it}^{k} - A_{it} x_{2,it}\right) - x_{1,it} x_{2,it} \left(Z_{\Theta_1,it}^{k} + U_{\Theta_1,it}^{k} - A_{it} x_{1,it}\right)}{1 + x_{1,it}^2 + x_{2,it}^2},$$

where

$$A := W^{k} + Z_{\Theta_0}^{k} + U_{W}^{k} - Y.$$

To solve (C.3), by singular value thresholding estimations, the update for Θ_0^{k+1} is

$$\Theta_0^{k+1} \leftarrow P_0 D_{0, \frac{\nu_1 NT}{\rho}} Q_0',$$

where $Z_{\Theta_0}^{k} + U_{\Theta_0}^{k} = P_0 D_0 Q_0'$, and $D_{0, \frac{\nu_0}{\rho}, ii} = \max(D_{0,ii} - \frac{\nu_0}{\rho}, 0)$. Similarly for $Z_{\Theta_1}^{k+1}$ and $Z_{\Theta_2}^{k+1}$,

$$Z_{\Theta_1}^{k+1} \leftarrow P_1 D_{1, \frac{\nu_1 NT}{\rho}} Q_1',$$

$$Z_{\Theta_2}^{k+1} \leftarrow P_2 D_{2,\frac{\nu_2 NT}{\rho}} Q_2',$$

where $\Theta_1^{k+1} - U_{\Theta_1}^k = P_1 D_1 Q_1'$, $\Theta_2^{k+1} - U_{\Theta_2}^k = P_2 D_2 Q_2'$, $D_{1,\frac{\nu_1}{\rho},ii} = \max\left(D_{1,ii} - \frac{\nu_1}{\rho}, 0\right)$, and $D_{2,\frac{\nu_2}{\rho},ii} = \max\left(D_{2,ii} - \frac{\nu_2}{\rho}, 0\right)$.

Finally, let $\tilde{A} := -Y + X_1 \odot \Theta_1^{k+1} + X_2 \odot \Theta_2^{k+1} + U_W^{k+1}$, $\tilde{B} := -V^{k+1} - U_v^k$, $\tilde{C} := -\tilde{\Theta_0}^{k+1} + U_{\Theta_0}^k$, then

$$Z_{\Theta_0}^{k+1} \leftarrow \frac{-\tilde{A} - 2\tilde{C} + \tilde{B}}{3},$$
$$W^{k+1} \leftarrow -\tilde{A} - \tilde{C} - 2Z_{\Theta_0}^{k+1}.$$