Discussion Paper No. 1019

INSTRUMENTAL VARIABLE ESTIMATION OF DYNAMIC LINEAR PANEL DATA MODELS WITH DEFACTORED REGRESSORS AND A MULTIFACTOR ERROR STRUCTUR E

Milda Norkuté
Vasilis Sarafidis
Takashi Yamagata
Guowei Cui

Revised April 2019
February 2018

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

Instrumental Variable Estimation of Dynamic Linear Panel Data Models with Defactored Regressors and a Multifactor Error Structure[◇]

Milda Norkutė[*]
Lund University

Vasilis Sarafidis[†]
Monash University

Takashi Yamagata[‡]
University of York and Osaka University

Guowei Cui[§]
Monash University

30 April 2019

Abstract

This paper develops two instrumental variable (IV) estimators for dynamic panel data models with exogenous covariates and a multifactor error structure when both cross-sectional and time series dimensions, N and T respectively, are large. Our approach initially projects out the common factors from the exogenous covariates of the model, and constructs instruments based on this defactored covariates. For models with homogeneous slope coefficients, we propose a two-step IV estimator: the first step IV estimator is obtained using the defactored covariates as instruments. In the second step, the entire model is defactored by the extracted factors from the residuals of the first step estimation and subsequently obtain the final IV estimator. For models with heterogeneous slope coefficients, we propose a mean-group type estimator, which is the cross-sectional average of first-step IV estimators of cross-section specific slopes. It is noteworthy that our estimators do not require us to seek for instrumental variables outside the model. Furthermore, our estimators are linear hence computationally robust and inexpensive. Moreover, they require no bias correction, and they are not subject to the small sample bias of least squares type estimators. The finite sample performances of the proposed estimators and associated statistical tests are investigated, and the results show that the estimators and the tests perform well even for small N and T.

Key Words: method of moments; dynamic panel data; cross-sectional dependence; factor model

JEL Classification: C13, C15, C23.

[*]Department of Economics, Lund University, P.O. Box 7082, S-220 07 Lund, Sweden. E-mail address: Milda.Norkute@nek.lu.se.

[†]Department of Econometrics and Business Statistics, Monash University, VIC 3145, Australia. E-mail address: vasilis.sarafidis@monash.edu.

[‡]Corresponding author. Department of Economics and Related Studies, University of York, York YO10 5DD, UK; E-mail address: takashi.yamagata@york.ac.uk.

This author gratefully acknowledges the partial support of JSPS KAKENHI JP15H05728 and JP18K01545.

[§]Department of Econometrics and Business Statistics, Monash University, VIC 3145, Australia. E-mail address: cuigw610@163.com.

1 Introduction

The rapid increase in the availability of panel data during the last few decades has invoked a large interest in developing ways to model and analyse them effectively. In particular, the issue of how to characterise 'between group' or cross-sectional dependence, and then creating consistent estimation methods and making asymptotically valid inferences, has proven both popular and challenging. The factor structure approach has been widely used to model cross-sectional dependence. It escapes from the curse of dimensionality by asserting that there exists a common component, which is a linear combination of a finite number of time-varying common factors with individual-specific factor loadings. One can provide different interpretations of this approach, depending on the application in mind. In macroeconomic panels the unobserved factors are frequently viewed as economy-wide shocks, affecting all individuals albeit with different intensities; see e.g. Favero et al. (2005). In microeconomic panels the factor error structure may be thought to reflect distinct sources of unobserved individual-specific heterogeneity, the impact of which varies over time. For instance, in a model of wage determination the factor loadings may represent several unmeasured skills, specific to each individual, while the factors may capture the price of these skills, which changes intertemporally in an arbitrary way; see e.g. Carneiro et al. (2003) and Heckman et al. (2006).

A large body of the literature has focused on developing statistical inferential methods for models with an error factor structure. For large panels, two estimation approaches have been popular: Pesaran (2006) proposed the Common Correlated Effects (CCE) estimator, that consists of approximating the unobserved factors by the linear combinations of cross-sectional averages of the dependent and explanatory variables. Bai (2009a) proposed an iterative least squares estimator with bias corrections, approximating the unobserved factors by principal component (PC) estimator.[1] For both estimators it is assumed that the regressors are strictly exogenous with respect to the idiosyncratic error component, whereas possible correlation between the regressors and the error factor component is permitted. Under somewhat weaker assumptions, Moon and Weidner (2015) show that the estimator of Bai (2009a) is interpretable as a quasi maximum likelihood estimator (QMLE), the consistency of which is maintained even when the number of factors is not specified correctly, so long as it is larger than or equal to the true number of factors.

In this paper we consider estimation of linear dynamic panel data models with an error factor structure in large panels.[2] Recently, the CCE and the PC estimators have been extended to accommodate this case as well. In particular, Chudik and Pesaran (2015a) propose mean group CCE (CCEMG) estimation of panel autoregressive distributed lag models. The dynamic structure considered therein is very general for two reasons. Firstly, it permits cross-sectionally heterogeneous slope coefficients. Secondly, their model can be seen as a structural transformation of a multivariate dynamic process, such as a vector autoregressive model. Chudik and Pesaran (2015a) employ a mean group type estimator to deal with the slope heterogeneity, and propose to augment the regression with the cross-sectional averages of dependent variables and covariates and their lags, in order to control the common components.

On the other hand, Moon and Weidner (2017) propose a bias-corrected QMLE (BC-QMLE) estimator for dynamic panel data models with homogeneous slopes, and put forward bias-corrected likelihood based tests. Unlike CCEMG and the approach proposed in the present paper, they allow covariates to be correlated to the common component in disturbances without

[1] See Westerlund and Urbain (2015) for a comparison analysis of the CCE and PC estimation. Chudik and Pesaran (2015b), Sarafidis and Wansbeek (2012) and Bai and Wang (2016) also provide excellent surveys on the related literature.

[2] Estimation of such models for short panel data is considered by Ahn et al. (2013) and Robertson and Sarafidis (2015).

having necessarily a linear factor structure. Furthermore, the precision of the estimator is expected to be higher than existing estimators under certain regularity conditions.

This paper develops two instrumental variable (IV) estimators for dynamic panel data models with exogenous covariates and a multifactor error structure when both cross-sectional and time series dimensions, N and T respectively, are large. Our approach initially projects out the common factors from the exogenous covariates of the model, and constructs instruments based on defactored covariates in order to build a consistent first step IV estimator.[3]

At the beginning, we consider a two-step IV estimator for models with homogeneous slope coefficients. The first step IV estimator is obtained using the defactored covariates as instruments. In the second step, the entire model is defactored by the extracted factors from the residuals of the first step estimation and subsequently obtain the final IV estimator. We show the $\sqrt{NT}$-consistency of the two-step estimator and establish its asymptotic normality. Although both our approach and the QMLE approach of Moon and Weidner (2017) is based on the PC estimator, there are important differences in practice; firstly, since our estimator is an instrumental variable estimator, it is not subject to the small sample bias, known as "Nickell bias" that arises with least squares type estimators in dynamic panel data models. Secondly, our estimator is linear hence robust and computationally inexpensive, whereas obtaining the QMLE estimator requires nonlinear optimisation, which can be more costly and can fail to reach at the global minimum.[4] Thirdly, our estimator does not have asymptotic bias unlike the QMLE estimator, which requires bias-corrections to re-centre the limiting distribution of the original estimator.

Next, we consider estimation of models with heterogeneous slope coefficients. In particular, we propose a mean group IV estimator, which is the cross-sectional average of first-step IV estimators of cross-section specific coefficients. We establish the $\sqrt{N}$-consistency of our estimator to the population average of the slopes and its asymptotic normality. Our estimator has some advantages over the CCEMG estimator of Chudik and Pesaran (2015a). Firstly, we employ the PC approach for defactoring the exogenous covariates, therefore we do not need to seek external variables to approximate the factors when the number of unobserved factors is larger than the number of covariates plus one. By contrast, in this situation the CCE estimation requires additional sets of variables, which are not in the original model of interest but expected to form a part of the dynamic system. In practice, this may not be a trivial exercise. Secondly, the CCE estimator is subject to the small T bias of least squares estimators, whilst our estimator is not, since it is based on the IV estimation. Chudik and Pesaran (2015a) propose to adjust the bias using the jackknife method, which might not be very effective for small or moderate T, especially so with persistent data.

The required assumption underlying our IV approach is that any sources of endogeneity of the covariates arise solely due to the non-zero correlation between the common components in the covariates and in the model disturbances. Notably, this assumption can be tested using an overidentifying restrictions test.

Incidentally, our approach can be regarded as the opposite one employed by Bai and Ng (2010) and Kapetanios and Marcellino (2010). In specific, in their model the idiosyncratic errors of the reduced form regression of the covariates cause endogeneity, therefore, no error factor structure is considered in the structural model of interest. They propose finding instruments for this endogenous covariates by extracting the common components from external variables and the endogenous covariates in the model. Our approach essentially complements theirs.

Using simulated data it is shown that the proposed approach performs satisfactorily under all circumstances examined. In particular, in comparison to the aforementioned alternative methods, both IV estimators appear to have little or negligible bias in most circumstances,

[3]Our methodology can be regarded as an extension of the approach taken by Sarafidis et al. (2009).

[4]See Moon and Weidner (2019) for more details.

and correct size of the t-test even for small sample size. Furthermore, the overidentifying restrictions test appears to have high power when the key assumption of the model is violated, namely the exogeneity of the covariates with respect to the purely idiosyncratic disturbance. In addition, the test tends to have good power under slope parameter heterogeneity, unless the number of degrees of freedom of the test statistic is very small. By contrast, the CCEMG estimator can suffer from non-negligible bias and large size distortions of the associated t-test. Similarly, although under slope homogeneity BC-QMLE tends to exhibit the smallest dispersion, it suffers from large bias and substantial size distortions of the associated bias-corrected test, unless both N and T is large.

The paper is organised as follows. Section 2 sets out the model with homogeneous slopes and assumptions to introduce the two-step IV estimator and its optimal version with the associated overidentifying restrictions test, then investigates their asymptotic properties. Section 3 develops consistent estimators of cross-sectionally heterogeneous slope coefficients and its averages, then establishes their asymptotic normality. Section 4 studies the finite sample performance of the proposed estimators along with the CCE estimator of Chudik and Pesaran (2015a) the QMLE estimator of Moon and Weidner (2017) using simulated data. Section 5 contains some concluding remarks. Proofs of propositions, theorems and corollaries, together with used lemmas, are contained in Appendix A. Appendix B gives proofs of all the lemmas and Appendix C provides extra experimental results, which are available in Supplemental Material.

2 Model and Estimation Method

In this section we consider the following autoregressive distributed lag, ARDL$(1,0)$, panel data model with homogeneous slopes and a multifactor error structure[5]

$$y_{it} = \rho y_{i,t-1} + \boldsymbol{\beta}'\mathbf{x}_{it} + u_{it}; \quad i = 1, 2, ..., N; \quad t = 1, 2, ..., T, \tag{1}$$

with

$$u_{it} = \boldsymbol{\gamma}_{yi}^{0\prime}\mathbf{f}_{y,t}^0 + \varepsilon_{it}, \tag{2}$$

where $|\rho| < 1$, $\boldsymbol{\beta} = (\beta_1, \beta_2, ..., \beta_k)'$ with at least one of $\{\beta_\ell\}_{\ell=1}^k$ being non-zero, $\mathbf{x}_{it} = (x_{1it}, x_{2it}, ..., x_{kit})'$ is a $k \times 1$ vector of regressors, $\mathbf{f}_{y,t}^0 = (f_{y,1t}^0, f_{y,2t}^0, ..., f_{y,m_y t}^0)'$ denotes an $m_y \times 1$ *true* vector of unobservable factors. The $m_x \times 1$ vector $\boldsymbol{\gamma}_{yi}^0$ contains the *true* factor loadings associated with $\mathbf{f}_{y,t}^0$, whereas ε_{it} is an idiosyncratic error. $\mathbf{x}_{it}$ is subject to the following process:

$$\mathbf{x}_{it} = \boldsymbol{\Gamma}_{xi}^{0\prime}\mathbf{f}_{x,t}^0 + \mathbf{v}_{it}, \tag{3}$$

where $\boldsymbol{\Gamma}_{xi}^0 = (\boldsymbol{\gamma}_{1i}^0, \boldsymbol{\gamma}_{2i}^0, ..., \boldsymbol{\gamma}_{ki}^0)$ denotes the true $m_x \times k$ factor loading matrix, $\mathbf{f}_{x,t}^0 = (f_{x,1t}^0, f_{x,2t}^0, ..., f_{x,m_x t}^0)'$ denotes an $m_x \times 1$ vector of true factors, and $\mathbf{v}_{it} = (v_{1it}, v_{2it}, ..., v_{kit})'$ is an idiosyncratic error term which is independent of ε_{it}.

Remark 1 Our approach permits correlation between and within $\boldsymbol{\gamma}_{yi}^0$ and $\boldsymbol{\Gamma}_{xi}^0$, and (non)overlapping elements in $\mathbf{f}_{y,t}^0$ and $\mathbf{f}_{x,t}^0$ may be correlated to each other. Importantly, our approach can control for endogeneity of $\mathbf{x}_{it}$ which stems from the common components, whereas it is assumed to be strongly exogenous with respect to the idiosyncratic errors, ε_{it}.

Remark 2 When time invariant effects and cross-sectionally invariant time effects exist in u_{it} and $\mathbf{x}_{it}$, all the results in this paper will remain unchanged if $\{y_{it}, \mathbf{x}_{it}\}$ is replaced with the transformed variables $\{\dot{y}_{it}, \dot{\mathbf{x}}_{it}'\}$, where $\dot{y}_{it} = y_{it} - \bar{y}_i - \bar{y}_t + \bar{y}$ and $\dot{\mathbf{x}}_{it} = \mathbf{x}_{it} - \bar{\mathbf{x}}_i - \bar{\mathbf{x}}_t + \bar{\mathbf{x}}$

[5]The main results of this paper naturally extend to the models with higher order lags, i.e. ARDL(p,q) for $p > 0$ and $q \geq 0$. Also the models with heterogeneous slopes are considered in Section 3.

with $\bar{y}_i = T^{-1}\sum_{t=0}^{T} y_{it}$, $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$, $\bar{y} = N^{-1}\sum_{i=1}^{N} \bar{y}_i$, and $\bar{\mathbf{x}}_i$, $\bar{\mathbf{x}}_t$ and $\bar{\mathbf{x}}$ are defined analogously. Indeed, the experiments for our proposed estimators and tests implemented in Section 4 are based on the transformed variables.

Stacking the T observations for each i yields

$$\mathbf{y}_i = \rho \mathbf{y}_{i,-1} + \mathbf{X}_i \boldsymbol{\beta} + \mathbf{u}_i \text{ with } \mathbf{u}_i = \mathbf{F}_y^0 \boldsymbol{\gamma}_{yi}^0 + \boldsymbol{\varepsilon}_i, \tag{4}$$

where $\mathbf{y}_i = (y_{i1}, y_{i2}, ..., y_{iT})'$, $\mathbf{y}_{i,-1} = L^1 \mathbf{y}_i = (y_{i0}, y_{i1}, ..., y_{iT-1})'$ with L^j being the j^{th} lag operator, $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, ..., \mathbf{x}_{iT})'$, $\mathbf{u}_i = (u_{i1}, u_{i2}, ..., u_{iT})'$, $\mathbf{F}_y^0 = (\mathbf{f}_{y,1}^0, \mathbf{f}_{y,2}^0, ..., \mathbf{f}_{y,T}^0)'$ and $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iT})'$. Similarly,

$$\mathbf{X}_i = \mathbf{F}_x^0 \boldsymbol{\Gamma}_{xi}^0 + \mathbf{V}_i, \tag{5}$$

where $\mathbf{F}_x^0 = (\mathbf{f}_{x,1}^0, \mathbf{f}_{x,2}^0, ..., \mathbf{f}_{x,T}^0)'$ and $\mathbf{V}_i = (\mathbf{v}_{i1}, \mathbf{v}_{i2}, ..., \mathbf{v}_{iT})'$.

Let $\mathbf{W}_i = (\mathbf{y}_{i,-1}, \mathbf{X}_i)$ and $\boldsymbol{\theta} = (\rho, \boldsymbol{\beta}')'$. The model in (4) can be written more concisely as

$$\mathbf{y}_i = \mathbf{W}_i \boldsymbol{\theta} + \mathbf{u}_i. \tag{6}$$

Our estimation approach involves two steps. In the first step, we asymptotically eliminate the common factors in $\mathbf{X}_i$ by projecting them out, then using the defactored regressors as instruments to consistently estimate the structural parameters of the model. To illustrate the first step estimator, momentarily supposing $\mathbf{F}_x^0$ observable, consider the following projection matrices:

$$\mathbf{M}_{F_x^0} = \mathbf{I}_T - \mathbf{F}_x^0 \left(\mathbf{F}_x^{0\prime} \mathbf{F}_x^0\right)^{-1} \mathbf{F}_x^{0\prime}; \; \mathbf{M}_{F_{x,-1}^0} = \mathbf{I}_T - \mathbf{F}_{x,-1}^0 \left(\mathbf{F}_{x,-1}^{0\prime} \mathbf{F}_{x,-1}^0\right)^{-1} \mathbf{F}_{x,-1}^{0\prime}, \tag{7}$$

where $\mathbf{F}_{x,-1}^0 = L^1 \mathbf{F}_x^0$. If $\mathbf{F}_x^0$ were observed, premultiplying $\mathbf{X}_i$ by $\mathbf{M}_{F_x^0}$ would yield $\mathbf{M}_{F_x^0} \mathbf{X}_i = \mathbf{M}_{F_x^0} \mathbf{V}_i$. Assuming $\mathbf{V}_i$ is independent of $\boldsymbol{\varepsilon}_i, \mathbf{F}_x^0, \mathbf{F}_y^0$ and $\boldsymbol{\gamma}_{yi}^0$, it is easily seen that $E(\mathbf{X}_i' \mathbf{M}_{F_x^0} \mathbf{u}_i) = E(\mathbf{V}_i' \mathbf{M}_{F_x^0} \mathbf{u}_i) = \mathbf{0}$. Furthermore, let $\mathbf{X}_{i,-j} = L^j \mathbf{X}_i$. So long as $\{y_{it}, \mathbf{x}_{it}'\}$, $t = 0, 1, ..., T$ is observed, the $T \times k$ matrix $\mathbf{X}_{i,-1}$ is also observed. Using similar assumptions, one can show that $E(\mathbf{X}_{i,-1}' \mathbf{M}_{F_{x,-1}^0} \mathbf{u}_i) = E(\mathbf{V}_{i,-1}' \mathbf{M}_{F_{x,-1}^0} \mathbf{u}_i) = \mathbf{0}$. Now collect the set of the instrumental variables:

$$\mathbf{Z}_i = \left(\mathbf{M}_{F_x^0} \mathbf{X}_i, \mathbf{M}_{F_{x,-1}^0} \mathbf{X}_{i,-1}\right) \; (T \times 2k). \tag{8}$$

Given the model in equation (6), it is clear that $\mathbf{Z}_i$ satisfies $E(\mathbf{Z}_i' \mathbf{u}_i) = \mathbf{0}$ and also $E(\mathbf{Z}_i' \mathbf{W}_i) \neq \mathbf{0}$, thus, it is a valid instrument set.[6]

Having obtained the consistent first step estimator, in the second step of our approach we estimate the factors in the error term, $\mathbf{F}_y^0$, using the residuals in the first step IV regression. Then we asymptotically eliminate $\mathbf{F}_y^0$ from the entire model by projecting them out from $\{\mathbf{y}_i, \mathbf{W}_i\}$ and use the instrumental variables $\mathbf{Z}_i$ to obtain the second step estimator. To portray the second step estimator, temporarily supposing $\mathbf{F}_y^0$ observable, define the projection matrix

$$\mathbf{M}_{F_y^0} = \mathbf{I}_T - \mathbf{F}_y^0 \left(\mathbf{F}_y^{0\prime} \mathbf{F}_y^0\right)^{-1} \mathbf{F}_y^{0\prime}. \tag{9}$$

Premultiplying the model (6) by $\mathbf{M}_{F_y^0}$ we obtain

$$\mathbf{M}_{F_y^0} \mathbf{y}_i = \mathbf{M}_{F_y^0} \mathbf{W}_i \boldsymbol{\theta} + \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i, \tag{10}$$

where the factor component $\mathbf{F}_y^0 \boldsymbol{\gamma}_{yi}^0$ in the error term is swept away. With a similar reasoning given in the first step estimation, we can easily see that $E(\mathbf{Z}_i' \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i) = \mathbf{0}$ and $E(\mathbf{Z}_i' \mathbf{M}_{F_y^0} \mathbf{W}_i) \neq$

[6] In general, for ARDL(p, q) models, $T \times (s+1)k$ instruments set $\left\{\mathbf{M}_{F_{x,-r}^0} \mathbf{X}_{i,-r}\right\}_{r=0}^{s}$, where $s = q + \lceil p/k \rceil$ with $\lceil . \rceil$ being the ceiling function, is necessary.

$\mathbf{0}$. Therefore, it is straightforward to apply instrumental variable (IV) estimation using $\mathbf{Z}_i$ to the transformed model in (10).[7]

In practice, the factors $\mathbf{F}_x^0$, $\mathbf{F}_{x,-1}^0$ and $\mathbf{F}_y^0$ are usually not observed. As will be discussed in detail below, we replace these factors with the estimated ones based on the principal components approach, as advanced in Bai (2003) and Bai (2009a), among many others.[8]

In this section and the next, we treat the number of factors, m_x and m_y as given. In practice, these should be estimated. m_x can be estimated from the raw data $\mathbf{x}_{it}$, $t = 0, ..., T$, $i = 1, ..., N$, by using the methods which have been proposed in the literature, such as information criteria of Bai and Ng (2002) and the eigenvalue methods of Ahn and Horenstein (2013). m_y can be estimated using the above mentioned methods from the residual covariance matrix.[9] In the Monte Carlo section below, we use the various existing methods to determine the number of factors, and it will be shown that these provide quite accurate determination of the number of factors with our experimental design.

Remark 3 Since our approach makes use of the transformed x's as instruments, identification of ρ requires that *at least one element in $\boldsymbol{\beta}$ is not equal to zero*, given model (3). We believe this is a mild condition, especially compared to a restriction that *all the elements in $\boldsymbol{\beta}$ are non-zero*. Specifically, identification of the autoregressive parameter can be achieved based on the covariate(s) and lagged value(s) corresponding to the non-zero slope coefficient(s). Notably, it is not necessary to know which covariates have non-zero coefficients since by construction the IV estimation procedure does not require that all instruments are relevant to all endogenous regressors.

Remark 4 More instruments potentially are available when further lags of $\mathbf{x}_{it}$ are observed. In particular, given model (3), when $\{\mathbf{x}_{it}\}_{t=1-j}^{T}$ for $j \geq 1$ are observable, $(j+1)k$ instruments can be used instead of (8):

$$\mathbf{Z}_i = \left(\mathbf{M}_{F_x^0}\mathbf{X}_i, \mathbf{M}_{F_{x,-1}^0}\mathbf{X}_{i,-1}, ..., \mathbf{M}_{F_{x,-j}^0}\mathbf{X}_{i,-j}\right) \ (T \times (j+1)k). \tag{11}$$

It is well documented in the literature that the larger the number of instruments, the more efficient but more biased the estimator will become. In this paper we assume a small finite number $j \geq 1$ which does not depend on sample size, T, in particular.[10] Without loss of generality, we set $j = 1$ for the theoretical analysis in Sections 2 and 3. In Section 4 we conduct finite sample experiment with different values of j.

To obtain our results it is sufficient to make the following assumptions, where $\text{tr}[\mathbf{A}]$ and $||\mathbf{A}|| = \sqrt{\text{tr}[\mathbf{A}'\mathbf{A}]}$ denote the trace and Frobenius (Euclidean) norm of matrix $\mathbf{A}$, respectively, and Δ is a finite positive constant.

Assumption 1 (idiosyncratic error in y): ε_{it} is independently distributed across i and t, with $E(\varepsilon_{it}) = 0$, $E(\varepsilon_{it}^2) = \sigma_{\varepsilon,it}^2$, and $E|\varepsilon_{it}|^{8+\delta} \leq \Delta < \infty$ for a small positive constant δ.

Assumption 2 (idiosyncratic error in x): (i) $v_{\ell it}$ is independently distributed across i and group-wise independent from ε_{it}; (ii) $E(v_{\ell it}) = 0$ and $E|v_{\ell it}|^{8+\delta} \leq \Delta < \infty$; (iii)

[7]This IV estimation is equivalent to the one using the transformed instrument set, $\mathbf{M}_{F_y^0}\mathbf{Z}_i$, for the original model (6).

[8]One could employ Pesaran's (2006) approach to estimate the common factors in the regressors.

[9]See Bai (2009b, C.3) for discussion on estimation of the number of factors in disturbances.

[10]The limit behaviour of the estimators when the number of instruments increases with T might be of theoretical interest, however, it is beyond the scope of this paper. See Alvarez and Arellano (2003) among others for a related analysis.

$T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}E\left|v_{\ell is}v_{\ell it}\right|^{1+\delta}\leq\Delta<\infty$; (iv) $E\left|N^{-1/2}\sum_{i=1}^{N}\left[v_{\ell is}v_{\ell it}-E\left(v_{\ell is}v_{\ell it}\right)\right]\right|^{4}\leq\Delta<\infty$ for every ℓ, t and s; (v) $N^{-1}T^{-2}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\sum_{r=1}^{T}\sum_{w=1}^{T}\left|\text{cov}\left(v_{\ell is}v_{\ell it},v_{\ell ir}v_{\ell iw}\right)\right|\leq\Delta<\infty$; (vi) the largest eigenvalue of $E\left(\mathbf{v}_{\ell i}\mathbf{v}_{\ell i}'\right)$ is bounded uniformly for every ℓ, i and T.

Assumption 3 (stationary factors): $\mathbf{f}_{x,t}^{0}=\boldsymbol{C}_{x}(L)\mathbf{e}_{f_x,t}$ and $\mathbf{f}_{y,t}^{0}=\boldsymbol{C}_{y}(L)\mathbf{e}_{f_y,t}$, where $\boldsymbol{C}_{x}(L)$ and $\boldsymbol{C}_{y}(L)$ are absolutely summable, $\mathbf{e}_{f_x,t}\sim i.i.d.(\mathbf{0},\boldsymbol{\Sigma}_{fx})$ and $\mathbf{e}_{f_y,t}\sim i.i.d.(\mathbf{0},\boldsymbol{\Sigma}_{fy})$, where $\boldsymbol{\Sigma}_{fx}$ and $\boldsymbol{\Sigma}_{fy}$ are positive definite matrices. Each element of $\mathbf{e}_{f_x,t}$ and $\mathbf{e}_{f_y,t}$ has finite fourth order moments and are group-wise independent from $\mathbf{v}_{it}$ and ε_{it}.

Assumption 4 (random factor loadings): $\boldsymbol{\Gamma}_{xi}^{0}\sim i.i.d.(\mathbf{0},\boldsymbol{\Sigma}_{\Gamma x})$, $\boldsymbol{\gamma}_{yi}^{0}\sim i.i.d.(\mathbf{0},\boldsymbol{\Sigma}_{\gamma y})$, where $\boldsymbol{\Sigma}_{\Gamma x}$ and $\boldsymbol{\Sigma}_{\gamma y}$ are positive definite matrices, and each element of $\boldsymbol{\Gamma}_{xi}^{0}$ and $\boldsymbol{\gamma}_{yi}^{0}$ has finite fourth order moments. $\boldsymbol{\Gamma}_{xi}^{0}$ and $\boldsymbol{\gamma}_{yi}^{0}$ are independent groups from ε_{it}, $\mathbf{v}_{it}$, $\mathbf{e}_{f_x,t}$ and $\mathbf{e}_{f_y,t}$.

Assumption 5 (identification of $\boldsymbol{\theta}$): (i) $\tilde{\mathbf{A}}_{i,T}=T^{-1}\mathbf{Z}_i'\mathbf{W}_i$, $\tilde{\mathbf{B}}_{i,T}=T^{-1}\mathbf{Z}_i'\mathbf{Z}_i$, $\mathbf{A}_{i,T}=T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_y^0}\mathbf{W}_i$ and $\mathbf{B}_{i,T}=T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_i$ have full column rank for all i for a sufficiently large T; (ii) $E\left\|\tilde{\mathbf{A}}_{i,T}\right\|^{2+2\delta}\leq\Delta<\infty$, $E\left\|\tilde{\mathbf{B}}_{i,T}\right\|^{2+2\delta}\leq\Delta<\infty$, $E\left\|\mathbf{A}_{i,T}\right\|^{2+2\delta}\leq\Delta<\infty$ and $E\left\|\mathbf{B}_{i,T}\right\|^{2+2\delta}\leq\Delta<\infty$ for all i for a sufficiently large T; (iii) $E\left\|\boldsymbol{\varphi}_{FiT}\right\|^{2+\delta}\leq\Delta<\infty$ for all i for a sufficiently large T, where $\boldsymbol{\varphi}_{FiT}=T^{-1/2}\mathbf{Z}_i{}'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i$, and $E(\boldsymbol{\varphi}_{FiT}\boldsymbol{\varphi}_{FiT}')$ is a positive definite matrix for all i for a sufficiently large T. In addition, $\lim_{N,T\to\infty}N^{-1}\sum_{i=1}^{N}E\left(\boldsymbol{\varphi}_{FiT}\boldsymbol{\varphi}_{FiT}'\right)=\boldsymbol{\Omega}$, which is a fixed positive definite matrix.

The assumptions above require some discussion. First of all, notice that Assumption 1 allows non-normality and (unconditional) times-series and cross-sectional heteroskedasticity in the idiosyncratic errors in the equation for y. Assumptions 2 and 3 allow for serial correlation in the idiosyncratic errors in the equation for x and the factors. Assumption 2 is in line with Bai (2003) but assumes independence across i, which can be relaxed such that the factors and $(\varepsilon_{it},\mathbf{v}_{it})$ and/or ε_{jt} and ε_{is} are weakly dependent, provided that there exist higher order moments; see Assumptions D-F in Bai (2003)[11]. Assumptions 3 and 4 are standard in the principal components literature; see e.g. Bai (2003) among others. Assumption 3 permits correlations between $\mathbf{f}_{x,t}^{0}$ and $\mathbf{f}_{y,t}^{0}$, and within each of them. Assumption 4 allows for possible non-zero correlations between $\boldsymbol{\gamma}_{yi}^{0}$ and $\boldsymbol{\Gamma}_{xi}^{0}$, and within each of them, which are the loadings associated with the factors $\mathbf{f}_{x,t}^{0}$ and $\mathbf{f}_{y,t}^{0}$. Since the variables y_{it} and $\mathbf{x}_{it}$ of the same individual unit i can be affected in a related manner by the common shocks, allowing for this possibility is potentially important in practice. Finally, Assumption 5(i)-(ii) is common in overidentified instrumental variable (IV) estimation; for example, see Wooldridge (2002, Ch5). Assumption 5(iii) is required for identification of the estimator, the consistency property of the variance-covariance estimator and the asymptotic normality of the estimator as N and T tend to infinity jointly.

Let us begin with the discussion of the first step IV estimator of our approach. Given m_x, the factors are extracted using principal components (PC) from $\{\mathbf{X}_i\}_{i=1}^{N}$. Define $\hat{\mathbf{F}}_x$ as $\sqrt{T}$ times the eigenvectors corresponding to the m_x largest eigenvalues of the $T\times T$ matrix $\sum_{i=1}^{N}\mathbf{X}_i\mathbf{X}_i'/NT$. $\hat{\mathbf{F}}_{x,-1}$ is defined in the same way, but this time based on $\sum_{i=1}^{N}\mathbf{X}_{i,-1}\mathbf{X}_{i,-1}'/NT$.

Remark 5 Note that $\mathbf{F}_x^0$, $\boldsymbol{\Gamma}_{xi}^0$, $\mathbf{F}_{x,-1}^0$, ($\mathbf{F}_y^0$ and $\boldsymbol{\gamma}_{yi}^0$) can be identified and estimated up to an invertible $m_x\times m_x$ (and $m_y\times m_y$) matrix transformation; see, Bai and Ng (2013), among others. For example, $\hat{\mathbf{F}}_x$ is a consistent estimator of $\mathbf{F}_x=\mathbf{F}_x^0\mathbf{G}_x$, where $\mathbf{G}_x$ is an invertible matrix which makes $\mathbf{F}_x'\mathbf{F}_x/T=\mathbf{I}_T$ and $\boldsymbol{\gamma}_{\ell i}=\mathbf{G}_x^{-1}\boldsymbol{\gamma}_{\ell i}^0$ such that $\sum_{\ell=1}^{k}\sum_{i=1}^{N}\boldsymbol{\gamma}_{\ell i}\boldsymbol{\gamma}_{\ell i}'$ is a diagonal

[11]This includes conditional heteroskedasticity, such as ARCH or GARCH processes.

matrix. We define $\mathbf{F}_{x,-1}$, ($\mathbf{F}_y$ and $\boldsymbol{\gamma}_{yi}$) in an analogous manner, which are estimated by the PC method.

The empirical counterparts of the projection matrices defined in (7) and (9) are given by

$$\mathbf{M}_{\hat{F}_x} = \mathbf{I}_T - \hat{\mathbf{F}}_x \left(\hat{\mathbf{F}}_x' \hat{\mathbf{F}}_x\right)^{-1} \hat{\mathbf{F}}_x'; \ \mathbf{M}_{\hat{F}_{x,-1}} = \mathbf{I}_T - \hat{\mathbf{F}}_{x,-1} \left(\hat{\mathbf{F}}_{x,-1}' \hat{\mathbf{F}}_{x,-1}\right)^{-1} \hat{\mathbf{F}}_{x,-1}'. \tag{12}$$

The associated transformed instrument matrix discussed above is

$$\hat{\mathbf{Z}}_i = \left(\mathbf{M}_{\hat{F}_x} \mathbf{X}_i, \mathbf{M}_{\hat{F}_{x,-1}} \mathbf{X}_{i,-1}\right). \tag{13}$$

We propose the following first-step instrumental variable (IV) estimator of $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}}_{IV} = \left(\hat{\hat{\mathbf{A}}}_{NT}' \hat{\hat{\mathbf{B}}}_{NT}^{-1} \hat{\hat{\mathbf{A}}}_{NT}\right)^{-1} \hat{\hat{\mathbf{A}}}_{NT}' \hat{\hat{\mathbf{B}}}_{NT}^{-1} \hat{\hat{\mathbf{g}}}_{NT}, \tag{14}$$

where

$$\hat{\hat{\mathbf{A}}}_{NT} = \frac{1}{NT} \sum_{i=1}^{N} \hat{\mathbf{Z}}_i' \mathbf{W}_i, \ \hat{\hat{\mathbf{B}}}_{NT} = \frac{1}{NT} \sum_{i=1}^{N} \hat{\mathbf{Z}}_i' \hat{\mathbf{Z}}_i, \ \hat{\hat{\mathbf{g}}}_{NT} = \frac{1}{NT} \sum_{i=1}^{N} \hat{\mathbf{Z}}_i' \mathbf{y}_i. \tag{15}$$

Firstly we show consistency of the above estimator. To begin with, from (6) and (14) we obtain

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}\right) = \left(\hat{\hat{\mathbf{A}}}_{NT}' \hat{\hat{\mathbf{B}}}_{NT}^{-1} \hat{\hat{\mathbf{A}}}_{NT}\right)^{-1} \hat{\hat{\mathbf{A}}}_{NT}' \hat{\hat{\mathbf{B}}}_{NT}^{-1} \left(\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \hat{\mathbf{Z}}_i' \mathbf{u}_i\right). \tag{16}$$

Since the asymptotic properties of the estimator are primarily determined by those of $\sum_{i=1}^{N} \hat{\mathbf{Z}}_i' \mathbf{u}_i / \sqrt{NT}$, we focus on this term. The formal analysis is provided as a proposition below, where $(N,T) \xrightarrow{j} \infty$ signifies that N and T tend to infinity jointly.

Proposition 1 *Under Assumptions 1-5, as* $(N,T) \xrightarrow{j} \infty$ *such that* $N/T \to c$ *with* $0 < c < \infty$,

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \hat{\mathbf{Z}}_i' \mathbf{u}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \tilde{\mathbf{Z}}_i' \mathbf{u}_i + \sqrt{\frac{T}{N}} \mathbf{b}_{1NT} + \sqrt{\frac{N}{T}} \mathbf{b}_{2NT} + o_p(1),$$

where $\hat{\mathbf{Z}}_i$ *is defined by (13),* $\tilde{\mathbf{Z}}_i = (\mathbf{M}_{F_x^0} \tilde{\mathbf{X}}_i, \mathbf{M}_{F_{x,-1}^0} \tilde{\mathbf{X}}_{i,-1})$, $\tilde{\mathbf{X}}_i = \mathbf{X}_i - \frac{1}{N} \sum_{n=1}^{N} \mathbf{X}_n \boldsymbol{\Gamma}_{xn}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \boldsymbol{\Gamma}_{xi}^0$, $\tilde{\mathbf{X}}_{i,-1} = \mathbf{X}_{i,-1} - \frac{1}{N} \sum_{n=1}^{N} \mathbf{X}_{n,-1} \boldsymbol{\Gamma}_{xn}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \boldsymbol{\Gamma}_{xi}^0$, $\boldsymbol{\Upsilon}_{xkN}^0 = \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \boldsymbol{\gamma}_{\ell i}^0 \boldsymbol{\gamma}_{\ell i}^{0\prime}$ *and* $\mathbf{b}_{1NT} = [\mathbf{b}_{11NT}', \mathbf{b}_{12NT}']'$, $\mathbf{b}_{2NT} = [\mathbf{b}_{21NT}', \mathbf{b}_{22NT}']'$ *with*

$$\mathbf{b}_{11NT} = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\tilde{\mathbf{V}}_i' \mathbf{V}_j}{T} \boldsymbol{\Gamma}_{xj}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \left(\frac{\mathbf{F}_x' \mathbf{F}_x^0}{T}\right)^{-1} \frac{\mathbf{F}_x^{0\prime} \mathbf{u}_i}{T};$$

$$\mathbf{b}_{12NT} = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\tilde{\mathbf{V}}_{i,-1}' \mathbf{V}_{j,-1}}{T} \boldsymbol{\Gamma}_{xj}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \left(\frac{\mathbf{F}_{x,-1}^{0\prime} \mathbf{F}_{x,-1}^0}{T}\right)^{-1} \frac{\mathbf{F}_{x,-1}^{0\prime} \mathbf{u}_i}{T};$$

$$\mathbf{b}_{21NT} = -\frac{1}{NT} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T}\right)^{-1} \mathbf{F}_x^{0\prime} \bar{\boldsymbol{\Sigma}}_{kNT} \mathbf{M}_{F_x^0} \mathbf{u}_i;$$

$$\mathbf{b}_{22NT} = -\frac{1}{NT} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \left(\frac{\mathbf{F}_{x,-1}^{0\prime} \mathbf{F}_{x,-1}^0}{T}\right)^{-1} \mathbf{F}_{x,-1}^{0\prime} \bar{\boldsymbol{\Sigma}}_{kNT,-1} \mathbf{M}_{F_{x,-1}^0} \mathbf{u}_i,$$

$\tilde{\mathbf{V}}_i = \mathbf{V}_i - \frac{1}{N} \sum_{n=1}^{N} \mathbf{V}_n \boldsymbol{\Gamma}_{xn}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \boldsymbol{\Gamma}_{xi}^0$, $\tilde{\mathbf{V}}_{i,-1} = \mathbf{V}_{i,-1} - \frac{1}{N} \sum_{n=1}^{N} \mathbf{V}_{n,-1} \boldsymbol{\Gamma}_{xn}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \boldsymbol{\Gamma}_{xi}^0$, $\bar{\boldsymbol{\Sigma}}_{kNT} = \frac{1}{N} \sum_{\ell=1}^{k} \sum_{j=1}^{N} E\left(\mathbf{v}_{\ell j} \mathbf{v}_{\ell j}'\right)$ *and* $\bar{\boldsymbol{\Sigma}}_{kNT,-1} = \frac{1}{N} \sum_{\ell=1}^{k} \sum_{j=1}^{N} E\left(\mathbf{v}_{\ell j,-1} \mathbf{v}_{\ell j,-1}'\right)$.

Remark 6 The source of the bias term in Proposition 1 is different than the bias terms reported in Bai (2009a) and Moon and Weidner (2017). In particular, the bias term of our estimator arises primarily due to the correlation between the factor loadings associated with $\mathbf{F}_x$ in x and the error term in the equation of y, $\mathbf{u}_i$. On the other hand, the two bias terms in Bai (2009a) and Moon and Weidner (2017) arise from error serial dependence and weak cross-sectional dependence. In our case, error serial correlation in the idiosyncratic part of the x process, $v_{\ell it}$, does not result in bias because $v_{\ell it}$ is not correlated with the error term in the y equation, ε_{it}. Also note that Moon and Weidner (2017) report additional bias term that generalises the small T bias, called "Nickell bias," which typically occurs in the least square estimation of dynamic panel models. Our estimator is not subject to such a bias as it is based on instrumental variable method.

From the result stated in Proposition 1 it can be shown that $\sum_{i=1}^{N}\tilde{\mathbf{Z}}_i'\mathbf{u}_i/\sqrt{NT}$ is $O_p(1)$ and tends to a multivariate distribution. In addition, $\sqrt{T/N}\mathbf{b}_{1NT}$ and $\sqrt{N/T}\mathbf{b}_{2NT}$ are $O_p(1)$ as $(N,T)\xrightarrow{j}\infty$ such that $N/T\to c$ with $0<c<\infty$. Therefore, in such situation the IV estimator is $\sqrt{NT}$-consistent.

The above discussion is formally summarised in the following theorem:

Theorem 1 *Consider model (1)-(3) and suppose that Assumptions 1-5 hold true. Then,*

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{IV}-\boldsymbol{\theta}\right)=O_p\left(1\right)$$

as $(N,T)\xrightarrow{j}\infty$ *such that* $N/T\to c$ *with* $0<c<\infty$*, where* $\hat{\boldsymbol{\theta}}_{IV}$ *is defined in (14).*

Even though the estimator $\hat{\boldsymbol{\theta}}_{IV}$ is $\sqrt{NT}$-consistent, under our assumptions the limiting distribution of $\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{IV}-\boldsymbol{\theta}\right)$ will contain asymptotic bias terms such as the limits of $\mathbf{b}_{1NT}$ and $\mathbf{b}_{2NT}$, which are defined in Proposition 1.[12] Rather than bias-correcting this estimator, we propose to compute a potentially more efficient second-step estimator, by asymptotically projecting out $\mathbf{F}_y$ from the model using $\hat{\boldsymbol{\theta}}_{IV}$.

To compute the second step estimator, firstly the factors $\mathbf{F}_y$ are estimated using principal components from $\{\hat{\mathbf{u}}_i\}_{i=1}^{N}$, where $\hat{\mathbf{u}}_i=\mathbf{y}_i-\mathbf{W}_i\hat{\boldsymbol{\theta}}_{IV}$ with $\hat{\boldsymbol{\theta}}_{IV}$ being a first-step IV estimator defined in (14). We define $\hat{\mathbf{F}}_y$ as $\sqrt{T}$ times the eigenvectors corresponding to the m_y largest eigenvalues of the $T\times T$ matrix $\sum_{i=1}^{N}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'/NT$.

The (sub-optimal) second step IV estimator is defined as

$$\hat{\hat{\boldsymbol{\theta}}}_{IV}=\left(\hat{\mathbf{A}}_{NT}'\hat{\mathbf{B}}_{NT}^{-1}\hat{\mathbf{A}}_{NT}\right)^{-1}\hat{\mathbf{A}}_{NT}'\hat{\mathbf{B}}_{NT}^{-1}\hat{\mathbf{g}}_{NT}, \tag{17}$$

where

$$\hat{\mathbf{A}}_{NT}=\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{W}_i,\ \hat{\mathbf{B}}_{NT}=\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{Z}}_i,\ \hat{\mathbf{g}}_{NT}=\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{y}_i \tag{18}$$

with

$$\mathbf{M}_{\hat{F}_y}=\mathbf{I}_T-\hat{\mathbf{F}}_y\left(\hat{\mathbf{F}}_y'\hat{\mathbf{F}}_y\right)^{-1}\hat{\mathbf{F}}_y'. \tag{19}$$

In order to show consistency of $\hat{\hat{\boldsymbol{\theta}}}_{IV}$, we use again (6) and (17) to obtain:

$$\sqrt{NT}\left(\hat{\hat{\boldsymbol{\theta}}}_{IV}-\boldsymbol{\theta}\right)=\left(\hat{\mathbf{A}}_{NT}'\hat{\mathbf{B}}_{NT}^{-1}\hat{\mathbf{A}}_{NT}\right)^{-1}\hat{\mathbf{A}}_{NT}'\hat{\mathbf{B}}_{NT}^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{u}_i\right). \tag{20}$$

[12] Of course we could consistently estimate these bias terms.

The asymptotic property of the key term $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{u}_i$ in (20) is stated in the following proposition.

Proposition 2 *Under Assumptions 1-5, as $(N,T)\xrightarrow{j}\infty$ such that $N/T\rightarrow c$ with $0<c<\infty$,*

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{u}_i=\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{Z}_i'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i+o_p(1),$$

where $\hat{\mathbf{Z}}_i$ is defined by (13).

From Proposition 2 we see that the estimation effect in $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{u}_i$ is asymptotically ignorable. Since $\boldsymbol{\varepsilon}_i$ is independent of $\mathbf{Z}_i$ and $\mathbf{F}_y^0$ with zero mean, the limiting distribution of $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{u}_i$ is centred at zero. The following theorem provides asymptotic normality of the distribution of $\hat{\hat{\boldsymbol{\theta}}}_{IV}$, based on Hansen's (2007) law of large numbers and central limit theorem, which are restated as Lemmas 1 and 2 in Appendix A.

Theorem 2 *Suppose that Assumptions 1-5 hold true under model (1)-(3). Then, as $(N,T)\xrightarrow{j}\infty$ such that $N/T\rightarrow c$ with $0<c<\infty$,*

(i)

$$\sqrt{NT}\left(\hat{\hat{\boldsymbol{\theta}}}_{IV}-\boldsymbol{\theta}\right)\xrightarrow{d}N\left(\mathbf{0},\boldsymbol{\Psi}\right),$$

where $\hat{\hat{\boldsymbol{\theta}}}_{IV}$ is defined by (17) and

$$\boldsymbol{\Psi}=\left(\mathbf{A}'\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}'\mathbf{B}^{-1}\boldsymbol{\Omega}\mathbf{B}^{-1}\mathbf{A}\left(\mathbf{A}'\mathbf{B}^{-1}\mathbf{A}\right)^{-1}$$

is a positive definite matrix, $\mathbf{A}=\operatorname{plim}_{N,T\rightarrow\infty}\hat{\mathbf{A}}_{NT}$ and $\mathbf{B}=\operatorname{plim}_{N,T\rightarrow\infty}\hat{\mathbf{B}}_{NT}$ with $\hat{\mathbf{A}}_{NT}$ and $\hat{\mathbf{B}}_{NT}$ defined in (18), and $\boldsymbol{\Omega}$ is defined in Assumption 5.

(ii) $\hat{\boldsymbol{\Psi}}_{NT}-\boldsymbol{\Psi}\xrightarrow{p}\mathbf{0}$, where

$$\hat{\boldsymbol{\Psi}}_{NT}=\left(\hat{\mathbf{A}}_{NT}'\hat{\mathbf{B}}_{NT}^{-1}\hat{\mathbf{A}}_{NT}\right)^{-1}\hat{\mathbf{A}}_{NT}'\hat{\mathbf{B}}_{NT}^{-1}\hat{\boldsymbol{\Omega}}_{NT}\hat{\mathbf{B}}_{NT}^{-1}\hat{\mathbf{A}}_{NT}\left(\hat{\mathbf{A}}_{NT}'\hat{\mathbf{B}}_{NT}^{-1}\hat{\mathbf{A}}_{NT}\right)^{-1}, \tag{21}$$

with

$$\hat{\boldsymbol{\Omega}}_{NT}=\frac{1}{NT}\sum_{i=1}^{N}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{Z}}_i \tag{22}$$

and $\hat{\mathbf{u}}_i=\mathbf{y}_i-\mathbf{W}_i\hat{\boldsymbol{\theta}}_{IV}$.

Finally we propose the optimal second step estimator, which we recommend to use:

$$\hat{\hat{\boldsymbol{\theta}}}_{IV2}=\left(\hat{\mathbf{A}}_{NT}'\hat{\boldsymbol{\Omega}}_{NT}^{-1}\hat{\mathbf{A}}_{NT}\right)^{-1}\hat{\mathbf{A}}_{NT}'\hat{\boldsymbol{\Omega}}_{NT}^{-1}\hat{\mathbf{g}}_{NT}, \tag{23}$$

where $\hat{\mathbf{A}}_{NT}$ and $\hat{\mathbf{B}}_{NT}$ are defined in (18), and $\hat{\boldsymbol{\Omega}}_{NT}$ is given in (22). The following corollary describes the asymptotic properties of the estimator:

Corollary 1 *Suppose that Assumptions 1-5 hold true under model (1)-(3). Then, as $(N,T)\xrightarrow{j}\infty$ such that $N/T\rightarrow c$ with $0<c<\infty$,*

$$\sqrt{NT}\left(\hat{\hat{\boldsymbol{\theta}}}_{IV2}-\boldsymbol{\theta}\right)\xrightarrow{d}N\left(\mathbf{0},\left(\mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{A}\right)^{-1}\right)$$

and

$$\hat{\mathbf{A}}_{NT}'\hat{\boldsymbol{\Omega}}_{NT}^{-1}\hat{\mathbf{A}}_{NT}-\mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{A}\xrightarrow{p}\mathbf{0},$$

where $\hat{\hat{\boldsymbol{\theta}}}_{IV2}$ is defined by (23), $\mathbf{A}=\operatorname{plim}_{N,T\rightarrow\infty}\hat{\mathbf{A}}_{NT}$ and $\boldsymbol{\Omega}$ is defined in Assumption 5.

The associated overidentifying restrictions test statistic is given by

$$S_{NT} = \frac{1}{NT}\left(\sum_{i=1}^{N} \hat{\mathbf{u}}_i' \mathbf{M}_{\hat{F}_y} \hat{\mathbf{Z}}_i\right) \hat{\boldsymbol{\Omega}}_{NT}^{-1} \left(\sum_{i=1}^{N} \hat{\mathbf{Z}}_i' \mathbf{M}_{\hat{F}_y} \hat{\mathbf{u}}_i\right), \tag{24}$$

where $\hat{\hat{\mathbf{u}}}_i = \mathbf{y}_i - \mathbf{W}_i \hat{\hat{\boldsymbol{\theta}}}_{IV2}$, and $\hat{\boldsymbol{\Omega}}_{NT}$ is defined by (22). The limit distribution of the overidentifying restrictions test statistic is established in the following theorem:

Theorem 3 *Suppose that Assumptions 1-5 hold true under model (1)-(3). Then, as* $(N, T) \xrightarrow{j} \infty$ *such that* $N/T \to c$ *with* $0 < c < \infty$*,*

$$S_{NT} \xrightarrow{d} \chi^2_{k-1}, \tag{25}$$

for $k > 1$*, where* S_{NT} *is defined in (24).*

Remark 7 The overidentifying restrictions test is particularly useful in our approach. Firstly, it is expected to pick up the violation of exogeneity assumption on the idiosyncratic error in the equation for x. Secondly, if the slope vector, $\boldsymbol{\theta}$, is cross-sectionally heterogeneous, the orthogonality property of the instruments may be lost hence the proposed estimators in this section may become inconsistent. In such a case the test is expected to reject the null hypothesis.

In the next section we discuss estimation of models with heterogeneous slope coefficients.

3 The Model with Heterogeneous Coefficients

So far we have discussed estimation of the model with homogeneous slopes. In this section we consider a model where the coefficients are heterogeneous across i:

$$\mathbf{y}_i = \mathbf{W}_i \boldsymbol{\theta}_i + \mathbf{u}_i, \tag{26}$$

where $\mathbf{W}_i = (\mathbf{y}_{i,-1}, \mathbf{X}_i)$ with $\mathbf{X}_i$ follows the factor structure defined by (5), $\boldsymbol{\theta}_i = (\rho_i, \boldsymbol{\beta}_i')'$ with $\sup_{1\leq i\leq N} |\rho_i| < 1$ and $\mathbf{u}_i$ is defined by (4). It is widely known that, for dynamic panel data models, the pooled estimator, including $\hat{\hat{\boldsymbol{\theta}}}_{IV2}$, will be inconsistent to, say, $\boldsymbol{\theta} = E(\boldsymbol{\theta}_i)$, if the slopes are cross-sectionally heterogeneous.[13] Henceforth, we introduce an estimator of $\boldsymbol{\theta}_i$ then propose a mean group IV estimator of the population average of $\boldsymbol{\theta}_i$ and establish its consistency and asymptotic normality.

To begin with, we introduce the following additional assumptions about the heterogeneous slopes, $\boldsymbol{\theta}_i$:

Assumption 6 (random coefficients): (i) $\boldsymbol{\theta}_i = \boldsymbol{\theta} + \boldsymbol{\eta}_i$, $\boldsymbol{\eta}_i \sim i.i.d.\,(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$, where $\boldsymbol{\Sigma}_\eta$ is a fixed positive definite matrix; (ii) $\boldsymbol{\eta}_i$ is independent of $\boldsymbol{\Gamma}^0_{xi}$, γ^0_{yi}, ε_{it}, $\mathbf{v}_{it}$, $\mathbf{e}_{f_x,t}$ and $\mathbf{e}_{f_y,t}$; (iii) $\boldsymbol{\eta}_i$ satisfies the tail bound:

$$P(|\eta_{ir}| > z) \leq 2\exp\left(-\frac{1}{2} \times \frac{z^2}{a + bz}\right)$$

for all z (and all i) and fixed $a, b > 0$, where η_{ir} is the r-th element of $\boldsymbol{\eta}_i$ for $2 \leq r \leq k+1$.

[13]See Pesaran and Smith (1995).

Assumption 7 (moment condition): (i) $E\|\boldsymbol{\eta}_i\|^4 \leq \Delta$; (ii) $E\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|^4 \leq \Delta$;
(iii) $E\|N^{-1/2}T^{-1/2}\sum_{\ell=1}^{k}\sum_{j=1}^{N}(\mathbf{V}_i'\mathbf{v}_{\ell j} - E(\mathbf{V}_i'\mathbf{v}_{\ell j}))\boldsymbol{\gamma}_{\ell j}^{0\prime}\|^4 \leq \Delta$; In addition,
(iv) $E(T^{-1/2}\sum_{\ell=1}^{k}\sum_{t=1}^{T}(v_{\ell it}^2 - Ev_{\ell it}^2))^2 \leq \Delta$.

Assumption 8 (identification of $\boldsymbol{\theta}_i$): $\mathbf{A}_i = p\lim_{T\to\infty}\tilde{\mathbf{A}}_{i,T}$ has full column rank, $\mathbf{B}_i = p\lim_{T\to\infty}\tilde{\mathbf{B}}_{i,T}$ and $\boldsymbol{\Sigma}_i = p\lim_{T\to\infty}T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i\mathbf{u}_i'\mathbf{M}_{F_x^0}\mathbf{Z}_i'$ are positive definite, uniformly.

Assumptions 6(i)-(ii) are standard in the random coefficient literature; see, for example, Pesaran (2006). Assumptions 6(iii), 7 and 8 are required for the estimators of $\boldsymbol{\theta}_i$ tending to their limiting distributions, uniformly.

Now it is ready to introduce the first-step IV estimator of $\boldsymbol{\theta}_i$:

$$\hat{\boldsymbol{\theta}}_{IV,i} = \left(\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}\right)^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{g}}}_{i,T}, \tag{27}$$

where

$$\hat{\tilde{\mathbf{A}}}_{i,T} = \frac{1}{T}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{W}_i,\ \hat{\tilde{\mathbf{B}}}_{i,T} = \frac{1}{T}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\hat{\mathbf{Z}}_i,\ \hat{\tilde{\mathbf{g}}}_{i,T} = \frac{1}{T}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{y}_i. \tag{28}$$

We can see from (28) that the instrument set used here is $\mathbf{M}_{\hat{F}_x}\hat{\mathbf{Z}}_i$. It is equivalent to making use of $\hat{\mathbf{Z}}_i$ for the model (26) premultiplied by $\mathbf{M}_{\hat{F}_x}$, which is expected to produce more efficient first step IV estimator of $\boldsymbol{\theta}_i$ if the span of $\mathbf{F}_y^0$ includes a subset of $\mathbf{F}_x^0$.[14] Using (26) and (27) we have

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i\right) = \left(\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}\right)^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\left(T^{-1/2}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i\right). \tag{29}$$

The limiting property of the $T^{-1/2}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i$ is given by the following proposition.

Proposition 3 *Consider the model in equation (26). Under Assumptions 1–6, as $(N,T)\xrightarrow{j}\infty$ such that $N/T\to c$ with $0<c<\infty$, we have*

$$T^{-1/2}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i = T^{-1/2}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i + \sqrt{T}O_p\left(\delta_{NT}^{-2}\right),$$

where $\hat{\mathbf{Z}}_i$, $\mathbf{M}_{\hat{F}_x}$ and $\mathbf{Z}_i$ are defined by (8), (12) and (13), respectively, and $\delta_{NT} = \min\left\{\sqrt{T},\sqrt{N}\right\}$.

Using the result stated in Proposition 3 we see that $T^{-1/2}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i$ is $O_p(1)$ and tends to a random vector as $(N,T)\xrightarrow{j}\infty$ such that $N/T\to c$ with $0<c<\infty$. The formal result is summarised in the theorem below:

Theorem 4 *Consider model (26) and suppose that Assumptions 1–8 hold true. Then, as $(N,T)\xrightarrow{j}\infty$ such that $N/T\to c$ with $0<c<\infty$, for each i,*

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i\right)\xrightarrow{d} N\left(\mathbf{0}, (\mathbf{A}_i'\mathbf{B}_i^{-1}\mathbf{A}_i)^{-1}\mathbf{A}_i'\mathbf{B}_i^{-1}\boldsymbol{\Sigma}_i\mathbf{B}_i^{-1}\mathbf{A}_i(\mathbf{A}_i'\mathbf{B}_i^{-1}\mathbf{A}_i)^{-1}\right) \tag{30}$$

where $\hat{\boldsymbol{\theta}}_{IV,i}$ is defined in (27), and $\mathbf{A}_i$, $\mathbf{B}_i$ and $\boldsymbol{\Sigma}_i$ are defined in Assumption 8.

[14] We could construct the first-step IV estimator of $\boldsymbol{\theta}$ in Section 2 using $\mathbf{M}_{\hat{F}_x}\hat{\mathbf{Z}}_i$ instead of $\hat{\mathbf{Z}}_i$, however, the second-step estimator will be asymptotically equivalent to the proposed one which is based on $\mathbf{M}_{\hat{F}_y}\hat{\mathbf{Z}}_i$ when the span of $\mathbf{F}_y^0$ includes a subset of $\mathbf{F}_x^0$.

Therefore, the estimator $\hat{\boldsymbol{\theta}}_{IV,i}$ is $\sqrt{T}$-consistent to $\boldsymbol{\theta}_i$.

Using a similar line of the discussion for the IV estimator in Section 2, we could consider a mean group IV estimator using the second step estimator, attempting to asymptotically project out $\mathbf{F}_y^0$ from the model, i.e., $\mathbf{M}_{F_y^0}\mathbf{y}_i = \mathbf{M}_{F_y^0}\mathbf{W}_i\boldsymbol{\theta}_i + \mathbf{M}_{F_y^0}\mathbf{u}_i$, then apply our IV method to estimate $\boldsymbol{\theta}_i$. However, to deal with the heterogeneous slopes, $\mathbf{F}_y$ should be estimated using the residuals from the time series IV regression, $\hat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{W}_i\hat{\boldsymbol{\theta}}_{IV,i}$. Since $\hat{\boldsymbol{\theta}}_{IV,i}$ is $\sqrt{T}$-consistent, not $\sqrt{NT}$-consistent, the estimation of $\mathbf{F}_y$ may become very inefficient. Due to this we will not pursue such an estimator here. Note that the estimation of $\mathbf{F}_x$ for the first step IV estimator does not suffer from a similar problem, because it can be estimated using the raw data $\{\mathbf{X}_i\}_{i=1}^N$.

Now define the mean group estimator of $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}}_{IVMG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\boldsymbol{\theta}}_{IV,i}, \tag{31}$$

where $\hat{\boldsymbol{\theta}}_{IV,i}$ is given in (27). From (26), (29) and Assumptions 1-8 it can be shown that[15]

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}}_{IVMG} - \boldsymbol{\theta}\right) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\eta}_i + o_p(1). \tag{32}$$

It is easily seen that $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\eta}_i \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_\eta)$ as $N \to \infty$, which implies that $\hat{\boldsymbol{\theta}}_{IVMG}$ is $\sqrt{N}$-consistent. The asymptotic normality of $\hat{\boldsymbol{\theta}}_{IVMG}$ and the consistency of an estimator of $\boldsymbol{\Sigma}_\eta$ are summarised in the following theorem:

Theorem 5 *Consider model (26) with (5) and suppose that Assumptions 1–7 hold true. Then, as* $(N,T) \xrightarrow{j} \infty$ *such that* $N/T \to c$ *with* $0 < c < \infty$,

(i)

$$\sqrt{N}\left(\hat{\boldsymbol{\theta}}_{IVMG} - \boldsymbol{\theta}\right) \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_\eta), \tag{33}$$

where $\hat{\boldsymbol{\theta}}_{IVMG}$ *is defined in (31);*

(ii)

$$\widehat{\boldsymbol{\Sigma}}_\eta - \boldsymbol{\Sigma}_\eta \xrightarrow{p} \mathbf{0} \tag{34}$$

where

$$\widehat{\boldsymbol{\Sigma}}_\eta = \frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\boldsymbol{\theta}}_{IV,i} - \hat{\boldsymbol{\theta}}_{IVMG}\right)\left(\hat{\boldsymbol{\theta}}_{IV,i} - \hat{\boldsymbol{\theta}}_{IVMG}\right)', \tag{35}$$

$\hat{\boldsymbol{\theta}}_{IV,i}$ *and* $\hat{\boldsymbol{\theta}}_{IVMG}$ *are given by (27) and (31), respectively.*

4 Monte Carlo Experiments

This section investigates the finite sample behaviour of the proposed estimators by means of Monte Carlo experiments, based on bias, root mean squared error (RMSE), empirical size and power of the t-test. In particular, we examine the optimal two-step IV estimator (IV2), which is defined in (23), and the mean group IV estimator (IVMG), defined in (31). To see the effects of choice of the number of instruments (see Remark 4), we consider two sets of instruments for IV2 and IVMG, $\hat{\mathbf{Z}}_i$:

$$\text{IV set } a\text{: } \left(\mathbf{M}_{\hat{F}_x}\mathbf{X}_i; \mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1}\right)\ (T \times 2k)$$

[15] See the proof of Theorem 5.

$$\text{IV set } b\text{: } \left(\mathbf{M}_{\hat{F}_x}\mathbf{X}_i, \mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1}, \mathbf{M}_{\hat{F}_{x,-2}}\mathbf{X}_{i,-2}\right)\ (T\times 3k). \tag{36}$$

The instrument sets used for the IV estimators are denoted by the superscripts a and b (e.g. IV2^a).

For the purposes of comparison, we also investigate the performance of the bias-corrected quasi maximum likelihood estimator (BC-QMLE) recently proposed by Moon and Weidner (2017), as well as the CCE mean group (CCEMG) estimator and its bias-corrected version (BC-CCEMG), put forward by Chudik and Pesaran (2015a).

The bias-corrected QMLE estimator, $\hat{\boldsymbol{\theta}}_{BC-QMLE}$, is defined as[16]

$$\hat{\boldsymbol{\theta}}_{BC-QMLE} = \hat{\boldsymbol{\theta}}_{QMLE} - \hat{\boldsymbol{b}}_{QMLE}, \tag{37}$$

where

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{QMLE} &= \arg\min_{\boldsymbol{\theta}\in\Theta} L_{NT}(\boldsymbol{\theta}); \\ L_{NT}(\boldsymbol{\theta}) &= \min_{\boldsymbol{\Gamma}_y, \mathbf{F}_y} \mathcal{L}_{NT}(\boldsymbol{\theta}, \boldsymbol{\Gamma}_y, \mathbf{F}_y); \\ \mathcal{L}_{NT}(\boldsymbol{\theta}, \boldsymbol{\Gamma}_y, \mathbf{F}_y) &= \min_{\mathbf{F}_y} \frac{1}{NT}\sum_{i=1}^{N} (\mathbf{y}_i - \mathbf{W}_i\boldsymbol{\theta})' \mathbf{M}_{F_y} (\mathbf{y}_i - \mathbf{W}_i\boldsymbol{\theta}), \end{aligned}$$

with $\boldsymbol{\Gamma}_y = \left(\boldsymbol{\gamma}_{y,1}, \ldots, \boldsymbol{\gamma}_{y,N}\right)'$, whereas the estimator of the bias, $\hat{\boldsymbol{b}}_{QMLE}$, is defined in Definition 1 in Moon and Weidner (2017). The "t-test" in our experiment is computed using the estimator of the variance-covariance matrix for $\hat{\boldsymbol{\theta}}_{BC-QMLE}$ (Moon and Weidner, 2017; p.174). It should be highlighted that Moon and Weidner (2017) do not assume a linear factor process in $\mathbf{x}_{it}$, which is specified by (3), hence they may permit more general processes for the covariates.

The CCEMG estimator is given by

$$\hat{\boldsymbol{\theta}}_{CCEMG} = N^{-1}\sum_{i=1}^{N} \hat{\boldsymbol{\theta}}_{CCE,i}, \tag{38}$$

where $\hat{\boldsymbol{\theta}}_{CCE,i} = (\mathbf{W}_i'\mathbf{M}_{\bar{H}}\mathbf{W}_i)^{-1}\mathbf{W}_i'\mathbf{M}_{\bar{H}}\mathbf{y}_i$, $\mathbf{M}_{\bar{H}} = \mathbf{I}_T - \bar{\mathbf{H}}\left(\bar{\mathbf{H}}'\bar{\mathbf{H}}\right)^{-1}\bar{\mathbf{H}}'$, $\bar{\mathbf{H}} = N^{-1}\sum_{i=1}^{N}\mathbf{H}_i$. $\mathbf{H}_i$ contains $(\mathbf{y}_i; \mathbf{X}_i)$ and their lags:

$$\mathbf{H}_i = \left(\mathbf{y}_i, \mathbf{y}_{i,-1}, ..., \mathbf{y}_{i,-p_y}, \mathbf{X}_i, \mathbf{X}_{i,-1}, ..., \mathbf{X}_{i,-p_x}, \boldsymbol{\iota}_T\right), \tag{39}$$

where $\boldsymbol{\iota}_T$ is a $T\times 1$ vector of ones, $\mathbf{y}_{i,-j} = L^j\mathbf{y}_i$ and $\mathbf{X}_{i,-j} = L^j\mathbf{X}_i$. In view of the strict exogeneity of $\mathbf{X}_i$ in our experimental design, which is discussed shortly below, $p_y = p$ lags of $\mathbf{y}_i$ are included but no lags of $\mathbf{X}_i$, namely $p_x = 0$ in $\mathbf{H}_i$; see Chudik and Pesaran (2015a, equation 38).[17] Following Chudik and Pesaran (2015a) we choose the integer part of $T^{1/3}$ as the value of p. The "t-test" is computed using the estimated variance-covariance matrix, $\widehat{\boldsymbol{\Sigma}}_{MGCCE} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\boldsymbol{\theta}}_{CCE,i} - \hat{\boldsymbol{\theta}}_{MGCCE}\right)\left(\hat{\boldsymbol{\theta}}_{CCE,i} - \hat{\boldsymbol{\theta}}_{MGCCE}\right)'$. The bias-corrected CCEMG estimator, $\hat{\boldsymbol{\theta}}_{BC-CCEMG}$, is given by

$$\hat{\boldsymbol{\theta}}_{BC-CCEMG} = 2\hat{\boldsymbol{\theta}}_{CCEMG} - \frac{1}{2}\left(\hat{\boldsymbol{\theta}}^{(1)}_{CCEMG} + \hat{\boldsymbol{\theta}}^{(2)}_{CCEMG}\right), \tag{40}$$

[16] We are grateful to Martin Weidner for providing us the computational algorithm for the BC-QMLE estimator.

[17] We have also considered $p_y = p_x = p$, such that $\mathbf{y}_i$ and $\mathbf{X}_i$ have the same number of lags in $\mathbf{H}_i$. The performance of the CCEMG estimator is slightly worse in this case. The results are reported in Tables C11-C12 in Appendix C.

where $\hat{\boldsymbol{\theta}}^{(1)}_{MGCCE}$ denotes the mean group CCE estimator computed from the first half of the available time period and $\hat{\boldsymbol{\theta}}^{(2)}_{MGCCE}$ from the second half. See Chudik and Pesaran (2015a) for more details.[18]

Following Remark 2, before computing the proposed IV estimators, the data are demeaned using the within transformation in order to eliminate individual-specific effects. $\hat{m}_x$ and $\hat{m}_y$ are obtained in each replication, based on the eigenvalue ratio (ER) statistic proposed by Ahn and Horenstein (2013, p.1207). In our experiment we set $m_x = 2$ and $m_y = 3$, as will be shown shortly. For the estimation we set the maximum number of factors equal to three for $\hat{m}_x$ and four for $\hat{m}_y$. For the CCEMG estimator, the untransformed data, $(\mathbf{y}_i, \mathbf{W}_i)$, are used but a $T \times 1$ vector of ones is included along with the cross-sectional averages, as described above. Finally, for the computation of BC-QMLE, following the practice of Moon and Weidner (2015) we use the within-transformed data, like in our IV estimators. To avoid introducing further uncertainty by estimating the number of factors in u_{it}, the BC-QMLE is computed given the true number of factors, m_y.

4.1 Design

We consider the following dynamic panel data model with two covariates and three factors:

$$y_{it} = \alpha_i + \rho_i y_{it-1} + \sum_{\ell=1}^{k} \beta_{\ell i} x_{\ell it} + u_{it};\ u_{it} = \sum_{s=1}^{m_y} \gamma^0_{si} f^0_{s,t} + \varepsilon_{it}, \tag{41}$$

$i = 1, ..., N$, $t = -49, ..., T$, where

$$f^0_{s,t} = \rho_{fs} f^0_{s,t-1} + (1 - \rho^2_{fs})^{1/2} \zeta_{s,t}, \tag{42}$$

with $\zeta_{s,t} \sim i.i.d.N(0,1)$ for $s = 1, ..., m_y$. We set $k = 2$ and $m_y = 3$. We set $\rho_{fs} = 0.5$ for all s.

The idiosyncratic error, ε_{it}, is non-normal and heteroskedastic across both i and t, such that $\varepsilon_{it} = \varsigma_\varepsilon \sigma_{it} (\epsilon_{it} - 1)/\sqrt{2}$, $\epsilon_{it} \sim i.i.d.\chi^2_1$, with $\sigma^2_{it} = \eta_i \varphi_t$, $\eta_i \sim i.i.d.\chi^2_2/2$, and $\varphi_t = t/T$ for $t = 0, 1, ..., T$ and unity otherwise.

It is straightforward to see that the average variance of ε_{it} depends only on ς^2_ε. Let π_u denote the proportion of the average variance of u_{it} due to ε_{it}. That is, we define $\pi_u := \varsigma^2_\varepsilon / (m_y + \varsigma^2_\varepsilon)$. Thus, for example, $\pi_u = 3/4$ means that the variance of the idiosyncratic error accounts for 75% of the total variance in u. In this case most of the variation in the total error is due to the idiosyncratic component and the factor structure has relatively minor significance. Solving in terms of ς^2_ε yields

$$\varsigma^2_\varepsilon = \frac{\pi_u}{(1 - \pi_u)} m_y. \tag{43}$$

We set ς^2_ε such that $\pi_u \in \{1/4, 3/4\}$.

The process for the covariates is given by

$$x_{\ell it} = \mu_{\ell i} + \sum_{s=1}^{m_x} \gamma^0_{\ell si} f^0_{s,t} + v_{\ell it};\ i = 1, 2, ..., N; t = -49, -48, ..., T, \tag{44}$$

for $\ell = 1, 2$.

We set $m_x = 2$. This implies that the first two factors in u_{it}, f^0_{1t}, f^0_{2t}, are also contained in $x_{\ell it}$, for $\ell = 1, 2$ whilst f^0_{3t} is included in u_{it} only. Observe that, using notation of earlier sections, $\mathbf{f}^0_{y,t} = (f^0_{1t}, f^0_{2t}, f^0_{3t})'$ and $\mathbf{f}^0_{x,t} = (f^0_{1t}, f^0_{2t})'$.

[18] We are grateful to Alex Chudik for sharing his code to compute the (BC-)CCEMG estimator with us.

The idiosyncratic errors of the process for the covariates are serially correlated, such that

$$v_{\ell it} = \rho_{v,\ell} v_{\ell it-1} + (1-\rho_{v,\ell}^2)^{1/2} \varpi_{\ell it},\ \varpi_{\ell it} \sim i.i.d.N(0, \varsigma_v^2), \tag{45}$$

for $\ell = 1, 2$. We set $\rho_{v,\ell} = 0.5$ for all ℓ.

Initially, all individual-specific effects and factor loadings are generated as correlated and *mean-zero* random variables, and they are distinguished using the superscript "*". In particular, the *mean-zero* individual-specific effects are drawn as

$$\alpha_i^* \sim i.i.d.N(0, (1-\rho_i)^2),\ \mu_{\ell i}^* = \rho_{\mu,\ell}\alpha_i^* + (1-\rho_{\mu,\ell}^2)^{1/2}\omega_{\ell i}, \tag{46}$$

where $\omega_{\ell i} \sim i.i.d.N(0, (1-\rho_i)^2)$, for $\ell = 1, 2$. We set $\rho_{\mu,\ell} = 0.5$ for $\ell = 1, 2$.
Moreover, the *mean-zero* factor loadings in u_{it} are generated as $\gamma_{si}^{0*} \sim i.i.d.N(0, 1)$ for $s = 1, ..., m_y = 3$, and the factor loadings in x_{1it} and x_{2it} are drawn as

$$\gamma_{1si}^{0*} = \rho_{\gamma,1s}\gamma_{3i}^{0*} + (1-\rho_{\gamma,1s}^2)^{1/2}\xi_{1si};\ \xi_{1si} \sim i.i.d.N(0, 1); \tag{47}$$

$$\gamma_{2si}^{0*} = \rho_{\gamma,2s}\gamma_{si}^{0*} + (1-\rho_{\gamma,2s}^2)^{1/2}\xi_{2si};\ \xi_{2si} \sim i.i.d.N(0, 1); \tag{48}$$

respectively, for $s = 1, ..., m_x = 2$. The process (47) allows the factor loadings to $f_{1,t}^0$ and $f_{2,t}^0$ in x_{1it} to be correlated with the factor loadings corresponding to the factor specific in u_{it}, $f_{3,t}^0$. On the other hand, (48) ensures that the factor loadings to $f_{1,t}^0$ and $f_{2,t}^0$ in x_{2it} are allowed to be correlated with the factor loadings corresponding to the same factors in u_{it}, $f_{1,t}^0$ and $f_{2,t}^0$. We consider $\rho_{\gamma,11} = \rho_{\gamma,12} \in \{0,\ 0.5\}$, whilst $\rho_{\gamma,21} = \rho_{\gamma,22} = 0.5$.

Finally, the factor loadings entering the model are generated such that

$$\mathbf{\Gamma}_i^0 = \mathbf{\Gamma}^0 + \mathbf{\Gamma}_i^{0*} \tag{49}$$

where

$$\mathbf{\Gamma}_i^0 = \begin{pmatrix} \gamma_{1i}^0 & \gamma_{11i}^0 & \gamma_{21i}^0 \\ \gamma_{2i}^0 & \gamma_{12i}^0 & \gamma_{22i}^0 \\ \gamma_{3i}^0 & 0 & 0 \end{pmatrix} \text{ and } \mathbf{\Gamma}_i^{0*} = \begin{pmatrix} \gamma_{1i}^{0*} & \gamma_{11i}^{0*} & \gamma_{21i}^{0*} \\ \gamma_{2i}^{0*} & \gamma_{12i}^{0*} & \gamma_{22i}^{0*} \\ \gamma_{3i}^{0*} & 0 & 0 \end{pmatrix}.$$

Observe that, using notation of earlier sections, $\boldsymbol{\gamma}_{yi}^0 = (\gamma_{1i}^0, \gamma_{2i}^0, \gamma_{3i}^0)'$ and $\mathbf{\Gamma}_{x,i}^0 = (\boldsymbol{\gamma}_{1i}^0, \boldsymbol{\gamma}_{2i}^0)'$ with $\boldsymbol{\gamma}_{\ell i}^0 = (\gamma_{\ell 1i}^0, \gamma_{\ell 2i}^0)'$ for $\ell = 1, 2$. Also it is easily seen that the average of the factor loadings are given by $E\left(\mathbf{\Gamma}_i^0\right) = \mathbf{\Gamma}^0$. To ensure that the rank condition for CCEMG is satisfied, we set[19]

$$\mathbf{\Gamma}^0 = \begin{pmatrix} 1/4 & 1/4 & -1 \\ 1/2 & -1 & 1/4 \\ 1/2 & 0 & 0 \end{pmatrix}. \tag{50}$$

We note that our estimators and BC-QMLE do not require this condition, and we also consider the experiment with $\mathbf{\Gamma}^0 = \mathbf{0}$.[20]

In a similar manner, the individual effects entering the data generating process are such that

$$\alpha_i = \alpha + \alpha_i^*, \mu_{\ell i} = \mu_\ell + \mu_{\ell i}^*, \tag{51}$$

for $\ell = 1, 2$, setting $\alpha = 1/2$, $\mu_1 = 1$, $\mu_2 = -1/2$.

The slope coefficients are generated as

$$\rho_i = \rho + \eta_{\rho i},\ \beta_{1i} = \beta_1 + \eta_{\beta_1 i} \text{ and } \beta_{2i} = \beta_2 + \eta_{\beta_2 i}. \tag{52}$$

[19]See Assumption 6 in Chudik and Pesaran (2015a).
[20]The results are reported in Tables C7 and C8 in Appendix C.

We consider $\rho \in \{0.5, 0.8\}$. Following Bai (2009a), we set $\beta_1 = 3$ and $\beta_2 = 1$ as a benchmark case. In order to investigate the properties of the estimator when one of the slope coefficients equals zero, we also consider $\beta_1 = 3$ and $\beta_2 = 0$.

For the homogeneous slopes design, we set $\rho_i = \rho$, $\beta_{1i} = \beta_1$ and $\beta_{2i} = \beta_2$. For the heterogeneous slopes design, we specify $\eta_{\rho i} \sim i.i.d.U\left[-c, +c\right]$, and

$$\eta_{\beta_\ell i} = \left[(2c)^2/12\right]^{1/2} \rho_\beta \xi_{\beta_\ell i} + \left(1 - \rho_\beta^2\right)^{1/2} \eta_{\rho i},$$

where $\xi_{\beta_\ell i}$ is the standardised squared idiosyncratic errors in $x_{\ell it}$, computed as

$$\xi_{\beta_\ell i} = \frac{\overline{v_{\ell i}^2} - \overline{v_\ell^2}}{\left[N^{-1}\sum_{i=1}^N \left(\overline{v_{\ell i}^2} - \overline{v_\ell^2}\right)^2\right]^{1/2}},$$

with $\overline{v_{\ell i}^2} = T^{-1}\sum_{t=1}^T v_{\ell it}^2$, $\overline{v_\ell^2} = N^{-1}\sum_{i=1}^N \overline{v_{\ell i}^2}$, for $\ell = 1, 2$. We set $c = 1/5$, $\rho_\beta = 0.4$ for $\ell = 1, 2$.

Denoting $\rho_v = \rho_{v,\ell}$, $\ell = 1, 2$, we define the signal-to-noise ratio (SNR) for the homogeneous model, conditional on the factor structure and the individual-specific effects, as follows:

$$SNR := \frac{\text{var}\left[\left(y_{it} - \varepsilon_{it}\right)|\mathcal{L}\right]}{\overline{\text{var}}\left(\varepsilon_{it}\right)} = \frac{\left(\frac{\beta_1^2 + \beta_2^2}{1-\rho_v^2}\right)\varsigma_v^2 + \frac{\varsigma_\varepsilon^2}{1-\rho_v^2} - \varsigma_\varepsilon^2}{\varsigma_\varepsilon^2}, \tag{53}$$

where $\mathcal{L}$ is the information set that contains the factor structure and the individual-specific effects,[21] and $\overline{\text{var}}\left(\varepsilon_{it}\right)$ is the overall average of $E\left(\varepsilon_{it}^2\right)$ over i and t. Solving for ς_v^2 yields

$$\varsigma_v^2 = \varsigma_\varepsilon^2\left[SNR - \frac{\rho_v^2}{1-\rho_v^2}\right]\left(\frac{\beta_1^2 + \beta_2^2}{1-\rho_v^2}\right)^{-1}. \tag{54}$$

We set $SNR = 4$, which lies with the range $\{3, 9\}$ considered by the simulation study of Bun and Kiviet (2006). We consider all the combinations of (T, N), for $T \in \{25, 50, 100, 200\}$ and $N \in \{25, 50, 100, 200\}$.

In order to investigate the power of the overidentifying restrictions test, which is defined in (24), we consider violations of the null due to slope heterogeneity and endogeneity as a result of the contemporaneous correlation between $\mathbf{x}_{it}$ and ε_{it}. For the slope heterogeneity, we use the DGP specified in (52). For the case of endogeneity, we replace the DGP given by (45) with $v_{\ell it} = \rho_{v,\ell} v_{\ell it-1} + (1 - \rho_\ell^2)^{1/2}\varpi_{\ell it} + \varepsilon_{it}$, where $\varpi_{\ell it}$ is defined previously, $\ell = 1, 2$.

All results are obtained based on 2,000 replications, and all tests are conducted at the 5% significance level. For the size of the "t-test", $H_0 : \rho = \rho^0$ (or $H_0 : \beta_\ell = \beta_\ell^0$ for $\ell = 1, 2$), where $\rho^0, \beta_1^0, \beta_2^0$ are the true parameter values. For the power of the test, $H_0 : \rho = \rho^0 + 0.1$ (or $H_0 : \beta_\ell = \beta_\ell^0 + 0.1$ for $\ell = 1, 2$) against two sided alternatives are considered. The power of the "t-test" reported below is the size-corrected power, for which the 5% critical values used are obtained as the 2.5% and 97.5% quantiles of the empirical distribution of the t-ratio under the null hypothesis.[22]

4.2 Results

Tables 1-4 report the bias and RMSE of IV2^b, BC-QMLE of Moon and Weidner (2017), IVMG^b, CCEMG of Chudik and Pesaran (2015a), and the size (nominal level is 5%) and power (size-adjusted) of the associated t-tests for the panel ARDL(1,0) model with $\rho = 0.5$, $\beta_1 = 3$, $\beta_2 = 1$

[21] The reason we condition on these variables is that they influence both the composite error in the equation for the dependent variable and the covariates.

[22] The size adjusted power is employed in this experiment as the finite T bias of the CCEMG and BC-QMLE estimators and the size distortion of the associated statistical tests often make the power comparison too confusing.

with $\pi_u = 3/4$.[23] To compare the sensitivity of the estimators to the correlation structure of the factor loadings in $\mathbf{x}_{it}$ and u_{it}, in Tables 1 and 2 we consider independent factor loadings in $\mathbf{x}_{it}$ and u_{it}, and in Tables 3 and 4 we consider correlated loadings.

We have investigated two different instruments sets for our estimators as explained by (36). IV2^a (IVMG^a) uses $2k$ instruments and IV2^b (IVMG^b) $3k$. As one might expect, the former has smaller bias but the latter smaller dispersion. In terms of RMSE, the latter always performed better. Therefore, we only report results for IV2^b and IVMG^b.[24] Moreover, we do not report results for BC-CCEMG, since in our experiments it did not reduce bias of CCEMG, neither mitigated the size-distortion of the associated t-tests.[25]

Table 1 reports results for the model under slope homogeneity. Panel A corresponds to ρ and Panel B to β_1. The results for β_2 are not reported because they are qualitatively similar to those for β_1.[26] As we can see, IV2^b appears to have virtually no bias. In particular, the largest reported value of absolute bias is 0.1 for $T = 50, N = 25$. In comparison, absolute bias of BC-QMLE appears to be much larger, perhaps reflecting that bias correction is not able to completely remove the bias under these circumstances. However, absolute bias steadily declines with larger values of N and T. In terms of RMSE, BC-QMLE outperforms IV2^b and other estimators, which reflects the higher efficiency of the maximum likelihood approach over IV and least-squares. However, for larger values of N or T (especially N), the RMSE values of IV2^b are very close to those of BC-QMLE. The bias of IVMG^b is similar to that of BC-QMLE, whilst the bias of CCEMG tends to be much larger, especially when $N = 25$ or $T = 25, 50$. IVMG^b mostly outperforms CCEMG in terms of RMSE.

In regards to inference, the size of the t-test associated with IV2^b is close to the nominal value in most cases, with moderate size distortions observed for $N = 25$. The size of IVMG^b appears to be very accurate unless T is much smaller than N. By contrast, both BC-QMLE and CCEMG exhibit substantial size distortions, which may be partially attributed to the relatively large bias of these estimators. In view of such size-distortions, size-adjusted power is reported. As expected, under slope parameter homogeneity, the power of the estimators, IV2^b and BC-QMLE, is higher than MG-type estimators, at least when N and T are both relatively small.

Next, we turn our attention to Panel B of Table 1, which reports results for β_1. In comparison to Panel A, the bias of IV2^b is slightly larger for small N and T, although it remains smaller relative to other estimators. For instance, absolute bias of BC-QMLE is large when $N = 25$ but steadily declines as the sample size increases. As a result, IV2^b mostly outperforms BC-QMLE in terms RMSE. Also, because of the large bias of BC-QMLE, it suffers from large size-distortions. On the other hand, the properties of the absolute bias of IVMG^b and CCEMG are similar to that for ρ, in the sense that the former has smaller bias than the latter. In contrast, the size of the IV2^b is close to its nominal level with moderate distortion for $N = 25$. The size of the t-test of IVMG^b is very close to 5% for all combinations of N and T, whilst CCEMG exhibits moderate size distortions even for large values of N or T.

Table 2 reports results for the model with heterogeneous slopes. Notice that IV2^b and BC-QMLE are not asymptotically justified in this case. This is confirmed in finite samples. In particular, it is evident that IV2^b exhibits systematic bias, fluctuating around 0.01 across all combinations of N and T. The bias of BC-QMLE is much larger, reaching values close to 0.03 for ρ (Panel A), for large values of N and T. This outcome is accompanied by large size distortions for both estimators. By contrast, for IVMG^b and CCEMG bias appears to

[23]The results for the specifications where $\{\rho, \beta_1, \beta_2\} = \{0.8, 3, 1\}$,$\{0.5, 3, 0\}$ and $\pi_u = 1/4$ are qualitatively very similar. See Tables C1-C6 in Appendix C.

[24]The results for IV2^a and IVMG^a are reported in Tables C9 and C10 in Appendix C.

[25]The results for BC-CCEMG are provided in Tables C11 and C12 in Appendix C.

[26]These are available upon request from the authors.

behave in a similar manner to the homogenous case in Table 1. Similarly, IVMG^b continues to perform well in terms of size, whereas for CCEMG size properties improve substantially compared to the homogeneous case, although they still deviate from the nominal value to a significant extent, at least for small values of N or T. Similar conclusions apply to Panel B, with the main difference being that the size of the CCEMG estimator is closer to its nominal level for all combinations of N and T, and power of the t-test appears to be smaller across all estimators.

Now let us turn our attention to the case in which the factor loadings in x_{1it} are correlated with those in u_{it}. The results for homogeneous slopes are reported in Table 3 and those for heterogeneous slopes are shown in Table 4. The performance of IV2^b and IVMG^b is very similar to that shown in Tables 1 and 2, which suggests the robustness of our approach against such correlations in factor loadings. In contrast, for β_1, the performance of BC-QMLE and CCEMG appears to deteriorate when the factor loadings are correlated. For example, for $T = 100$ and $N = 25, 50, 100, 200$ the bias($\times 100$) for BC-QMLE equals -4.0, 2.1, -0.5, 0.4, whereas the corresponding values in the uncorrelated loadings design (Table 1 Panel B) are -1.6, -0.5, 0.2, 0.4. Consequently, in terms of RMSE and the size of the test, IV2^b outperforms BC-QMLE.

For the models with heterogeneous slopes, interestingly the bias of the CCEMG estimator for β_1 does not decrease as sample size rises. For example, in the correlated loadings design with heterogeneous slopes (Table 4 Panel B) the bias($\times 100$) of CCEMG for β_1 for $T = N = 25, 50, 100, 200$ are 0.7, -0.8, -1.3, -1.5 whilst in the uncorrelated loadings design (Table 2 Panel B) those are 2.2, 1.3, 0.1, -0.3. Consequently, IVMG^b mostly outperforms CCEMG in terms of bias, RMSE and size by a substantial margin.

Finally, we look at the finite sample behaviour of the overidentifying restrictions test based on IV2^b estimator, which is summarised in Table 5. As emphasised in Remark 7, we would like the test to reject the null when the exogeneity assumption on $\mathbf{x}_{it}$ is violated and/or when the slope coefficients are cross-sectionally heterogeneous. Table 5 contains two column blocks: the left one entitled IV2^a shows results using $2k$ instruments. In the right block entitled IV2^b shows results using additional instruments, which raises the total number of instruments to $3k$. The latter case provides for more degrees of freedom of the overidentifying restrictions test. As we can see, the size of the test is sufficiently close to its nominal level for both sets of instruments. On the other hand, there appear to be substantial differences in terms of the power of the test against slope heterogeneity. In particular, when $2k$ instruments are employed, there is no power, resulting in rejection frequencies equal to 4.7% and 5.0% for $N = T = 100$ and $N = T = 200$. By contrast, when we add two more instruments, power increases dramatically, such that for $N = T = 100$ and $N = T = 200$ power rises to 23.9% and 77.2%, respectively. Finally the test has substantial power for both sets of instruments when the exogeneity of $\mathbf{x}_{it}$ is violated - specifically, ε_{it} is correlated with $\mathbf{v}_{it}$. For example, for $N = T = 100$ and $N = T = 200$ the power of the test with $2k$ instruments is 45.2% and 95.9%, whilst that with $3k$ instruments is 36.7% and 91.4%, respectively.

In conclusion, we recommend the use of the (optimal) second-step IV estimator, $\widehat{\hat{\boldsymbol{\theta}}}_{IV2}$ defined by (23) for slope homogeneous models, and the mean group IV estimator, $\hat{\boldsymbol{\theta}}_{IVMG}$ defined by (31) for slope heterogeneous models with a moderate number of degrees of freedom. This is because $\widehat{\hat{\boldsymbol{\theta}}}_{IV2}$ is more efficient than $\hat{\boldsymbol{\theta}}_{IVMG}$ for homogeneous slope specification but it becomes unreliable for the models with heterogeneous slopes. We also note that $\hat{\boldsymbol{\theta}}_{IVMG}$ and the associated t-test seem reliable for the models with heterogeneous as well as homogeneous slope coefficients in our experiment. Both estimators appear to be reasonably precise, and, notably, robust in cases where factor loadings are mutually correlated. The size of the associated tests typically is far more accurate compared to BC-QMLE and CCEMG, and they have sound power. The choice between the two estimators depends on the assertion of heterogeneity in the slope coefficients. The overidentifying restrictions test associated with the optimal second-step IV estimator has good power to reject the null under slope heterogeneity with sufficient degrees of overidentification, which could be used as a guide.

5 Concluding Remarks

This paper develops two instrumental variable estimators for consistent estimation of homogeneous and heterogeneous dynamic panel data models with a multifactor error structure, when both N and T are large. The proposed approach initially involves extracting the common factors from the exogenous regressors of the model. Subsequently, the model is defactored, and defactored regressors are used as instruments to build a consistent first-step IV estimator, denoted as $\hat{\boldsymbol{\theta}}_{IV}$. For the case of homogeneous slope coefficients, subsequently the factors entering the disturbance of the y process are extracted from the residuals obtained from the one-step estimates; next, the entire model is defactored accordingly, and the defactored regressors used in the first stage as instruments, are employed in order to build an optimal two-step IV estimator, denoted as $\hat{\hat{\boldsymbol{\theta}}}_{IV}$. For the case of heterogeneous slope coefficients, the proposed approach involves mean group estimation of individual-specific IV regressions, based on the aforementioned first-step instruments.

The two-step IV estimator is $\sqrt{NT}$-consistent under slope parameter homogeneity. Notably, it requires no bias correction, unlike Moon and Weidner (2017). Similarly, the proposed mean group IV estimator does not require any small T bias correction, unlike Chudik and Pesaran (2015a), which employs a subsampling method for such adjustment.

The finite sample evidence reported in the paper suggests that the proposed estimators perform reasonably well under all circumstances examined, and therefore it presents a good alternative method of estimation to existing approaches. In particular, in comparison to alternative methods examined, both IV estimators appear to have little or negligible bias in most circumstances, and correct size of the t-test. Furthermore, the experimental results of the overidentifying restrictions test show that it has high power when a key assumption of the model is violated, namely the exogeneity of x.

Naturally, it is recommended to employ the optimal two-step IV estimator for slope homogeneous models, and the mean group IV estimator for slope heterogeneous models. This is because $\widehat{\hat{\boldsymbol{\theta}}}_{IV2}$ is more efficient than $\hat{\boldsymbol{\theta}}_{IVMG}$ for homogeneous slope specification but it becomes unreliable for the models with heterogeneous slopes. We also note that $\hat{\boldsymbol{\theta}}_{IVMG}$ and the associated t-test seem reliable for the models with heterogeneous as well as homogeneous slope coefficients in our experiment. The choice of the estimators depends on the assertion of heterogeneity in the slope coefficients. The experimental results show that the overidentifying restriction test associated with the optimal second-step IV estimator in general has good power to reject the null of slope homogeneity, unless the degrees of freedom of the test is very small. Thus, the development of a direct test for slope heterogeneity is of importance. We leave this avenue for future research.

In this paper we assumed that the covariates in the model is strongly exogenous with respect to its idiosyncratic errors. This assumption may not be too restrictive for many applications, thought, relaxing to weakly exogenous regressors, such that $\mathbf{x}_{it} = \mathbf{\Gamma}^{0\prime}_{xi}\mathbf{f}^0_{x,t} + \boldsymbol{\kappa}\varepsilon_{i,t-1} + \mathbf{v}_{it}$ with $\boldsymbol{\kappa} = (\kappa_1, ..., \kappa_k)'$ may be of interest and worth a further investigation.

Finally, we note that our approach is quite general and actually applicable to a large class of linear panel data models. For example, our method is applicable to the model considered by Pesaran (2006), Bai and Li (2014), and Westerlund and Urbain (2015) among others: $y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \boldsymbol{\gamma}^{0\prime}_{yi}\mathbf{f}^0_{y,t} + \varepsilon_{it}$ with $\mathbf{x}_{it} = \mathbf{\Gamma}^{0\prime}_{xi}\mathbf{f}^0_{x,t} + \mathbf{v}_{it}$. Comparing our approach to the above mentioned existing approaches may be an interesting research theme.

Table 1: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, IVMG^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$, $\pi_u = 3/4$, independent factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	0.0	0.0	0.0	0.0	-0.5	-0.6	-0.8	-1.0	-0.5	-0.7	-0.6	-0.7	-3.2	-3.4	-3.7	-3.9
50	-0.1	0.0	0.0	0.0	0.0	-0.3	-0.4	-0.5	-0.5	-0.4	-0.4	-0.4	-0.8	-1.0	-1.2	-1.5
100	0.0	0.0	0.0	0.0	0.1	-0.1	-0.2	-0.3	-0.2	-0.2	-0.2	-0.2	0.4	0.1	-0.1	-0.4
200	0.0	0.0	0.0	0.0	0.2	0.0	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	0.9	0.7	0.4	0.1
	RMSE (×100)															
25	3.2	2.2	1.7	1.1	1.5	1.2	1.2	1.2	3.3	2.4	1.8	1.4	4.1	4.0	4.0	4.1
50	2.1	1.4	1.0	0.7	1.0	0.8	0.6	0.6	2.3	1.6	1.2	0.9	1.8	1.5	1.5	1.6
100	1.4	1.0	0.7	0.4	0.7	0.5	0.4	0.3	1.4	1.1	0.7	0.5	1.1	0.9	0.6	0.6
200	1.0	0.7	0.4	0.3	0.5	0.4	0.3	0.2	1.0	0.7	0.5	0.4	1.2	0.9	0.7	0.4
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	9.5	7.4	6.9	4.8	18.3	22.7	37.4	59.1	5.5	6.1	7.5	10.7	28.2	51.1	81.3	97.7
50	10.2	6.0	6.0	5.7	13.5	15.1	22.8	45.0	6.5	6.4	6.9	9.2	13.8	22.7	49.8	81.8
100	8.4	6.4	6.3	5.2	13.5	14.8	17.4	27.0	5.6	5.8	6.4	7.3	12.4	14.4	18.4	39.0
200	9.8	6.6	5.7	5.6	16.8	11.6	12.2	17.2	6.2	4.9	4.7	7.3	32.8	36.1	34.5	24.7
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	87.5	98.2	99.9	100.0	99.8	100.0	100.0	100.0	80.7	94.8	99.5	100.0	48.2	54.0	57.4	63.2
50	98.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	96.1	99.9	100.0	100.0	99.7	100.0	100.0	100.0
100	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0
200	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
PANEL B: Results for β_1, homogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	-0.2	-0.2	-0.1	0.0	-2.3	-1.1	-0.5	0.0	1.3	1.0	1.5	1.4	2.9	2.9	3.3	3.2
50	0.2	0.1	0.1	0.0	-1.8	-0.5	-0.1	0.4	0.9	0.7	0.8	0.7	1.1	1.4	1.9	2.0
100	-0.1	0.0	0.1	0.0	-1.6	-0.5	0.2	0.4	0.3	0.4	0.4	0.4	-0.6	-0.2	0.3	0.7
200	-0.2	0.0	0.0	0.0	-1.5	-0.3	0.1	0.2	0.0	0.2	0.1	0.2	-1.8	-1.2	-0.8	-0.2
	RMSE (×100)															
25	12.1	8.6	6.1	4.4	14.2	9.9	7.1	5.3	16.9	11.7	8.5	6.1	17.1	11.9	9.0	6.7
50	8.2	5.6	4.0	2.9	12.4	8.0	5.4	3.3	10.2	6.9	5.1	3.6	9.7	6.7	5.1	3.9
100	5.7	3.9	2.8	1.9	11.1	6.6	3.6	2.0	6.4	4.3	3.2	2.3	6.4	4.2	3.2	2.2
200	4.1	2.8	1.9	1.4	9.8	4.8	2.3	1.3	4.4	3.1	2.2	1.6	4.7	3.3	2.3	1.5
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	9.1	7.0	5.9	5.8	37.8	30.8	27.3	23.2	6.4	5.8	6.3	6.1	7.2	6.4	8.1	8.8
50	8.7	6.1	5.7	5.8	44.6	32.3	22.7	14.6	5.9	5.2	6.2	6.9	6.8	6.0	6.8	10.1
100	8.6	6.7	6.3	6.2	48.2	32.1	16.0	10.2	6.7	5.3	5.6	5.9	7.0	6.5	7.2	7.3
200	8.8	6.1	6.7	6.2	51.6	25.5	10.5	7.7	5.8	5.0	5.7	5.7	7.8	9.2	9.0	8.7
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	17.6	25.4	43.2	66.8	6.9	12.4	19.6	40.4	11.1	18.2	30.7	49.4	12.5	21.9	32.4	54.0
50	27.0	47.4	72.2	92.6	8.1	15.6	38.5	88.2	22.3	38.6	59.5	85.0	21.9	41.4	70.1	89.8
100	47.8	73.6	94.0	100.0	8.4	24.8	83.4	99.8	36.3	65.8	90.0	99.5	34.7	62.9	89.8	99.7
200	71.3	95.0	99.8	100.0	9.4	49.0	99.0	100.0	62.7	90.7	99.2	100.0	45.3	76.4	97.5	100.0

Notes: The data generating process is $y_{it} = \alpha_i + \rho_i y_{it-1} + \sum_{\ell=1}^{2} \beta_{\ell i} x_{\ell it} + u_{it}$, $u_{it} = \sum_{s=1}^{3} \gamma_{si}^0 f_{st}^0 + \varepsilon_{it}$, $x_{\ell it} = \mu_{\ell i} + \sum_{s=1}^{2} \gamma_{\ell si}^0 f_{st}^0 + v_{\ell it}$ $\ell = 1, 2$; $i = 1, .., N$; $t = -50, .., T$ and the first 50 observations are discarded; $f_{st}^0 = \rho_{fs} f_{st-1}^0 + (1 - \rho_{fs}^2)^{1/2} \zeta_{st}$, $\zeta_{st} \sim i.i.d.N(0, 1)$, $\gamma_{si}^0 = \gamma_s + \gamma_{si}^*$, $\gamma_{si}^* \sim i.i.d.N(0, 1)$ for $s = 1, 2, 3$, $\varepsilon_{it} = \varsigma_\varepsilon \sigma_{it} (\epsilon_{it} - 1)/\sqrt{2}$, $\epsilon_{it} \sim i.i.d.\chi_1^2$ with $\sigma_{it}^2 = \eta_i \varphi_t$, $\eta_i \sim i.i.d.\chi_2^2/2$, and $\varphi_t = t/T$ for $t = 0, 1, ..., T$ and unity otherwise; $\gamma_{\ell si}^0 = \gamma_{\ell s} + \gamma_{\ell si}^{0*}$, $\gamma_{1si}^{0*} = \rho_{\gamma,1s} \gamma_{3i}^{0*} + (1 - \rho_{\gamma,1s}^2)^{1/2} \xi_{1si}$, $\gamma_{2si}^{0*} = \rho_{\gamma,2s} \gamma_{si}^{0*} + (1 - \rho_{\gamma,2s}^2)^{1/2} \xi_{2si}$, $\xi_{\ell si} \sim i.i.d.N(0, 1)$, $v_{\ell it} = \rho_{\upsilon,\ell} v_{\ell it-1} + (1 - \rho_{\upsilon,\ell}^2)^{1/2} \varpi_{\ell it}$, $\varpi_{\ell it} \sim i.i.d.N(0, \varsigma_\upsilon^2 \sigma_{\varpi_{\ell,i}}^2)$, $\sigma_{\varpi_{\ell,i}}^2 \sim i.i.d.U[0.5, 1.5]$ for $\ell = 1, 2$, $s = 1, 2$. We set $\rho_i = \rho$, $\beta_{1i} = \beta_1$ and $\beta_{2i} = \beta_2$, $\rho_{fs} = \rho_{\gamma,2s} = \rho_{\upsilon,\ell} = 0.5$ and $\rho_{\gamma,1s} = 0.0$ for all ℓ, s. IV2^b and IVMG^b are given by (23), (31) with (36), BC-QMLE and CCEMG by (37), (38). The rank condition for CCEMG is met.

Table 2: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, IVMG^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$, $\pi_u = 3/4$, independent factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	0.7	0.8	0.9	0.8	1.1	0.8	0.5	0.3	-0.7	-0.7	-0.6	-0.7	-3.1	-3.3	-3.6	-3.8
50	1.1	1.1	1.1	1.1	2.0	1.9	1.8	1.8	-0.3	-0.3	-0.4	-0.4	-0.6	-0.9	-1.2	-1.5
100	1.0	1.2	1.2	1.2	2.5	2.5	2.3	2.3	-0.2	-0.2	-0.3	-0.2	0.4	0.3	-0.1	-0.3
200	1.1	1.2	1.3	1.3	2.7	2.6	2.6	2.6	-0.1	-0.1	-0.1	-0.1	1.0	0.8	0.5	0.2
	RMSE ($\times$100)															
25	4.4	3.2	2.3	1.7	4.1	2.9	2.2	1.7	4.2	3.0	2.1	1.6	4.8	4.3	4.1	4.1
50	3.4	2.6	2.0	1.5	4.0	3.1	2.5	2.1	3.1	2.2	1.6	1.2	2.8	2.2	1.9	1.8
100	3.0	2.3	1.8	1.6	4.1	3.3	2.8	2.6	2.7	1.9	1.4	1.0	2.5	1.9	1.3	1.0
200	2.9	2.1	1.8	1.6	4.1	3.4	3.0	2.8	2.5	1.7	1.2	0.9	2.6	1.8	1.3	0.9
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	10.7	9.9	9.5	10.8	52.0	47.3	47.4	48.6	6.2	7.0	7.0	9.2	18.6	30.7	56.7	84.2
50	11.6	12.0	13.1	17.1	63.6	67.5	73.1	79.7	5.2	6.3	5.9	6.2	8.3	10.2	16.2	36.4
100	12.5	13.3	17.1	27.7	75.0	80.2	87.9	94.4	5.9	6.3	5.9	5.9	7.3	8.0	8.0	9.6
200	13.6	13.8	20.1	34.8	82.7	86.4	94.4	98.9	5.9	5.4	5.0	5.2	9.3	8.0	9.0	7.1
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	72.2	91.8	99.9	100.0	69.8	92.7	98.0	99.7	65.1	85.9	99.1	100.0	31.0	40.6	46.9	51.8
50	89.7	98.6	100.0	100.0	82.4	97.6	99.9	100.0	84.5	98.3	99.9	100.0	89.2	98.6	100.0	100.0
100	95.9	100.0	100.0	100.0	82.9	99.2	100.0	100.0	93.1	99.8	100.0	100.0	98.1	100.0	100.0	100.0
200	97.7	100.0	100.0	100.0	88.6	99.6	100.0	100.0	96.4	100.0	100.0	100.0	99.2	100.0	100.0	100.0
PANEL B: Results for β_1, homogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	-1.9	-1.1	-1.3	-1.3	-4.9	-4.1	-4.4	-3.7	1.1	1.7	1.5	1.6	2.2	3.3	2.9	3.2
50	-1.6	-1.3	-1.1	-1.3	-5.4	-4.0	-3.6	-3.7	0.5	0.7	0.9	0.4	0.8	1.3	1.8	1.9
100	-1.0	-1.2	-1.2	-1.2	-4.7	-4.0	-3.5	-3.4	0.5	0.2	0.3	0.3	-0.4	-0.4	0.1	0.6
200	-1.0	-1.1	-1.1	-1.2	-5.0	-4.0	-3.7	-3.6	0.2	0.1	0.1	0.2	-1.6	-1.4	-0.9	-0.3
	RMSE ($\times$100)															
25	13.8	9.7	7.2	5.0	16.0	12.1	9.5	7.3	16.8	11.9	8.6	6.1	16.8	12.2	9.0	6.7
50	9.6	6.9	4.8	3.6	14.5	9.8	7.3	5.7	10.4	7.2	5.2	3.6	9.8	7.2	5.3	3.8
100	7.0	5.0	3.5	2.6	12.6	8.7	6.0	4.8	7.0	4.9	3.3	2.4	6.7	4.7	3.2	2.3
200	5.1	3.6	2.8	2.1	11.7	7.7	5.4	4.4	4.9	3.4	2.5	1.7	5.3	3.8	2.6	1.7
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	11.0	8.0	8.1	7.3	41.6	37.3	37.6	37.0	5.6	5.4	6.1	6.8	6.7	6.5	7.6	7.8
50	9.8	7.4	6.4	8.0	47.5	39.9	35.6	41.2	5.9	5.6	6.0	4.9	6.1	6.3	7.5	9.1
100	9.5	8.8	7.3	8.6	52.8	47.4	42.9	50.3	5.4	5.5	5.0	5.4	5.7	6.4	5.3	6.3
200	9.5	7.3	9.5	11.6	60.4	54.7	56.0	67.6	6.1	5.8	5.3	4.9	8.5	9.4	8.6	6.8
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	10.5	15.2	24.2	40.5	4.1	5.5	4.5	7.4	12.1	19.4	29.7	47.6	12.8	19.7	30.9	53.4
50	14.7	28.5	46.5	70.1	3.3	5.6	7.5	13.0	19.9	33.6	56.4	83.5	19.8	38.1	62.6	89.9
100	28.1	46.7	73.7	94.5	4.4	6.4	12.7	22.4	36.1	57.9	86.6	99.1	33.4	51.3	87.7	99.5
200	45.3	68.2	92.4	99.8	4.5	6.4	13.7	27.5	52.2	81.2	97.9	100.0	34.6	64.0	94.4	99.9

Notes: The DGP is the same as that for Table 1 except that the slope coefficients are heterogeneous. Specifically, $\rho_i = \rho + \eta_{\rho i}$; $\beta_{\ell i} = \beta_\ell + \eta_{\beta_\ell i}$, $\eta_{\rho i} \sim i.i.d.U\left[-1/5, +1/5\right]$, and $\eta_{\beta_\ell i} = \left[(2/5)^2/12\right]^{1/2} \rho_\beta \xi_{\beta_\ell i} + (1-\rho_\beta^2)^{1/2} \zeta_{\beta_\ell i}$, where $\xi_{\beta_\ell i}$ is the standardised squared idiosyncratic errors in $x_{\ell it}$, computed as $\xi_{\beta_\ell i} = (\overline{v_{\ell i}^2} - \overline{v_\ell^2})/[N^{-1}\sum_{i=1}^N (\overline{v_{\ell i}^2} - \overline{v_\ell^2})^2]^{1/2}$ with $\overline{v_{\ell i}^2} = T^{-1}\sum_{t=1}^T v_{\ell it}^2$, $\overline{v_\ell^2} = N^{-1}\sum_{i=1}^N \overline{v_{\ell i}^2}$, for $\ell = 1, 2$, whereas $\zeta_{\beta_\ell i} \sim i.i.d.U\left(-\sqrt{3}, \sqrt{3}\right)$ for $\ell = 1, 2$.

Table 3: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, IVMG^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$, $\pi_u = 3/4$, correlated factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	0.1	0.0	0.0	0.0	-0.4	-0.7	-0.9	-1.0	-0.7	-0.7	-0.6	-0.7	-3.2	-3.5	-3.6	-3.8
50	0.0	0.0	0.0	0.0	0.0	-0.3	-0.5	-0.6	-0.5	-0.4	-0.4	-0.3	-0.9	-1.0	-1.2	-1.4
100	0.0	0.0	0.0	0.0	0.1	-0.1	-0.2	-0.3	-0.3	-0.2	-0.2	-0.2	0.3	0.2	-0.1	-0.4
200	0.0	0.0	0.0	0.0	0.2	0.0	-0.1	-0.1	-0.1	-0.1	-0.1	-0.1	0.9	0.7	0.4	0.1
	RMSE (×100)															
25	3.1	2.2	1.6	1.1	1.6	1.4	1.3	1.3	3.4	2.6	1.9	1.5	4.3	4.1	4.0	4.0
50	2.1	1.4	1.0	0.7	1.1	0.8	0.7	0.7	2.3	1.5	1.1	0.9	1.8	1.6	1.5	1.6
100	1.4	1.0	0.6	0.4	0.7	0.5	0.4	0.4	1.5	1.0	0.7	0.6	1.1	0.9	0.7	0.6
200	1.0	0.6	0.4	0.3	0.6	0.4	0.3	0.2	1.0	0.7	0.5	0.4	1.2	0.9	0.6	0.4
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	7.5	6.4	6.1	5.4	18.5	26.0	39.1	59.2	5.8	7.0	7.9	11.5	30.2	52.8	77.8	94.9
50	8.7	7.0	6.7	4.9	16.4	17.8	25.5	46.9	5.9	5.4	6.5	8.7	14.3	25.5	49.1	76.6
100	8.9	6.5	5.2	4.8	13.4	14.2	18.7	30.2	5.9	6.4	5.2	6.6	13.5	17.1	23.5	40.7
200	8.6	5.4	6.1	5.3	16.0	10.7	12.3	17.1	6.2	5.4	5.5	6.6	32.9	35.7	33.4	28.1
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	89.0	98.0	100.0	100.0	99.8	100.0	100.0	100.0	78.9	93.7	99.0	99.9	45.3	48.1	55.7	60.7
50	99.1	100.0	100.0	100.0	100.0	100.0	100.0	100.0	96.5	99.6	100.0	100.0	99.5	100.0	100.0	100.0
100	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0
200	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
PANEL B: Results for β_1, homogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	-0.2	0.2	0.0	-0.1	-4.4	-3.2	-2.9	-1.6	1.3	1.7	1.4	1.5	0.6	1.2	1.3	1.9
50	0.0	0.0	0.0	-0.1	-4.8	-2.9	-1.1	-0.1	0.9	0.7	0.8	0.5	-0.9	-0.4	0.2	0.7
100	0.2	-0.1	0.0	0.0	-4.0	-2.1	-0.5	0.4	0.6	0.2	0.3	0.4	-2.6	-2.3	-1.4	-0.5
200	0.0	0.0	0.0	0.0	-3.6	-1.6	-0.1	0.3	0.2	0.2	0.2	0.2	-3.9	-3.3	-2.4	-1.4
	RMSE (×100)															
25	11.8	8.7	6.1	4.3	16.6	12.4	9.9	7.4	16.9	11.9	8.5	6.2	16.6	12.4	8.8	6.4
50	8.1	5.6	4.0	2.7	14.5	10.7	7.3	4.7	10.0	6.8	5.0	3.5	10.1	7.1	4.9	3.5
100	5.8	3.9	2.8	1.9	13.4	8.6	4.9	2.5	6.4	4.4	3.2	2.2	7.2	5.2	3.6	2.3
200	3.9	2.8	1.9	1.4	12.2	6.8	3.1	1.4	4.3	3.1	2.1	1.5	6.3	4.8	3.2	2.1
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	8.5	7.5	6.2	6.1	44.5	40.8	37.9	32.2	5.8	5.9	5.9	6.3	6.8	7.4	6.9	7.5
50	8.8	5.4	6.3	4.4	50.4	40.0	27.4	18.9	5.1	4.9	6.0	5.0	8.3	7.8	5.9	6.8
100	8.0	6.3	6.4	5.0	52.6	34.7	17.9	9.9	5.1	4.4	5.4	4.3	10.8	12.5	11.1	7.9
200	8.2	5.7	5.0	7.0	55.5	30.8	13.0	8.0	5.7	5.0	4.6	5.5	18.2	24.8	24.2	21.5
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	19.6	26.2	41.4	66.3	5.4	4.8	6.0	9.0	13.0	20.3	29.9	50.2	10.2	16.4	26.2	48.3
50	31.8	47.7	73.1	95.4	3.6	4.4	9.7	67.4	22.3	37.6	60.8	86.3	15.5	25.5	57.9	87.0
100	50.4	75.8	94.4	99.9	4.9	5.5	58.6	99.2	43.2	66.0	89.7	99.6	17.1	29.8	67.6	98.8
200	73.5	94.8	99.7	100.0	3.8	8.6	98.0	100.0	65.9	91.2	99.2	100.0	14.9	27.9	73.8	99.7

Notes: The DGP is the same as that for Table 1 except that the factor loadings in x_{1it} & u_{it} are correlated: $\rho_{\gamma,1s} = 0.5$ in $\gamma^{0*}_{1si} = \rho_{\gamma,1s}\gamma^{0*}_{3i} + (1-\rho^2_{\gamma,1s})^{1/2}\xi_{1si}$.

Table 4: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, IVMG^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$, $\pi_u = 3/4$, correlated factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times 100$)															
25	0.8	0.9	0.9	0.9	0.9	0.8	0.5	0.2	-0.7	-0.7	-0.7	-0.7	-3.1	-3.3	-3.5	-3.8
50	0.9	1.1	1.2	1.2	2.2	2.1	1.9	1.9	-0.4	-0.3	-0.3	-0.3	-0.7	-0.9	-1.1	-1.3
100	1.1	1.2	1.2	1.2	2.8	2.7	2.6	2.6	-0.2	-0.3	-0.2	-0.2	0.4	0.2	-0.1	-0.3
200	1.1	1.3	1.3	1.3	3.1	2.9	2.8	2.9	-0.1	-0.1	-0.1	-0.1	0.9	0.8	0.4	0.2
	RMSE ($\times 100$)															
25	4.4	3.1	2.3	1.7	4.3	3.2	2.5	2.0	4.1	3.0	2.2	1.7	4.8	4.3	4.1	4.1
50	3.5	2.6	2.0	1.6	4.2	3.3	2.6	2.3	3.2	2.2	1.6	1.2	2.9	2.3	1.9	1.8
100	3.0	2.3	1.8	1.6	4.4	3.5	3.0	2.8	2.7	2.0	1.4	1.0	2.6	1.9	1.4	1.1
200	2.8	2.2	1.8	1.6	4.5	3.7	3.2	3.1	2.5	1.8	1.3	0.9	2.6	1.9	1.4	1.0
	SIZE: $H_0: \rho = 0.5$ against $H_1: \rho \neq 0.5$, at the 5% level															
25	11.7	9.2	9.3	10.5	53.2	50.6	51.5	53.7	6.3	6.6	7.4	8.9	17.3	30.0	54.9	81.9
50	12.7	11.5	12.9	19.8	66.2	67.6	72.3	81.5	6.2	5.1	6.8	6.8	9.0	11.6	17.3	35.3
100	11.6	13.1	15.9	26.3	77.5	82.0	87.6	95.8	5.9	5.6	5.5	6.2	8.6	8.2	8.3	13.0
200	12.3	14.1	21.9	33.3	84.4	89.3	93.7	99.0	5.5	5.3	4.9	5.4	9.4	10.1	10.9	9.6
	POWER (size-adjusted) : $H_0: \rho = 0.6$ against $H_1: \rho \neq 0.6$, at the 5% level															
25	71.9	92.8	99.6	100.0	64.8	90.1	96.6	99.0	61.2	87.9	98.3	100.0	30.9	41.5	44.3	46.1
50	87.2	98.9	100.0	100.0	78.4	97.3	99.9	100.0	80.7	98.2	100.0	100.0	89.1	98.8	100.0	100.0
100	96.3	100.0	100.0	100.0	83.5	99.0	100.0	100.0	93.6	99.7	100.0	100.0	97.6	100.0	100.0	100.0
200	98.6	100.0	100.0	100.0	84.8	99.7	100.0	100.0	97.3	100.0	100.0	100.0	99.2	100.0	100.0	100.0
PANEL B: Results for β_1, homogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times 100$)															
25	-1.2	-1.3	-1.4	-1.5	-7.6	-7.4	-6.5	-6.2	1.7	1.4	1.5	1.4	0.7	1.2	1.5	2.0
50	-1.0	-1.3	-1.2	-1.2	-8.5	-7.0	-5.9	-4.9	0.9	0.4	0.7	0.7	-0.8	-0.8	0.1	0.9
100	-1.2	-1.0	-1.1	-1.2	-8.0	-6.7	-5.2	-4.5	0.2	0.5	0.4	0.3	-2.7	-2.0	-1.3	-0.6
200	-1.3	-1.0	-1.1	-1.2	-7.9	-6.0	-4.9	-4.6	0.1	0.3	0.2	0.1	-3.9	-3.2	-2.3	-1.5
	RMSE ($\times 100$)															
25	13.4	9.8	6.8	5.1	18.9	15.2	12.1	10.3	16.9	11.9	8.5	6.1	17.0	12.3	9.0	6.4
50	9.6	6.9	4.8	3.5	17.5	13.5	10.0	7.6	10.2	7.1	5.2	3.7	10.2	7.4	5.1	3.8
100	6.9	5.1	3.5	2.6	15.9	11.4	8.1	6.0	6.7	4.8	3.4	2.3	7.7	5.5	3.8	2.6
200	5.3	3.8	2.7	2.1	15.1	9.7	6.7	5.5	5.0	3.6	2.5	1.7	6.9	5.0	3.5	2.4
	SIZE: $H_0: \beta_1 = 3$ against $H_1: \beta_1 \neq 3$, at the 5% level															
25	9.6	7.5	6.5	6.6	48.1	47.6	46.3	51.7	6.3	5.8	6.2	6.7	7.6	6.8	7.0	8.0
50	8.8	7.5	6.3	7.7	54.7	50.7	50.2	49.8	6.2	4.8	5.7	5.7	8.2	7.2	6.7	8.4
100	9.0	9.0	7.3	8.2	59.4	54.9	51.3	59.1	5.5	5.5	5.3	4.8	10.7	11.3	10.6	8.9
200	10.8	8.6	8.6	12.0	67.7	58.4	62.0	76.4	6.2	6.5	6.0	4.6	16.9	20.5	20.0	18.9
	POWER (size-adjusted) : $H_0: \beta_1 = 3.1$ against $H_1: \beta_1 \neq 3.1$, at the 5% level															
25	14.2	16.6	25.0	39.6	3.1	1.9	0.9	1.5	11.6	16.5	29.8	47.4	10.7	16.6	26.4	46.5
50	20.5	28.3	46.9	72.3	1.5	1.6	2.6	3.0	19.2	35.0	55.0	82.9	16.1	25.6	51.8	81.8
100	28.5	45.7	74.3	95.4	2.0	1.0	2.7	5.8	33.0	60.2	86.1	99.1	14.8	29.4	62.9	96.1
200	40.8	66.5	92.6	99.8	2.2	1.8	2.6	7.3	50.3	80.5	98.0	100.0	11.0	24.4	65.5	97.3

Notes: The DGP is the same as that for Table 2 except that the factor loadings in x_{1it} & u_{it} are correlated: $\rho_{\gamma,1s} = 0.5$ in $\gamma^{0*}_{1si} = \rho_{\gamma,1s}\gamma^{0*}_{3i} + (1-\rho^2_{\gamma,1s})^{1/2}\xi_{1si}$.

Table 5: Size and power of the overidentifying restrictions test for the panel ARDL(1,0) model with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$, $\pi_u = 3/4$, correlated factor loadings in x_{1it} & u_{it}

	$IV2^a$				$IV2^b$			
T,N	25	50	100	200	25	50	100	200
	Slope Homogeneity (Size)							
25	7.2	6.9	5.7	6.4	6.8	6.0	5.7	5.5
50	7.7	6.6	6.5	4.8	7.3	6.0	5.5	4.4
100	6.7	7.3	5.9	4.8	7.4	6.2	5.9	5.1
200	7.7	6.6	6.5	5.6	7.0	5.9	6.1	4.7
	Slope Heterogeneity (Power)							
25	7.5	6.2	6.0	5.4	8.3	7.9	8.8	10.1
50	6.9	5.9	5.4	5.7	8.6	10.4	12.6	22.0
100	7.4	5.6	4.7	4.9	11.2	13.6	23.9	44.3
200	7.0	5.9	5.8	5.0	14.6	24.0	45.3	77.2
	Endogeneous Idiosyncratic Error of X (Power)							
25	10.1	10.4	14.6	18.8	10.4	11.0	14.8	18.9
50	12.3	16.1	23.5	37.8	11.9	15.9	20.0	31.4
100	17.2	27.6	45.2	70.3	14.4	20.7	36.7	60.2
200	28.2	46.4	73.2	95.9	22.5	37.6	62.8	91.4

Notes: The table reports the size and the power of overidentifying restrictions tests based on the IV2 estimator using different set of instruments. $IV2^a$ uses $(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_{i,-1})$ and $IV2^b$ $(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_{i,-1}, \hat{\mathbf{X}}_{i,-2})$, where $\hat{\mathbf{X}}_i = \mathbf{M}_{F_x}\mathbf{X}_i$ and $\hat{\mathbf{X}}_{i,-j} = \mathbf{M}_{F_{x,-j}}\mathbf{X}_{i,-j}$ for $j = 1, 2$. The test statistic is defined by (24). The tests for $IV2^a$ and $IV2^b$ are referenced to the 95% quantiles of χ^2_1 and χ^2_3 distributions, respectively. The DGP for Slope Homogeneity is of Table 3, for Slope Heterogeneity is of Table 4, and for Endogeneous Idiosyncratic Error of X, the DGP of Table 3 is changed such that $v_{\ell it} = \rho_{\upsilon,\ell} v_{\ell it-1} + (1-\rho^2_\ell)^{1/2}\varpi_{\ell it}$, $\varpi_{\ell it} = \tau_\ell \varepsilon_{it} + (1-\tau^2_\ell)^{1/2}\varrho_{\ell it}$ with $\varrho_{\ell it} \sim i.i.d.N(0,1)$, $\ell = 1, 2$ (see notes to Table 1). We set $\tau_1 = 0.5$ and $\tau_2 = 0$ so that the idiosyncratic error of x_{1it} is contemporaneously correlated with ε_{it}

Appendix A: Proofs of Main Results

Lemma 1 *Under Assumptions 1-5, as $(N,T) \xrightarrow{j} \infty$ such that $N/T \to c$ with $0 < c < \infty$,*

$$T^{-1}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{Z}}_i - T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_i = o_p(1), \tag{A.1}$$

$$T^{-1}\left(\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y} - \mathbf{Z}_i'\mathbf{M}_{F_y^0}\right)\mathbf{W}_i = o_p(1). \tag{A.2}$$

Lemma 2 *Under Assumptions 1-5, as $(N,T) \xrightarrow{j} \infty$ such that $N/T \to c$ with $0 < c < \infty$,*

$$\frac{1}{NT}\sum_{i=1}^N \hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{Z}}_i - \mathbf{Z}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_i = o_p(1), \tag{A.3}$$

$$\frac{1}{NT}\sum_{i=1}^N \left(\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y} - \mathbf{Z}_i'\mathbf{M}_{F_y^0}\right)\mathbf{W}_i = o_p(1). \tag{A.4}$$

Lemma 3 *Under Assumptions 1-5, as $(N,T) \xrightarrow{j} \infty$ such that $N/T \to c$ with $0 < c < \infty$,*

$$\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^N \mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^0\right)^{-1}\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x'\left(\mathbf{\Sigma}_{kNT} - \bar{\mathbf{\Sigma}}_{kNT}\right)\mathbf{M}_{\hat{F}_x}\mathbf{u}_i$$
$$= O_p\left(T^{-1/2}\right) + O_p\left(\delta_{NT}^{-1}\right) + \sqrt{T}O_p\left(\delta_{NT}^{-2}\right), \tag{A.5}$$

and

$$\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^N \mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^0\right)^{-1}\left(\frac{\hat{\mathbf{F}}_{x,-1}'\mathbf{F}_{x,-1}^0}{T}\right)^{-1}\hat{\mathbf{F}}_{x,-1}'\left(\mathbf{\Sigma}_{kNT,-1} - \bar{\mathbf{\Sigma}}_{kNT,-1}\right)\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{u}_i$$
$$= O_p\left(T^{-1/2}\right) + O_p\left(\delta_{NT}^{-1}\right) + \sqrt{T}O_p\left(\delta_{NT}^{-2}\right), \tag{A.6}$$

where $\boldsymbol{\Sigma}_{kNT} = \frac{1}{N}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'$, $\boldsymbol{\Sigma}_{kNT,-1} = \frac{1}{N}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{v}_{\ell j,-1}\mathbf{v}_{\ell j,-1}'$ *and* $\bar{\boldsymbol{\Sigma}}_{kNT} = \frac{1}{N}\sum_{\ell=1}^{k}\sum_{j=1}^{N}E\left(\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'\right)$, $\bar{\boldsymbol{\Sigma}}_{kNT,-1} = \frac{1}{N}\sum_{\ell=1}^{k}\sum_{j=1}^{N}E\left(\mathbf{v}_{\ell j,-1}\mathbf{v}_{\ell j,-1}'\right)$.

Lemma 4 *Under Assumptions 1-5, as* $(N,T)\xrightarrow{j}\infty$ *such that* $N/T\to c$ *with* $0<c<\infty$,

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\mathbf{F}_{x}^{0\prime}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i\\
&=-\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\Gamma}_{xj}^{0}\mathbf{V}_j'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i\\
&-\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x'\boldsymbol{\Sigma}_{kNT}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i+o_p(1),
\end{aligned}\tag{A.7}$$

and

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\mathbf{F}_{x,-1}^{0\prime}\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{u}_i\\
&=-\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\Gamma}_{xj}^{0}\mathbf{V}_{j,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{u}_i\\
&-\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\hat{\mathbf{F}}_{x,-1}'\mathbf{F}_{x,-1}^0}{T}\right)^{-1}\hat{\mathbf{F}}_{x,-1}'\boldsymbol{\Sigma}_{kNT,-1}\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{u}_i+o_p(1).
\end{aligned}\tag{A.8}$$

Lemma 5 *Under Assumptions 1-5, as* $(N,T)\xrightarrow{j}\infty$ *such that* $N/T\to c$ *with* $0<c<\infty$,

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\mathbf{F}_{x}^{0\prime}\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_y}\mathbf{u}_i=o_p(1),\tag{A.9}$$

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\mathbf{F}_{x,-1}^{0\prime}\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_y}\mathbf{u}_i=o_p(1).\tag{A.10}$$

Lemma 6 *Under Assumptions 1-5, as* $(N,T)\xrightarrow{j}\infty$ *such that* $N/T\to c$ *with* $0<c<\infty$,

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\Gamma}_{xj}^{0}\mathbf{V}_j'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i\\
&=\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\Gamma}_{xj}^{0}\mathbf{V}_j'\mathbf{M}_{F_x^0}\mathbf{u}_i\\
&+\sqrt{\frac{T}{N}}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\Gamma}_{xn}^{0}\frac{\mathbf{V}_n'\mathbf{V}_j}{T}\boldsymbol{\Gamma}_{xj}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T}\\
&+o_p(1),
\end{aligned}\tag{A.11}$$

and

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\Gamma}_{xj}^{0}\mathbf{V}_{j,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{u}_i\\
&=\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\Gamma}_{xj}^{0}\mathbf{V}_{j,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{u}_i\\
&+\sqrt{\frac{T}{N}}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\Gamma}_{xn}^{0}\frac{\mathbf{V}_{n,-1}'\mathbf{V}_{j,-1}}{T}\boldsymbol{\Gamma}_{xj}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_{x,-1}^0}{T}\right)^{-1}\frac{\mathbf{F}_{x,-1}^{0\prime}\mathbf{u}_i}{T}\\
&+o_p(1).
\end{aligned}\tag{A.12}$$

Lemma 7 *Under Assumptions 1-5, as* $(N,T) \xrightarrow{j} \infty$ *such that* $N/T \to c$ *with* $0 < c < \infty$*,*

$$\frac{1}{\sqrt{N}T^{3/2}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1} \left(\frac{\hat{\mathbf{F}}_x^{\prime}\mathbf{F}_x^0}{T}\right)^{-1} \hat{\mathbf{F}}_x^{\prime} \bar{\boldsymbol{\Sigma}}_{kNT} \mathbf{M}_{\hat{F}_x} \mathbf{u}_i$$
$$= \frac{1}{\sqrt{N}T^{3/2}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1} \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \mathbf{F}_x^{0\prime} \bar{\boldsymbol{\Sigma}}_{kNT} \mathbf{M}_{F_x^0} \mathbf{u}_i + o_p(1), \tag{A.13}$$

and

$$\frac{1}{\sqrt{N}T^{3/2}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1} \left(\frac{\hat{\mathbf{F}}_{x,-1}^{\prime}\mathbf{F}_{x,-1}^0}{T}\right)^{-1} \hat{\mathbf{F}}_{x,-1}^{\prime} \bar{\boldsymbol{\Sigma}}_{kNT,-1} \mathbf{M}_{\hat{F}_{x,-1}} \mathbf{u}_i$$
$$= \frac{1}{\sqrt{N}T^{3/2}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1} \left(\frac{\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_{x,-1}^0}{T}\right)^{-1} \mathbf{F}_{x,-1}^{0\prime} \bar{\boldsymbol{\Sigma}}_{kNT,-1} \mathbf{M}_{F_{x,-1}^0} \mathbf{u}_i + o_p(1). \tag{A.14}$$

Lemma 8 *Under Assumptions 1-5, as* $(N,T) \xrightarrow{j} \infty$ *such that* $N/T \to c$ *with* $0 < c < \infty$*,*

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i^{\prime} \mathbf{M}_{\hat{F}_x} \mathbf{u}_i$$
$$= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i^{\prime} \mathbf{M}_{F_x^0} \mathbf{u}_i$$
$$- \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\mathbf{V}_i^{\prime}\mathbf{V}_j}{T} \boldsymbol{\Gamma}_{xj}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1} \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T} + o_p(1), \tag{A.15}$$

and

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_{i,-1}^{\prime} \mathbf{M}_{\hat{F}_{x,-1}} \mathbf{u}_i$$
$$= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_{i,-1}^{\prime} \mathbf{M}_{F_{x,-1}^0} \mathbf{u}_i$$
$$- \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\mathbf{V}_{i,-1}^{\prime}\mathbf{V}_{j,-1}}{T} \boldsymbol{\Gamma}_{xj}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1} \left(\frac{\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_{x,-1}^0}{T}\right)^{-1} \frac{\mathbf{F}_{x,-1}^{0\prime}\mathbf{u}_i}{T} + o_p(1). \tag{A.16}$$

Lemma 9 *Under Assumptions 1-5, as* $(N,T) \xrightarrow{j} \infty$ *such that* $N/T \to c$ *with* $0 < c < \infty$*,*

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i^{\prime} \mathbf{M}_{\hat{F}_x} \mathbf{M}_{\hat{F}_y} \mathbf{u}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i^{\prime} \mathbf{M}_{F_x^0} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i + o_p(1), \tag{A.17}$$

and

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_{i,-1}^{\prime} \mathbf{M}_{\hat{F}_{x,-1}} \mathbf{M}_{\hat{F}_y} \mathbf{u}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_{i,-1}^{\prime} \mathbf{M}_{F_{x,-1}^0} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i + o_p(1), \tag{A.18}$$

Proof of Proposition 1. Consider

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \hat{\mathbf{Z}}_i^{\prime} \mathbf{u}_i, \tag{A.19}$$

where $\hat{\mathbf{Z}}_i = \left[\mathbf{M}_{\hat{F}_x}\mathbf{X}_i, \mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1}\right]$. We begin with the first component of $\hat{\mathbf{Z}}_i$, which is $\mathbf{M}_{\hat{F}_x}\mathbf{X}_i$. Firstly, note that

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{X}_i^{\prime} \mathbf{M}_{\hat{F}_x} \mathbf{u}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \mathbf{F}_x^{0\prime} \mathbf{M}_{\hat{F}_x} \mathbf{u}_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i^{\prime} \mathbf{M}_{\hat{F}_x} \mathbf{u}_i. \tag{A.20}$$

By using the results of Lemmas 3, 4, 6 and 7, the first term in (A.20) is given by

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\mathbf{F}_x^{0\prime}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&= -\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\mathbf{\Gamma}_{xj}^{0}\mathbf{V}_j'\mathbf{M}_{F_x^0}\mathbf{u}_i \\
&+\sqrt{\frac{T}{N}}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\mathbf{\Gamma}_{xn}^{0}\frac{\mathbf{V}_n'\mathbf{V}_j}{T}\mathbf{\Gamma}_{xj}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^{0}}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T} \\
&-\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\mathbf{F}_x^{0\prime}\bar{\mathbf{\Sigma}}_{kNT}\mathbf{M}_{F_x^0}\mathbf{u}_i + o_p(1). \qquad \text{(A.21)}
\end{aligned}
$$

By making use of Lemma 8, the second term in (A.20) is given by

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i \\
&-\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{V}_j}{T}\mathbf{\Gamma}_{xj}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^{0}}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T} + o_p(1). \qquad \text{(A.22)}
\end{aligned}
$$

So, by adding (A.21) and (A.22) together, rearranging the terms and using $\mathbf{M}_{F_x^0}\mathbf{F}_x^0 = \mathbf{0}$, we get

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i \\
&-\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\mathbf{\Gamma}_{xj}^{0}\mathbf{X}_j'\mathbf{M}_{F_x^0}\mathbf{u}_i \\
&-\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\tilde{\mathbf{V}}_i'\mathbf{V}_j}{T}\mathbf{\Gamma}_{xj}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^{0}}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T} \\
&-\sqrt{\frac{N}{T}}\frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^{0}}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\bar{\mathbf{\Sigma}}_{kNT}\mathbf{M}_{F_x^0}\mathbf{u}_i + o_p(1) \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\tilde{\mathbf{X}}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i + \sqrt{\frac{T}{N}}\mathbf{b}_{11NT} + \sqrt{\frac{N}{T}}\mathbf{b}_{21NT} + o_p(1), \qquad \text{(A.23)}
\end{aligned}
$$

where

$$
\tilde{\mathbf{X}}_i = \mathbf{X}_i - \frac{1}{N}\sum_{n=1}^{N}\mathbf{X}_n\mathbf{\Gamma}_{xn}^{0\prime}(\mathbf{\Upsilon}_{xkN}^{0})^{-1}\mathbf{\Gamma}_{xi}^{0},
$$

$$
\tilde{\mathbf{V}}_i = \mathbf{V}_i - \frac{1}{N}\sum_{n=1}^{N}\mathbf{V}_n\mathbf{\Gamma}_{xn}^{0\prime}(\mathbf{\Upsilon}_{xkN}^{0})^{-1}\mathbf{\Gamma}_{xi}^{0},
$$

$$
\mathbf{b}_{11NT} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\tilde{\mathbf{V}}_i'\mathbf{V}_j}{T}\mathbf{\Gamma}_{xj}^{0\prime}(\mathbf{\Upsilon}_{xkN}^{0})^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^{0}}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_x^0}\mathbf{u}_i}{T},
$$

$$
\mathbf{b}_{21NT} = -\frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}(\mathbf{\Upsilon}_{xkN}^{0})^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^{0}}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\bar{\mathbf{\Sigma}}_{kNT}\mathbf{M}_{F_x^0}\mathbf{u}_i.
$$

As for the second component of $\hat{\mathbf{Z}}_i$, which is $\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1}$, by following the same steps as before and using again Lemmas 3, 4, 6, 7, 8, and using $\mathbf{M}_{F^0_{x,-1}}\mathbf{F}^0_{x,-1} = \mathbf{0}$, we obtain

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}'_{i,-1}\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}'_{i,-1}\mathbf{M}_{F^0_{x,-1}}\mathbf{u}_i \\
&- \frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}^{0\prime}_{xi}\left(\boldsymbol{\Upsilon}^0_{xkN}\right)^{-1}\boldsymbol{\Gamma}^0_{xj}\mathbf{X}'_{j,-1}\mathbf{M}_{F^0_{x,-1}}\mathbf{u}_i \\
&+ \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\tilde{\mathbf{V}}'_{i,-1}\mathbf{V}_{j,-1}}{T}\boldsymbol{\Gamma}^{0\prime}_{xj}\left(\boldsymbol{\Upsilon}^0_{xkN}\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}_{x,-1}\mathbf{F}^0_{x,-1}}{T}\right)^{-1}\frac{\mathbf{F}^{0\prime}_{x,-1}\mathbf{u}_i}{T} \\
&- \frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\boldsymbol{\Gamma}^{0\prime}_{xi}\left(\boldsymbol{\Upsilon}^0_{xkN}\right)^{-1}\left(\frac{\mathbf{F}^{0\prime}_{x,-1}\mathbf{F}^0_{x,-1}}{T}\right)^{-1}\mathbf{F}^{0\prime}_{x,-1}\bar{\boldsymbol{\Sigma}}_{kNT,-1}\mathbf{M}_{F^0_{x,-1}}\mathbf{u}_i + o_p(1) \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\tilde{\mathbf{X}}'_{i,-1}\mathbf{M}_{F^0_{x,-1}}\mathbf{u}_i + \sqrt{\frac{T}{N}}\mathbf{b}_{12NT} + \sqrt{\frac{N}{T}}\mathbf{b}_{22NT} + o_p(1),
\end{aligned} \tag{A.24}
$$

where

$$
\begin{aligned}
\tilde{\mathbf{X}}_{i,-1} &= \mathbf{X}_{i,-1} - \frac{1}{N}\sum_{n=1}^{N}\mathbf{X}_{n,-1}\boldsymbol{\Gamma}^{0\prime}_{xn}(\boldsymbol{\Upsilon}^0_{xkN})^{-1}\boldsymbol{\Gamma}^0_{xi}, \\
\tilde{\mathbf{V}}_{i,-1} &= \mathbf{V}_{i,-1} - \frac{1}{N}\sum_{n=1}^{N}\mathbf{V}_{n,-1}\boldsymbol{\Gamma}^{0\prime}_{xn}(\boldsymbol{\Upsilon}^0_{xkN})^{-1}\boldsymbol{\Gamma}^0_{xi}, \\
\mathbf{b}_{12NT} &= \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\tilde{\mathbf{V}}'_{i,-1}\mathbf{V}_{j,-1}}{T}\boldsymbol{\Gamma}^{0\prime}_{xj}(\boldsymbol{\Upsilon}^0_{xkN})^{-1}\left(\frac{\mathbf{F}^{0\prime}_{x,-1}\mathbf{F}^0_{x,-1}}{T}\right)^{-1}\frac{\mathbf{F}^{0\prime}_{x,-1}\mathbf{u}_i}{T}, \\
\mathbf{b}_{22NT} &= -\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}^{0\prime}_{xi}(\boldsymbol{\Upsilon}^0_{xkN})^{-1}\left(\frac{\mathbf{F}^{0\prime}_{x,-1}\mathbf{F}^0_{x,-1}}{T}\right)^{-1}\mathbf{F}^{0\prime}_{x,-1}\bar{\boldsymbol{\Sigma}}_{kNT,-1}\mathbf{M}_{F^0_{x,-1}}\mathbf{u}_i.
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\hat{\mathbf{Z}}'_i\mathbf{u}_i &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\left[\mathbf{M}_{\hat{F}_x}\mathbf{X}_i, \mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1}\right]'\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\left[\mathbf{M}_{F^0_x}\tilde{\mathbf{X}}_i, \mathbf{M}_{F^0_{x,-1}}\tilde{\mathbf{X}}_{i,-1}\right]'\mathbf{u}_i + \sqrt{\frac{T}{N}}[\mathbf{b}'_{11NT}, \mathbf{b}'_{12NT}]' \\
&+ \sqrt{\frac{N}{T}}[\mathbf{b}'_{21NT}, \mathbf{b}'_{22NT}]' + o_p(1) \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\tilde{\mathbf{Z}}'_i\mathbf{u}_i + \sqrt{\frac{T}{N}}\mathbf{b}_{1NT} + \sqrt{\frac{N}{T}}\mathbf{b}_{2NT} + o_p(1),
\end{aligned}
$$

where $\tilde{\mathbf{Z}}_i = \left[\mathbf{M}_{F^0_x}\tilde{\mathbf{X}}_i, \mathbf{M}_{F^0_{x,-1}}\tilde{\mathbf{X}}_{i,-1}\right]$, $\mathbf{b}_{1NT} = [\mathbf{b}'_{11NT}, \mathbf{b}'_{12NT}]'$ and $\mathbf{b}_{2NT} = [\mathbf{b}'_{21NT}, \mathbf{b}'_{22NT}]'$, which provides the expression given in Proposition 1. ■

Proof of Proposition 2. Now consider

$$
\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\hat{\mathbf{Z}}'_i\mathbf{M}_{\hat{F}_y}\mathbf{u}_i, \tag{A.25}
$$

where $\hat{\mathbf{Z}}_i = \left[\mathbf{M}_{\hat{F}_x}\mathbf{X}_i, \mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1}\right]$. We start with the first component of $\hat{\mathbf{Z}}_i$, i.e. $\mathbf{M}_{\hat{F}_x}\mathbf{X}_i$, which can be written as

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_i'\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_y}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\mathbf{F}_x^{0\prime}\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_y}\mathbf{u}_i + \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_y}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_y}\mathbf{u}_i + o_p(1) \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{F_x^0}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i + o_p(1), \qquad (A.26)
\end{aligned}$$

where the second and third equalities is due to Lemma 5 and 9, respectively.

As for the second component of $\hat{\mathbf{Z}}_i$, which is $\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1}$, by following the same steps as before and using again Lemmas 5 and 9, we yield

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{X}_{i,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_y}\mathbf{u}_i = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_{i,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i + o_p(1). \qquad (A.27)$$

By combining the results above, we obtain the required expression. ■

Lemma 10 *Under Assumptions 1-5, as* $(N,T)\xrightarrow{j}\infty$ *such that* $N/T\to c$ *with* $0<c<\infty$,

$$\begin{aligned}
T^{-1/2}\mathbf{X}_i'\left(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0}\right)\mathbf{u}_i &= \sqrt{T}O_p\left(\delta_{NT}^{-2}\right), \qquad (A.28)\\
T^{-1/2}\mathbf{X}_i'\mathbf{M}_{\hat{F}_{x,-1}}\left(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0}\right)\mathbf{u}_i &= \sqrt{T}O_p\left(\delta_{NT}^{-2}\right), \qquad (A.29)\\
T^{-1/2}\mathbf{X}_i'\left(\mathbf{M}_{\hat{F}_{x,-1}}-\mathbf{M}_{F_{x,-1}^0}\right)\mathbf{M}_{F_x^0}\mathbf{u}_i &= \sqrt{T}O_p\left(\delta_{NT}^{-2}\right). \qquad (A.30)
\end{aligned}$$

Proof of Proposition 3. Consider $T^{-1/2}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i$ where $\hat{\mathbf{Z}}_i=\left[\mathbf{M}_{\hat{F}_x}\mathbf{X}_i,\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1}\right]$. Let us start with the first component of $\mathbf{M}_{\hat{F}_x}\hat{\mathbf{Z}}_i$, i.e. $\mathbf{M}_{\hat{F}_x}\mathbf{X}_i$. By adding and subtracting we get

$$\begin{aligned}
T^{-1/2}\mathbf{X}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i &= T^{-1/2}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i + T^{-1/2}\mathbf{X}_i'\left(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0}\right)\mathbf{u}_i\\
&= T^{-1/2}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i + \sqrt{T}O_p\left(\delta_{NT}^{-2}\right) \qquad (A.31)
\end{aligned}$$

where the second equality is due to result in (A.28) stated in Lemma 10.

Next is the second component of $\mathbf{M}_{\hat{F}_x}\hat{\mathbf{Z}}_i$, which is $\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1}$. Again, by adding and subtracting and using Lemma 10, we get

$$\begin{aligned}
&T^{-1/2}\mathbf{X}_{i,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i\\
&= T^{-1/2}\mathbf{X}_{i,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{F_x^0}\mathbf{u}_i + T^{-1/2}\mathbf{X}_i'\mathbf{M}_{\hat{F}_{x,-1}}\left(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0}\right)\mathbf{u}_i\\
&= T^{-1/2}\mathbf{X}_{i,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{F_x^0}\mathbf{u}_i + \sqrt{T}O_p\left(\delta_{NT}^{-2}\right)\\
&= T^{-1/2}\mathbf{X}_{i,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_x^0}\mathbf{u}_i + T^{-1/2}\mathbf{X}_{i,-1}'\left(\mathbf{M}_{\hat{F}_{x,-1}}-\mathbf{M}_{F_{x,-1}^0}\right)\mathbf{M}_{F_x^0}\mathbf{u}_i + \sqrt{T}O_p\left(\delta_{NT}^{-2}\right)\\
&= T^{-1/2}\mathbf{X}_{i,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_x^0}\mathbf{u}_i + \sqrt{T}O_p\left(\delta_{NT}^{-2}\right). \qquad (A.32)
\end{aligned}$$

Finally, by combining the results, we get

$$T^{-1/2}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i = T^{-1/2}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i + \sqrt{T}O_p\left(\delta_{NT}^{-2}\right), \qquad (A.33)$$

where $\mathbf{Z}_i=\left[\mathbf{M}_{F_x^0}\mathbf{X}_i,\mathbf{M}_{F_{x,-1}^0}\mathbf{X}_{i,-1}\right]$, which provides the expression given in Proposition 3. ■

Lemma 11 *Under Assumptions 1-5, as* $(N,T)\xrightarrow{j}\infty$ *such that* $N/T\to c$ *with* $0<c<\infty$, $\frac{1}{NT}\sum_{i=1}^{N}\hat{\boldsymbol{\xi}}_{\hat{F}iT}\hat{\boldsymbol{\xi}}_{\hat{F}iT}'=\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\xi}_{\hat{F}iT}\boldsymbol{\xi}_{\hat{F}iT}'+o_p(1)$, *where* $\boldsymbol{\xi}_{\hat{F}iT}=\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{u}_i$ *and* $\hat{\boldsymbol{\xi}}_{\hat{F}iT}=\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{u}}_i$.

Lemma 12 *Under Assumptions 1-5, as* $(N,T)\xrightarrow{j}\infty$ *such that* $N/T\to c$ *with* $0<c<\infty$, $\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\xi}_{\hat{F}iT}\boldsymbol{\xi}_{\hat{F}iT}'-\boldsymbol{\Omega}=o_p(1)$, *where* $\boldsymbol{\Omega}=\operatorname{plim}_{N,T\to\infty}\frac{1}{N}\sum_{i=1}^{N}E\left(T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_i\right)$.

Proposition 4 *Under Assumptions 1-5, as* $(N,T)\xrightarrow{j}\infty$ *such that* $N/T\to c$ *with* $0<c<\infty$,

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\xi}_{\hat{F}iT}\xrightarrow{d}N(\mathbf{0},\boldsymbol{\Omega}).$$

Proof. Proposition 2 and Lemma 12, together with Lemma B.2, yield the required result. ■

Lemma 13 *Under Assumptions 1-5, as* $(N,T)\xrightarrow{j}\infty$ *such that* $N/T\to c$ *with* $0<c<\infty$, $\hat{\mathbf{A}}_{NT}\xrightarrow{p}\mathbf{A}$, $\mathbf{B}_{NT}\xrightarrow{p}\mathbf{B}$, *where* $\hat{\mathbf{A}}_{NT}=\frac{1}{N}\sum_{i=1}^{N}T^{-1}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{W}_i$, $\hat{\mathbf{B}}_{NT}=\frac{1}{N}\sum_{i=1}^{N}T^{-1}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{Z}}_i$ *and* $\mathbf{A}=\lim_{N,T\to\infty}\frac{1}{N}\sum_{i=1}^{N}E(\mathbf{A}_{i,T})$, $\mathbf{B}=\lim_{N,T\to\infty}\frac{1}{N}\sum_{i=1}^{N}E(\mathbf{B}_{i,T})$, $\mathbf{A}_{i,T}=T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_y^0}\mathbf{W}_i$, $\mathbf{B}_{i,T}=T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_i$.

Lemma 14 (Lemma 2.2.10 of Van der Vaart and Wellner (1996)) *Let* $x_1,\cdots,x_N$ *be arbitrary random variables that satisfy the tail bound:*

$$P(|x_i|>z)\leq 2\exp\left(-\frac{1}{2}\times\frac{z^2}{a+bz}\right)$$

for all z *(and all* i*) and fixed* $a,b>0$. *Then,*

$$E\Big|\sup_{1\leq i\leq N}x_i\Big|\leq\Delta\left(b\times\ln(N+1)+\sqrt{a\times\ln(N+1)}\right)$$

for some positive constant Δ.

Lemma 15 *Under Assumptions 2 to 4, and Assumption 7, we have*

(a) $N^{-1}T^{-1}\sum_{i=1}^{N}\|\mathbf{X}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i-\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i\|=O_p(\delta_{NT}^{-2})$.

(b) $N^{-1}T^{-1}\sum_{i=1}^{N}\|\mathbf{X}_{i,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i-\mathbf{X}_{i,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_x^0}\mathbf{u}_i\|=O_p(\delta_{NT}^{-2})$.

(c) $\sup_{1\leq i,j\leq N}\|T^{-1}\mathbf{X}_j'\mathbf{M}_{\hat{F}_x}\mathbf{X}_i-T^{-1}\mathbf{X}_j'\mathbf{M}_{F_x^0}\mathbf{X}_i\|=O_p(N^{1/2}\delta_{NT}^{-2})$.

(d) $\sup_{1\leq i,j\leq N}\|T^{-1}\mathbf{X}_{j,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1}-T^{-1}\mathbf{X}_{j,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_x^0}\mathbf{M}_{F_{x,-1}^0}\mathbf{X}_{i,-1}\|=O_p(N^{1/2}\delta_{NT}^{-2})$.

(e) $\sup_{1\leq i,j\leq N}\|T^{-1}\mathbf{X}_{j,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_x}\mathbf{X}_i-T^{-1}\mathbf{X}_{j,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_x^0}\mathbf{X}_i\|=O_p(N^{1/2}\delta_{NT}^{-2})$.

Lemma 16 *Under Assumptions 1 to 7, we have*

$$(a)\quad \sup_{1\le i,j\le N}\|T^{-1}\mathbf{X}_i'\mathbf{M}_{\hat{F}_x}\mathbf{y}_{i,-1}-T^{-1}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{y}_{i,-1}\|$$
$$=O_p(N^{1/2}\delta_{NT}^{-2})+O_p(N^{3/4}T^{-1/2}\delta_{NT}^{-2})+O_p(NT^{-1}\delta_{NT}^{-2})+O_p(N^{1/4}T^{-1/2})\,.$$
$$(b)\quad \sup_{1\le i,j\le N}\|T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_x}\mathbf{y}_{i,-1}-T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_x^0}\mathbf{y}_{i,-1}\|$$
$$=O_p(N^{1/2}\delta_{NT}^{-2})+O_p(N^{3/4}T^{-1/2}\delta_{NT}^{-2})+O_p(NT^{-1}\delta_{NT}^{-2})+O_p(N^{1/4}T^{-1/2})\,.$$

Lemma 17 *Under Assumptions 1 to 7, we have*

$$(a)\quad \sup_{1\le i\le N}\|T^{-1}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{X}_i-T^{-1}E(\mathbf{V}_i'\mathbf{V}_i)\|=O_p(N^{1/4}T^{-1/2})+O_p(N^{1/2}T^{-1})\,,$$
$$(b)\quad \sup_{1\le i\le N}\|T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_x^0}\mathbf{M}_{F_{x,-1}^0}\mathbf{X}_{i,-1}-T^{-1}E(\mathbf{V}_{i,-1}'\mathbf{V}_{i,-1})\|=O_p(N^{1/4}T^{-1/2})+O_p(N^{1/2}T^{-1})\,,$$
$$(c)\quad \sup_{1\le i\le N}\|T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_x^0}\mathbf{X}_i-T^{-1}E(\mathbf{V}_{i,-1}'\mathbf{V}_i)\|=O_p(N^{1/4}T^{-1/2})+O_p(N^{1/2}T^{-1})\,,$$
$$(d)\quad \sup_{1\le i\le N}\|T^{-1}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{y}_{i,-1}-T^{-1}\sum_{s=1}^{\infty}E(\mathbf{V}_i'\mathbf{V}_{i,-s})\boldsymbol{\beta}_i\rho_i^{s-1}\|=O_p(N^{1/4}T^{-1/2})+O_p(N^{1/2}T^{-1})\,.$$
$$(e)\quad \sup_{1\le i\le N}\|T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_x^0}\mathbf{y}_{i,-1}-T^{-1}\sum_{s=1}^{\infty}E(\mathbf{V}_{i,-1}'\mathbf{V}_{i,-s})\boldsymbol{\beta}_i\rho_i^{s-1}\|=O_p(N^{1/4}T^{-1/2})+O_p(N^{1/2}T^{-1})\,.$$

Lemma 18 *Define*

$$\mathbf{A}_{i,T}=\begin{pmatrix}T^{-1}\sum_{s=1}^{\infty}\rho_i^{s-1}E(\mathbf{V}_i'\mathbf{V}_{i,-s})\boldsymbol{\beta}_i & T^{-1}E(\mathbf{V}_i'\mathbf{V}_i)\\ T^{-1}\sum_{s=1}^{\infty}\rho_i^{s-1}E(\mathbf{V}_{i,-1}'\mathbf{V}_{i,-s})\boldsymbol{\beta}_i & T^{-1}E(\mathbf{V}_{i,-1}'\mathbf{V}_i)\end{pmatrix}$$
$$\mathbf{B}_{i,T}=\begin{pmatrix}T^{-1}E(\mathbf{V}_i'\mathbf{V}_i) & T^{-1}E(\mathbf{V}_i'\mathbf{V}_{i,-1})\\ T^{-1}E(\mathbf{V}_{i,-1}'\mathbf{V}_i) & T^{-1}E(\mathbf{V}_{i,-1}'\mathbf{V}_{i,-1})\end{pmatrix}$$

under Assumptions 1 to 7, we have

$$(a)\quad \sup_{1\le i\le N}\|(\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}-(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\|$$
$$=O_p(N^{1/4}T^{-1/2}\ln N)+O_p(N^{1/2}(\ln N)^5\delta_{NT}^{-2})\,,$$
$$(b)\quad \sup_{1\le i\le N}\|(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\|=O_p((\ln N)^2)\,,$$
$$(c)\quad \sup_{1\le i\le N}\|[(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}-(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T})^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}]\|$$
$$=O_p(N^{1/4}T^{-1/2}(\ln N)^5)+O_p(N^{1/2}T^{-1}(\ln N)^5)\,.$$

Proof of Theorem 1. By using the expression in (16), the result of Proposition 1 from which $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\hat{\mathbf{Z}}_i'\mathbf{u}_i$ tends to a multivatiate random variable and is therefore $O_p(1)$, and $\sqrt{\frac{T}{N}}\mathbf{b}_{1NT}$ together with $\sqrt{\frac{N}{T}}\mathbf{b}_{2NT}$ are $O_p(1)$ as T/N tends to a finite positive constant c $(0<c<\infty)$ when N and $T\to\infty$ jointly. And so, $\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{IV}-\boldsymbol{\theta}\right)=O_p(1)$, which implies the required result. ■

Proof of Theorem 2. (i) $\sqrt{NT}\left(\hat{\hat{\boldsymbol{\theta}}}_{IV}-\boldsymbol{\theta}\right)=\left(\hat{\mathbf{A}}_{NT}'\hat{\mathbf{B}}_{NT}^{-1}\hat{\mathbf{A}}_{NT}\right)^{-1}\hat{\mathbf{A}}_{NT}'\hat{\mathbf{B}}_{NT}^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\xi}_{\hat{F}iT}\right)$ $=\left(\mathbf{A}'\mathbf{B}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}'\mathbf{B}^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\boldsymbol{\xi}_{FiT}\right)+o_p(1)$, by the results of Proposition 2 and Lemma 13. Next, by the result of Proposition 4, we have $\sqrt{NT}\left(\hat{\hat{\boldsymbol{\theta}}}_{IV}-\boldsymbol{\theta}\right)\xrightarrow{d}N(\mathbf{0},\boldsymbol{\Psi})$, as required. (ii) $\hat{\boldsymbol{\Psi}}-\boldsymbol{\Psi}=o_p(1)$ follows immediately from Lemmas 11, 12 and 13. ■

Proof of Theorem 3. Under Assumptions 1-5, noting $\hat{\mathbf{u}}_i = \mathbf{u}_i - \mathbf{W}_i\left(\hat{\hat{\boldsymbol{\theta}}}_{IV2} - \boldsymbol{\theta}\right)$ we have $\frac{1}{\sqrt{NT}}\sum_{i=1}^N \hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{u}}_i = \frac{1}{\sqrt{NT}}\sum_{i=1}^N \hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{u}_i - \hat{\mathbf{A}}_{NT}\sqrt{NT}\left(\hat{\hat{\boldsymbol{\theta}}}_{IV2} - \boldsymbol{\theta}\right)$. Since $\sqrt{NT}\left(\hat{\hat{\boldsymbol{\theta}}}_{IV2} - \boldsymbol{\theta}\right) = \left(\mathbf{A}'\boldsymbol{\Omega}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}'\boldsymbol{\Omega}^{-1}\frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{Z}_i'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i$ $+o_p(1)$ by Corollary 1 and defining $\mathbf{L} = \boldsymbol{\Omega}^{-1/2}\mathbf{A}$ we have $\hat{\boldsymbol{\Omega}}_{NT}^{-1/2}\frac{1}{\sqrt{NT}}\sum_{i=1}^N \hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{u}}_i = \mathbf{M}_L\boldsymbol{\Omega}^{-1/2}\frac{1}{\sqrt{NT}}\sum_{i=1}^N \mathbf{Z}_i'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i$ $+o_p(1)$ with $\mathbf{M}_L = \mathbf{I}_T - \mathbf{L}\left(\mathbf{L}'\mathbf{L}\right)^{-1}\mathbf{L}'$ whose rank is $k-1$, which yields $\frac{1}{NT}\sum_{i=1}^N \hat{\mathbf{u}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{Z}}_i'\hat{\boldsymbol{\Omega}}_{NT}^{-1}\sum_{i=1}^N \hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{u}}_i \xrightarrow{d} \chi^2_{k-1}$ as required. ■

Proof of Theorem 4. By Proposition 3 $T^{-1/2}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i = T^{-1/2}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i + o_p(1)$ as $(N,T) \xrightarrow{j} \infty$ as $N/T \to c$ for $0 < c < \infty$. It is immediate that, under Assumptions 1-8, for each i, $T^{-1/2}\mathbf{Z}_i'\mathbf{u}_i \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}_i)$. A similar line of the argument in the proof of Lemma 7 ensures that $\hat{\tilde{\mathbf{A}}}_{i,T} - \tilde{\mathbf{A}}_{i,T} \xrightarrow{p} \mathbf{0}$ and $\hat{\tilde{\mathbf{B}}}_{i,T} - \tilde{\mathbf{B}}_{i,T} \xrightarrow{p} \mathbf{0}$ as $T \to \infty$, and togther with Assumption 8 we see that $p\lim_{T\to\infty}\hat{\tilde{\mathbf{A}}}_{i,T} = \mathbf{A}_i$ and $p\lim_{T\to\infty}\hat{\tilde{\mathbf{B}}}_{i,T} = \mathbf{B}_i$, thus the required result follows. ■

Proof of Theorem 5. Note that the instrumental variable (IV) or two-stage least squares estimator of $\boldsymbol{\theta}_i$ is $\hat{\boldsymbol{\theta}}_{IV,i} = (\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{g}}}_{i,T}$, then we have

$$\widehat{\boldsymbol{\theta}}_{IVMG} - \boldsymbol{\theta} = N^{-1}\sum_{i=1}^N(\widehat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}) = N^{-1}\sum_{i=1}^N(\widehat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i) + N^{-1}\sum_{i=1}^N\boldsymbol{\eta}_i$$

where the first term is

$$\begin{aligned}
&N^{-1}\sum_{i=1}^N(\widehat{\boldsymbol{\theta}}_{IV,i} - \boldsymbol{\theta}_i) = N^{-1}\sum_{i=1}^N(\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}(T^{-1}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i)\\
=&N^{-1}\sum_{i=1}^N(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}(T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i)\\
&+N^{-1}\sum_{i=1}^N\left[(\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1} - (\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\right](T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i)\\
&+N^{-1}\sum_{i=1}^N\left[(\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1} - (\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\right](T^{-1}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i - T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i)\\
&+N^{-1}\sum_{i=1}^N(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}(T^{-1}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i - T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i)\\
=&\mathbb{G}_1 + \mathbb{G}_2 + \mathbb{G}_3 + \mathbb{G}_4
\end{aligned}$$

We first consider the terms $\mathbb{G}_2$, $\mathbb{G}_3$, and $\mathbb{G}_4$. With Lemma 15 (a)-(b), we have

$$N^{-1}T^{-1}\sum_{i=1}^N\|T^{-1}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i - T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i\| = O_p(\delta_{NT}^{-2}) \tag{A.34}$$

With the above facts, $\mathbb{G}_2$ is bounded in norm by

$$\begin{aligned}
&N^{-1}\sum_{i=1}^N\|T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i\| \cdot \sup_{1\le i\le N}\|(\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1} - (\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\|\\
=&O_p(N^{1/4}T^{-1}\ln N) + O_p(N^{1/2}T^{-1/2}(\ln N)^5\delta_{NT}^{-2})
\end{aligned}$$

Analogously, with (A.34), we can show that $\mathbb{G}_3 = O_p(N^{1/4}T^{-1/2}\ln N\delta_{NT}^{-2}) + O_p(N^{1/2}(\ln N)^5\delta_{NT}^{-4})$ and $\mathbb{G}_4 = O_p((\ln N)^2\delta_{NT}^{-2})$. Consider $\mathbb{G}_1$. Define

$$\begin{aligned}
\mathbb{H}_{1i} &= \begin{pmatrix}\mathbf{V}_i'\mathbf{F}_y\boldsymbol{\gamma}_{yi}\\ \mathbf{V}_{i,-1}'\mathbf{F}_y\boldsymbol{\gamma}_{yi}\end{pmatrix}, \mathbb{H}_{2i} = \begin{pmatrix}-\mathbf{V}_i'\mathbf{F}_x^0(\mathbf{F}_x^{0\prime}\mathbf{F}_x^0)^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_y\boldsymbol{\gamma}_{yi}\\ -\mathbf{V}_{i,-1}'\mathbf{F}_{x,-1}^0(\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_{x,-1}^0)^{-1}\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_y\boldsymbol{\gamma}_{yi}\end{pmatrix},\\
\mathbb{H}_{3i} &= \begin{pmatrix}\mathbf{V}_i'\boldsymbol{\varepsilon}_i\\ \mathbf{V}_{i,-1}'\boldsymbol{\varepsilon}_i\end{pmatrix}, \mathbb{H}_{4i} = \begin{pmatrix}-\mathbf{V}_i'\mathbf{F}_x^0(\mathbf{F}_x^{0\prime}\mathbf{F}_x^0)^{-1}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i\\ -\mathbf{V}_{i,-1}'\mathbf{F}_{x,-1}^0(\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_{x,-1}^0)^{-1}\mathbf{F}_{x,-1}^{0\prime}\boldsymbol{\varepsilon}_i\end{pmatrix},\\
\mathbb{H}_{5i} &= \begin{pmatrix}\mathbf{0}\\ \mathbf{V}_{i,-1}'\mathbf{P}_{\mathbf{F}_x^0}\mathbf{F}_y\boldsymbol{\gamma}_{yi}\end{pmatrix}, \mathbb{H}_{6i} = \begin{pmatrix}\mathbf{0}\\ -\mathbf{V}_{i,-1}'\mathbf{F}_{x,-1}^0(\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_{x,-1}^0)^{-1}\mathbf{F}_{x,-1}^{0\prime}\mathbf{P}_{\mathbf{F}_x^0}\mathbf{F}_y\boldsymbol{\gamma}_{yi}\end{pmatrix},\\
\mathbb{H}_{7i} &= \begin{pmatrix}\mathbf{0}\\ \mathbf{V}_{i,-1}'\mathbf{P}_{\mathbf{F}_x^0}\boldsymbol{\varepsilon}_i\end{pmatrix}, \mathbb{H}_{8i} = \begin{pmatrix}\mathbf{0}\\ -\mathbf{V}_{i,-1}'\mathbf{F}_{x,-1}^0(\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_{x,-1}^0)^{-1}\mathbf{F}_{x,-1}^{0\prime}\mathbf{P}_{\mathbf{F}_x^0}\boldsymbol{\varepsilon}_i\end{pmatrix},
\end{aligned}$$

with $\mathbf{u}_i = \mathbf{F}_y \boldsymbol{\gamma}_{yi} + \boldsymbol{\varepsilon}_i$, we have

$$\begin{aligned}\mathbb{G}_1 =& N^{-1}\sum_{i=1}^{N}[\left(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T}\right)^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1} - \left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}](T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i)\\ &+\sum_{\ell=1}^{8} N^{-1}\sum_{i=1}^{N}\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}T^{-1}\mathbb{H}_{\ell i}.\end{aligned}$$

With (18), we can show that the first term is $O_p(N^{1/4}(\ln N)^5T^{-1}) + O_p(N^{1/2}(\ln N)^5T^{-3/2})$. It's easy to show that the fifth term, the eighth term and the ninth term both are $O_p((\ln N)^2T^{-1})$. For the second term, we have

$$\begin{aligned}&E\|N^{-1}\sum_{i=1}^{N}\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}T^{-1}\mathbb{H}_{1i}\|^2\\ =&\mathrm{tr}\Big(N^{-2}T^{-2}\sum_{i\neq j}E\Big[\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\Big]\mathbf{B}_{i,T}^{-1}E\Big[\mathbb{H}_{1i}\mathbb{H}_{1j}'\Big]\mathbf{B}_{j,T}^{-1}E\Big[\mathbf{A}_{j,T}\left(\mathbf{A}_{j,T}'\mathbf{B}_{j,T}^{-1}\mathbf{A}_{j,T}\right)^{-1}\Big]\Big)\\ &+\mathrm{tr}\Big(N^{-2}T^{-2}\sum_{i=1}^{N}E\Big[\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}E\Big[\mathbb{H}_{1i}\mathbb{H}_{1i}'\Big]\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\Big]\Big)\\ \leq&\Delta\Big\|N^{-2}T^{-2}\sum_{i=1}^{N}E\mathrm{vec}\Big[\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}E\Big[\mathbb{H}_{1i}\mathbb{H}_{1i}'\Big]\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\Big]\Big\|\\ =&\Delta\Big\|N^{-2}T^{-2}\sum_{i=1}^{N}E\Big(\Big[\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\Big]\otimes\Big[\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\Big]\Big)\mathbf{vec}(E\Big[\mathbb{H}_{1i}\mathbb{H}_{1i}'\Big])\Big\|\\ \leq&\Delta N^{-2}T^{-2}\sum_{i=1}^{N}\Big\|E\Big(\Big[\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\Big]\otimes\Big[\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\Big]\Big)\Big\|\Big\|E(\mathbb{H}_{1i}\mathbb{H}_{1i}')\Big\|\\ \leq&\Delta N^{-2}T^{-2}\sum_{i=1}^{N}E\|(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T})^{-1}\mathbf{A}_{i,T}'\|^2\|\mathbf{B}_{i,T}^{-1}\|\|E(\mathbb{H}_{1i}\mathbb{H}_{1i}')\| \leq \Delta N^{-2}T^{-2}\sum_{i=1}^{N}\|E(\mathbb{H}_{1i}\mathbb{H}_{1i}')\|\\ \leq&\Delta N^{-2}T^{-2}\sum_{i=1}^{N}\sum_{s=1}^{T}\sum_{t=1}^{T}(\|E(\mathbf{v}_{is}\mathbf{v}_{it}')\| + 2\|E(\mathbf{v}_{is}\mathbf{v}_{i,t-1}')\| + \|E(\mathbf{v}_{i,s-1}\mathbf{v}_{i,t-1}')\|) \leq \Delta N^{-1}T^{-1}\end{aligned}$$

because

$$|E(\mathbf{f}_{y,s}\boldsymbol{\gamma}_{yi}\boldsymbol{\gamma}_{yj}'\mathbf{f}_{y,t}')| \leq \sqrt{E\|\mathbf{f}_{y,s}\|^4E\|\boldsymbol{\gamma}_{yi}\|^4} \leq \Delta$$

and

$$\begin{aligned}&E(\mathbb{H}_{1i}\mathbb{H}_{1j}')\\ =&\begin{pmatrix}\sum_{s=1}^{T}\sum_{t=1}^{T}E(\mathbf{v}_{is}\mathbf{v}_{jt}')E(\mathbf{f}_{y,s}\boldsymbol{\gamma}_{yi}\boldsymbol{\gamma}_{yj}'\mathbf{f}_{y,t}') & \sum_{s=1}^{T}\sum_{t=1}^{T}E(\mathbf{v}_{is}\mathbf{v}_{j,t-1}')E(\mathbf{f}_{y,s}\boldsymbol{\gamma}_{yi}\boldsymbol{\gamma}_{yj}'\mathbf{f}_{y,t}')\\ \sum_{s=1}^{T}\sum_{t=1}^{T}E(\mathbf{v}_{i,s-1}\mathbf{v}_{j,t}')E(\mathbf{f}_{y,s}\boldsymbol{\gamma}_{yi}\boldsymbol{\gamma}_{yj}'\mathbf{f}_{y,t}') & \sum_{s=1}^{T}\sum_{t=1}^{T}E(\mathbf{v}_{i,s-1}\mathbf{v}_{j,t-1}')E(\mathbf{f}_{y,s}\boldsymbol{\gamma}_{yi}\boldsymbol{\gamma}_{yj}'\mathbf{f}_{y,t}')\end{pmatrix},\end{aligned}$$

which indicates that $E(\mathbb{H}_{1i}\mathbb{H}_{1j}') = 0$ for $i \neq j$, then the second term is $O_p(N^{-1/2}T^{-1/2})$. Consider the third term. Note that

$$\begin{aligned}\mathbb{H}_{2i} = \mathrm{vec}(\mathbb{H}_{2i}) &= \begin{pmatrix}\boldsymbol{\gamma}_{yi}'\otimes(\mathbf{V}_i'\mathbf{F}_x^0)\mathrm{vec}\big[(\mathbf{F}_x^{0\prime}\mathbf{F}_x^0)^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_y\big]\\ \boldsymbol{\gamma}_{yi}'\otimes(\mathbf{V}_{i,-1}'\mathbf{F}_{x,-1}^0)\mathrm{vec}\big[(\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_{x,-1}^0)^{-1}\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_y\big]\end{pmatrix}\\ &= \begin{pmatrix}\boldsymbol{\gamma}_{yi}'\otimes(\mathbf{V}_i'\mathbf{F}_x^0) & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\gamma}_{yi}'\otimes(\mathbf{V}_{i,-1}'\mathbf{F}_{x,-1}^0)\end{pmatrix}\begin{pmatrix}\mathrm{vec}\big[(\mathbf{F}_x^{0\prime}\mathbf{F}_x^0)^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_y\big]\\ \mathrm{vec}\big[(\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_{x,-1}^0)^{-1}\mathbf{F}_{x,-1}^{0\prime}\mathbf{F}_y\big]\end{pmatrix}\\ &= \mathbb{H}_{2ia}\times\mathbb{H}_{2ib}.\end{aligned}$$

It's easy to prove that $\mathbb{H}_{2ib} = O_p(1)$. Following the argument in the proof of the second term, we can prove that

$$-N^{-1}\sum_{i=1}^{N}\left(\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\right)^{-1}\mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}T^{-1}\mathbb{H}_{2ia} = O_p(N^{-1/2}T^{-1/2}).$$

Then the third term is $O_p(N^{-1/2}T^{-1/2})$. Analogously, the forth term, the sixth term and the seventh term can be proved to be $O_p(N^{-1/2}T^{-1/2})$. Thus, $\mathbb{G}_1 = O_p(N^{1/4}(\ln N)^5T^{-1})+O_p(N^{1/2}(\ln N)^5T^{-3/2})+O_p(N^{-1/2}T^{-1/2})$.

Combining the above terms, we can show that

$$N^{-1}\sum_{i=1}^{N}(\widehat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i)=O_p(N^{1/4}(\ln N)^5T^{-1})+O_p(N^{1/2}(\ln N)^5T^{-3/2})+O_p((\ln N)^2\delta_{NT}^{-2}).$$

Note that $N^{-1}\sum_{i=1}^{N}\boldsymbol{\eta}_i=O_p(N^{-1/2})$, if $N^{3+\delta}/T^4\to 0$ for any $\delta>0$, we have

$$\sqrt{N}(\widehat{\boldsymbol{\theta}}_{IVMG}-\boldsymbol{\theta})=N^{-1/2}\sum_{i=1}^{N}\boldsymbol{\eta}_i+o_p(1)$$

and

$$\sqrt{N}(\widehat{\boldsymbol{\theta}}_{IVMG}-\boldsymbol{\theta})\xrightarrow{d}N(\mathbf{0},\boldsymbol{\Sigma}_\eta).$$

Next, we consider the consistency of $\widehat{\boldsymbol{\Sigma}}_\eta$. By decomposition, we have

$$\begin{aligned}&\sum_{i=1}^{N}(\hat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}+\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IVMG})(\hat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}+\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IVMG})'\\&=\sum_{i=1}^{N}\boldsymbol{\eta}_i\boldsymbol{\eta}_i'+\sum_{i=1}^{N}(\hat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i)(\hat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i)'+\sum_{i=1}^{N}(\hat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i)\boldsymbol{\eta}_i'+\sum_{i=1}^{N}\boldsymbol{\eta}_i(\hat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i)'\\&\quad-N(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IVMG})'(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IVMG}).\end{aligned}$$

Then

$$\begin{aligned}&\widehat{\boldsymbol{\Sigma}}_\eta-\boldsymbol{\Sigma}_\eta\\=&\frac{1}{N-1}\sum_{i=1}^{N}(\boldsymbol{\eta}_i\boldsymbol{\eta}_i'-\boldsymbol{\Sigma}_\eta)+\frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i\right)\left(\hat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i\right)'+\frac{1}{N-1}\sum_{i=1}^{N}\left(\hat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i\right)\boldsymbol{\eta}_i'\\&+\frac{1}{N-1}\sum_{i=1}^{N}\boldsymbol{\eta}_i\left(\hat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i\right)'-\frac{N}{N-1}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IVMG})'(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IVMG})\\=&\mathbb{J}_1+\cdots+\mathbb{J}_5.\end{aligned}$$

Easily, we can derive that $\mathbb{J}_1=O_p(N^{-1/2})$, $\mathbb{J}_5=O_p(N^{-1})$. Consider $\mathbb{J}_3$, which is

$$\begin{aligned}&\frac{1}{N-1}\sum_{i=1}^{N}(\widehat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i)\boldsymbol{\eta}_i'\\=&\frac{1}{N-1}\sum_{i=1}^{N}\left(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T}\right)^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}(T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i)\boldsymbol{\eta}_i'\\&+\frac{1}{N-1}\sum_{i=1}^{N}\left[(\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}-(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\right](T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i)\boldsymbol{\eta}_i'\\&+\frac{1}{N-1}\sum_{i=1}^{N}\left[(\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}-(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\right](T^{-1}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i-T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i)\boldsymbol{\eta}_i'\\&+\frac{1}{N-1}\sum_{i=1}^{N}(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}(T^{-1}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i-T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i)\boldsymbol{\eta}_i'.\end{aligned}$$

With $\sup_{1\le i\le N}\|\boldsymbol{\eta}_i\|=O_p(N^{1/4})$, we can follow the argument in the proof of the terms $\mathbb{G}_2$ to $\mathbb{G}_4$, to prove that the second term is $O_p(N^{3/4}(\ln N)^5T^{-1/2}\delta_{NT}^{-2})$, the third term is $O_p(N^{3/4}(\ln N)^5\delta_{NT}^{-4})$, the forth is $O_p(N^{1/4}(\ln N)^2\delta_{NT}^{-2})$. By Lemma 18 (b), the first term is bounded in norm by

$$(N-1)^{-1}\sum_{i=1}^{N}\|T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i\|\cdot\sup_{1\le i\le N}\|(\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\|\cdot\sup_{1\le i\le N}\|\boldsymbol{\eta}_i\|=O_p(N^{1/4}(\ln N)^2T^{-1/2}).$$

Then $\mathbb{J}_3=O_p(N^{3/4}(\ln N)^5T^{-1/2}\delta_{NT}^{-2})+O_p(N^{3/4}(\ln N)^5\delta_{NT}^{-4})+O_p(N^{1/4}(\ln N)^2\delta_{NT}^{-2})+O_p(N^{1/4}(\ln N)^2T^{-1/2})$. $\mathbb{J}_4$ is the same order of $\mathbb{J}_3$ since it is transpose of $\mathbb{J}_3$.

Consider $\mathbb{J}_2$, with $(a+b+c+d)^2 \leq 4(a^2+b^2+c^2+d^2)$, it is bounded in norm by

$$
\begin{aligned}
&\frac{1}{N-1}\sum_{i=1}^{N}\|\widehat{\boldsymbol{\theta}}_{IV,i}-\boldsymbol{\theta}_i\|^2 \\
\leq&\frac{4}{N-1}\sum_{i=1}^{N}\|\left(\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\tilde{\mathbf{A}}_{i,T}\right)^{-1}\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}(T^{-1}\mathbf{Z}'_i\mathbf{M}_{F_x^0}\mathbf{u}_i)\|^2 \\
&+\frac{4}{N-1}\sum_{i=1}^{N}\left\|\left[(\hat{\tilde{\mathbf{A}}}'_{i,T}\hat{\tilde{\mathbf{B}}}^{-1}_{i,T}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}'_{i,T}\hat{\tilde{\mathbf{B}}}^{-1}_{i,T}-(\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\right](T^{-1}\mathbf{Z}'_i\mathbf{M}_{F_x^0}\mathbf{u}_i)\right\|^2 \\
&+\frac{4}{N-1}\sum_{i=1}^{N}\left\|\left[(\hat{\tilde{\mathbf{A}}}'_{i,T}\hat{\tilde{\mathbf{B}}}^{-1}_{i,T}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}'_{i,T}\hat{\tilde{\mathbf{B}}}^{-1}_{i,T}-(\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\right](T^{-1}\hat{\mathbf{Z}}'_i\mathbf{M}_{\hat{F}_x}\mathbf{u}_i-T^{-1}\mathbf{Z}'_i\mathbf{M}_{F_x^0}\mathbf{u}_i)\right\|^2 \\
&+\frac{4}{N-1}\sum_{i=1}^{N}\|(\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}(T^{-1}\hat{\mathbf{Z}}'_i\mathbf{M}_{\hat{F}_x}\mathbf{u}_i-T^{-1}\mathbf{Z}'_i\mathbf{M}_{F_x^0}\mathbf{u}_i)\|^2.
\end{aligned}
$$

By Lemma 18 (b), the first term is bounded in norm by

$$4(N-1)^{-1}\sum_{i=1}^{N}\|T^{-1}\mathbf{Z}'_i\mathbf{M}_{F_x^0}\mathbf{u}_i\|^2\cdot(\sup_{1\leq i\leq N}\|(\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\|)^2=O_p((\ln N)^4T^{-1})$$

similarly, we can show that the second term is $O_p(N(\ln N)^{10}T^{-1}\delta_{NT}^{-4})$. Following the argument in the proof of Lemma 15 (a) and (b), we can show that $N^{-1}\sum_{i=1}^{N}\|T^{-1}\hat{\mathbf{Z}}'_i\mathbf{M}_{\hat{F}_x}\mathbf{u}_i-T^{-1}\mathbf{Z}'_i\mathbf{M}_{F_x^0}\mathbf{u}_i\|^2=O_p(\delta_{NT}^{-4})$. Then similar to the argument in the proof of the first term, we can prove that the third term is $O_p(N(\ln N)^{10}\delta_{NT}^{-8})$ and the forth term is $O_p((\ln N)^4\delta_{NT}^{-4})$. Then $\mathbb{J}_2=O_p((\ln N)^4T^{-1})+O_p(N(\ln N)^{10}T^{-1}\delta_{NT}^{-4})+O_p(N(\ln N)^{10}\delta_{NT}^{-8})+O_p((\ln N)^4\delta_{NT}^{-4})$.

Combining the terms $\mathbb{J}_1$ to $\mathbb{J}_5$, we can derive that $\widehat{\mathbf{\Sigma}}_\eta-\mathbf{\Sigma}_\eta=o_p(1)$. Thus, we complete the proof. ■

References

Ahn, S., Horenstein, A., 2013. Eigenvalue ratio test for the number of factors. Econometrica 81, 1203–1227.

Ahn, S., Lee, Y., Schmidt, P., 2013. Panel data models with multiple time varying effects. Journal of Econometrics 174, 1–14.

Alvarez, J., Arellano, M., 2003. The time series and cross-section asymptotics of dynamic panel data estimators. Econometrica 71, 1121–1159.

Bai, J., 2003. Inferential theory for factor models of large dimensions. Econometrica 71, 135–173.

Bai, J., 2009a. Panel data models with interactive fixed effects. Econometrica 77, 1229–1279.

Bai, J., 2009b. Supplement to "panel data models with interactive fixed effects": technical details and proofs. Econometrica Supplemental Material.

Bai, J., Li, K., 2014. Theory and methods of panel data models with interactive effects. Annals of Statistics 42, 142–170.

Bai, J., Ng, S., 2010. Instrumental variable estimation in a data rich environment. Econometric Theory 26 (6), 1577–1606.

Bai, J., Ng, S., 2013. Principal components estimation and identification of static factors. Journal of Econometrics 176, 18–29.

Bai, J., Wang, P., 2016. Econometric analysis of large factor models. Annual Review of Economics 8, 53–80.

Bun, M., Kiviet, J., 2006. The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. Journal of Econometrics 132, 409–444.

Carneiro, P., Hansen, K., Heckman, J., 2003. Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college. NBER working paper No. w9546.

Chudik, A., Pesaran, M. H., 2015a. Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. Journal of Econometrics 188, 393–420.

Chudik, A., Pesaran, M. H., 2015b. Large panel data models with cross-sectional dependence: A survey. In: Baltagi, B. H. (Ed.), The Oxford Handbook of Panel Data. Oxford University Press.

Favero, C., Marcellino, M., Neglia, F., 2005. Principal components at work: The empirical analysis of monetary policy with large data sets. Journal of Applied Econometrics 20, 603–620.

Heckman, J., Stixrud, J., Urzua, S., 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labour Economics 24, 411–482.

Kapetanios, G., Marcellino, M., 2010. Factor-GMM estimation with large sets of possibly weak instruments. Computational Statistics & Data Analysis 54 (11), 2655 – 2675, the Fifth Special Issue on Computational Econometrics.

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

Moon, H. R., Weidner, M., 2017. Dynamic linear panel regression models with interactive fixed effects. Econometric Theory 33, 158–195.

Moon, H. R., Weidner, M., 2019. Nuclear norm regularized estimation of panel regression models. cemmap working paper 14/19.

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

Robertson, D., Sarafidis, V., 2015. IV estimation of panels with factor residuals. Journal of Econometrics 185, 526–541.

Sarafidis, V., Wansbeek, T., 2012. Cross-sectional dependence in panel data analysis. Econometric Reviews 31, 483–531.

Sarafidis, V., Yamagata, T., Robertson, D., 2009. A test of cross section dependence for a linear dynamic panel model with regressors. Journal of Econometrics 148, 149–161.

Van Der Vaart, W. A., Wellner, J. A., 1996. Weak Convergence and Empirical Processes: With Applications to Statistics. Springer-Verlag, New York.

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

Wooldridge, J., 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press.

Supplemental Material to

"Instrumental Variable Estimation of Dynamic Linear Panel Data Models with Defactored Regressors and a Multifactor Error Structure"

by Milda Norkutė, Vasilis Sarafidis, Takashi Yamagata and Guowei Cui

Appendix B: Proofs of Lemmas

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

Proof of Lemma B.1. See Proof of Lemma 1 in Appendix, Hansen (2007). ■

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

Proof of Lemma B.2. See Proof of Lemma 2 in Appendix, Hansen (2007). ■

Lemma B.3 *Under Assumptions 1-5, the following statements hold for* $\ell = 1,2,...,k$ *and* $s = 1,2,...,T$*:*

$$E\left\|\frac{1}{\sqrt{N}T}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbf{f}_{x,t}^{0}\left[v_{\ell i,t}v_{\ell i,s} - E\left(v_{\ell i,t}v_{\ell i,s}\right)\right]u_{is}\right\|^{2} \leq \Delta < \infty, \tag{B.1}$$

$$E\left\|\frac{1}{\sqrt{N}T}\sum_{i=1}^{N}\sum_{t=1}^{T}\sum_{s=1}^{T}\mathbf{f}_{x,t}^{0}\left[v_{\ell i,t}v_{\ell i,s} - E\left(v_{\ell i,t}v_{\ell i,s}\right)\right]\mathbf{f}_{x,s}^{0\prime}\right\|^{2} \leq \Delta < \infty. \tag{B.2}$$

Proof of Lemma B.3. The proof of Lemma B.3 can be obtained in a similar manner based on the proof of Lemma A.2 provided in Bai (2009a, p.1268). ■

Lemma B.4 *Under Assumptions 1-5, as* $(N,T) \overset{j}{\to} \infty$ *such that* $N/T \to c$ *with* $0 < c < \infty$*, for* $i = 1,2,..,N$ *and* $\ell = 1,2,..,k$*,*

$$T^{-r/2}\left\|\hat{\mathbf{F}}_{x} - \mathbf{F}_{x}^{0}\mathbf{G}_{x}\right\|^{r} = T^{-r/2}\sum_{t=1}^{T}\left\|\hat{\mathbf{f}}_{x,t} - \mathbf{G}_{x}'\mathbf{f}_{x,t}^{0}\right\|^{r} = O_{p}\left(\delta_{NT}^{-r}\right),\ r = 1,2, \tag{B.3}$$

$$\frac{\left(\hat{\mathbf{F}}_{x} - \mathbf{F}_{x}^{0}\mathbf{G}_{x}\right)'\hat{\mathbf{F}}_{x}}{T} = O_{p}\left(\delta_{NT}^{-2}\right), \tag{B.4}$$

$$\frac{\left(\hat{\mathbf{F}}_{x} - \mathbf{F}_{x}^{0}\mathbf{G}_{x}\right)'\mathbf{F}_{x}^{0}}{T} = O_{p}\left(\delta_{NT}^{-2}\right), \tag{B.5}$$

$$\frac{\left(\hat{\mathbf{F}}_{x} - \mathbf{F}_{x}^{0}\mathbf{G}_{x}\right)'\mathbf{F}_{y}^{0}}{T} = O_{p}\left(\delta_{NT}^{-2}\right), \tag{B.6}$$

$$\frac{\left(\hat{\mathbf{F}}_{x} - \mathbf{F}_{x}^{0}\mathbf{G}_{x}\right)'\boldsymbol{\varepsilon}_{i}}{T} = O_{p}\left(\delta_{NT}^{-2}\right), \tag{B.7}$$

$$\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)' \mathbf{v}_{\ell i}}{T} = O_p\left(\delta_{NT}^{-2}\right), \tag{B.8}$$

$$\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}\right)' \mathbf{W}_i}{T} = O_p\left(\delta_{NT}^{-2}\right), \tag{B.9}$$

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)' \mathbf{v}_{\ell i}}{T} \boldsymbol{\gamma}_{\ell i}^{0\prime} = O\left(N^{-1/2}\right) + O_p\left(\delta_{NT}^{-2}\right), \tag{B.10}$$

$$\mathbf{G}_x \mathbf{G}_x' - \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T}\right)^{-1} = O_p\left(\delta_{NT}^{-2}\right), \tag{B.11}$$

$$\frac{\mathbf{F}_x^{0\prime} \hat{\mathbf{F}}_x}{T} \xrightarrow{p} \boldsymbol{\Lambda}_x \ as \ (N,T) \xrightarrow{j} \infty, \tag{B.12}$$

where we define $\mathbf{F}_x = \mathbf{F}_x^0 \mathbf{G}_x$ *and* $\boldsymbol{\Gamma}_{xi} = \mathbf{G}_x^{-1} \boldsymbol{\Gamma}_{xi}^0$*, where* $\mathbf{G}_x$ *and* $\boldsymbol{\Lambda}_x$ *are invertible* $m_x \times m_x$ *matrices.*

Proof of Lemma B.4. The proof of (B.3) is given in Bai (2009a, Proposition A.1). No modification is required because of our assumption of cross-sectional independence and serial correlation of $v_{it\ell}$, see Assumption 2. A similar point applies to the proofs of (B.4)-(B.11), which are given by Bai (2009a) as proofs of corresponding Lemmas A3(ii), A4(i), A4(ii), A3(iv), A4(iii) and A7(i). The result (B.12) is given as part of Proposition 1 in Bai (2003) with its proof therein. ■

Lemma B.5 *Under Assumptions 1-5, as* $(N,T) \xrightarrow{j} \infty$ *such that* $N/T \to c$ *with* $0 < c < \infty$*, for* $i = 1, 2, .., N$*,* $\ell = 1, 2, .., k$ *and* $r = 1, 2$

$$T^{-r/2} \left\| \hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y \right\|^r = T^{-r/2} \sum_{t=1}^{T} \left\| \hat{\mathbf{f}}_{y,t} - \mathbf{G}_y' \mathbf{f}_{y,t}^0 \right\|^r = O_p\left(\delta_{NT}^{-r}\right) + O_p\left(||\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}||^r\right), \tag{B.13}$$

$$\frac{\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y\right)' \hat{\mathbf{F}}_x}{T} = O_p\left(\delta_{NT}^{-2}\right) + O_p\left(\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}\right), \tag{B.14}$$

$$\frac{\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y\right)' \mathbf{F}_x^0}{T} = O_p\left(\delta_{NT}^{-2}\right) + O_p\left(\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}\right), \tag{B.15}$$

$$\frac{\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y\right)' \mathbf{F}_y^0}{T} = O_p\left(\delta_{NT}^{-2}\right) + O_p\left(\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}\right), \tag{B.16}$$

$$\frac{\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y\right)' \boldsymbol{\varepsilon}_i}{T} = O_p\left(\delta_{NT}^{-2}\right) + T^{-1/2} O_p\left(\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}\right), \tag{B.17}$$

$$\frac{\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y\right)' \mathbf{v}_{\ell i}}{T} = O_p\left(\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}\right) + O_p\left(\delta_{NT}^{-2}\right), \tag{B.18}$$

$$\frac{\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}\right)' \mathbf{W}_i}{T} = O_p\left(\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}\right) + O_p\left(\delta_{NT}^{-2}\right), \tag{B.19}$$

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y\right)' \boldsymbol{\varepsilon}_i}{T} \boldsymbol{\gamma}_{yi}^{0\prime} = N^{-1/2} O_p\left(\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}\right) + T^{-1/2} O_p\left(\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}\right) + O\left(N^{-1/2}\right) + O_p\left(\delta_{NT}^{-2}\right), \tag{B.20}$$

$$\mathbf{G}_y \mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime} \mathbf{F}_y^0}{T}\right)^{-1} = O_p\left(\hat{\boldsymbol{\theta}}_{IV} - \boldsymbol{\theta}\right) + O_p\left(\delta_{NT}^{-2}\right), \tag{B.21}$$

$$\frac{\mathbf{F}_y^{0\prime} \hat{\mathbf{F}}_y}{T} \xrightarrow{p} \boldsymbol{\Lambda}_y, \tag{B.22}$$

where we define $\mathbf{F}_y = \mathbf{F}_y^0 \mathbf{G}_y$ *and* $\boldsymbol{\gamma}_{yi} = \mathbf{G}_y^{-1} \boldsymbol{\gamma}_{yi}^0$*, with* $\mathbf{G}_y$ *being any invertible* $m_y \times m_y$ *matrix, whereas* $\boldsymbol{\Lambda}_y$ *is also an invertible* $m_y \times m_y$ *matrix.*

Proof of Lemma B.5. The proof is analogous to that of Lemma B.4 and is therefore omitted. ■

Lemma B.6 *Under Assumptions 1-5, as* $(N,T)\overset{j}{\to}\infty$ *such that* $N/T\to c$ *with* $0<c<\infty$, $\left\|\mathbf{P}_{\hat{F}_x}-\mathbf{P}_{F_x^0}\right\| = O_p(\delta_{NT}^{-1})$, $\left\|\mathbf{P}_{\hat{F}_{x,-1}}-\mathbf{P}_{F_{x,-1}^0}\right\| = O_p(\delta_{NT}^{-1})$ *and* $\left\|\mathbf{P}_{\hat{F}_y}-\mathbf{P}_{F_y^0}\right\| = o_p(1)$.

Proof of Lemma B.6. $\left\|\mathbf{P}_{\hat{F}_x}-\mathbf{P}_{F_x^0}\right\|^2 = tr\left[\left(\mathbf{P}_{\hat{F}_x}-\mathbf{P}_{F_x^0}\right)^2\right] = tr\left[\mathbf{P}_{\hat{F}_x}-\mathbf{P}_{\hat{F}_x}\mathbf{P}_{F_x^0}-\mathbf{P}_{F_x^0}\mathbf{P}_{\hat{F}_x}+\mathbf{P}_{F_x^0}\right] = tr\left[\mathbf{P}_{\hat{F}_x}\right]-2tr\left[\mathbf{P}_{\hat{F}_x}\mathbf{P}_{F_x^0}\right]+tr\left[\mathbf{P}_{F_x^0}\right] = 2m-2tr\left[T^{-1}\hat{\mathbf{F}}_x'\mathbf{P}_{F_x^0}\hat{\mathbf{F}}_x\right]$, where $T^{-1}\hat{\mathbf{F}}_x'\mathbf{P}_{F_x^0}\hat{\mathbf{F}}_x = T^{-1}\hat{\mathbf{F}}_x'\mathbf{F}_x^0\mathbf{G}_x\left(\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\mathbf{G}_x\right)^{-1}\mathbf{G}_x'\mathbf{F}_x^{0\prime}$ $T^{-2}\left(\hat{\mathbf{F}}_x'\mathbf{F}_x^0\mathbf{G}_x\right)\left(\mathbf{G}_x'\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x\right)$. By making use of (B.5), we have $T^{-1}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x = T^{-1}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\mathbf{G}_x + T^{-1}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right) =$ $\mathbf{I}_{m_x}+O_p\left(\delta_{NT}^{-2}\right)$. Hence, we have $\left\|\mathbf{P}_{\hat{F}_x}-\mathbf{P}_{F_x^0}\right\|^2 = 2m_x-2tr\left[\left(T^{-1}\hat{\mathbf{F}}_x'\mathbf{F}_x^0\mathbf{G}_x\right)\left(T^{-1}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x\right)\right] = 2m_x - 2tr\left[\mathbf{I}_{m_x}+O_p\left(\delta_{NT}^{-2}\right)\right] = O_p\left(\delta_{NT}^{-2}\right)$. Following similar arguments, it can be shown that $\left\|\mathbf{P}_{\hat{F}_{x,-1}}-\mathbf{P}_{F_{x,-1}^0}\right\|^2 = O_p\left(\delta_{NT}^{-2}\right)$. $\left\|\mathbf{P}_{\hat{F}_y}-\mathbf{P}_{F_y^0}\right\|^2 = o_p(1)$ is shown by the same steps as in the proof of Proposition 1 in Bai (2009a) and the result of Theorem 1. ■

Proof of Lemma 1. We begin with (A.1), which is given by

$$T^{-1}\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{Z}}_i - T^{-1}\mathbf{Z}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_i = \begin{bmatrix}\mathbf{Q}_{11} & \mathbf{Q}_{12}\\ \mathbf{Q}_{21} & \mathbf{Q}_{22}\end{bmatrix}, \tag{B.23}$$

where

$$\begin{aligned}
\mathbf{Q}_{11} &= T^{-1}\mathbf{X}_i'\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_y}\mathbf{M}_{\hat{F}_x}\mathbf{X}_i - T^{-1}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{M}_{F_y^0}\mathbf{M}_{F_x^0}\mathbf{X}_i,\\
\mathbf{Q}_{12} &= T^{-1}\mathbf{X}_i'\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_y}\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1} - T^{-1}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{M}_{F_y^0}\mathbf{M}_{F_{x,-1}^0}\mathbf{X}_{i,-1},\\
\mathbf{Q}_{21} &= T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_y}\mathbf{M}_{\hat{F}_x}\mathbf{X}_i - T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_y^0}\mathbf{M}_{F_x^0}\mathbf{X}_i,\\
\mathbf{Q}_{22} &= T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_y}\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{X}_{i,-1} - T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_y^0}\mathbf{M}_{F_{x,-1}^0}\mathbf{X}_{i,-1}.
\end{aligned}$$

Consider first $\mathbf{Q}_{11}$. By adding and subtracting the terms we have

$$\begin{aligned}
\|\mathbf{Q}_{11}\| &= \left\|T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_y}\mathbf{M}_{\hat{F}_x}\mathbf{X}_i - T^{-1}\mathbf{X}_{i,-1}'\mathbf{M}_{F_x^0}\mathbf{M}_{F_y^0}\mathbf{M}_{F_x^0}\mathbf{X}_i\right\|\\
&\leq \left\|T^{-1}\mathbf{X}_i'\left(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0}\right)\mathbf{M}_{\hat{F}_y}\mathbf{M}_{\hat{F}_x}\mathbf{X}_i\right\|\\
&+ \left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{M}_{\hat{F}_y}\left(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0}\right)\mathbf{X}_i\right\|\\
&+ \left\|T^{-1}\mathbf{X}_i'\mathbf{M}_{F_x^0}\left(\mathbf{M}_{\hat{F}_y}-\mathbf{M}_{F_y^0}\right)\mathbf{M}_{F_x^0}\mathbf{X}_i\right\|\\
&\leq \frac{\|\mathbf{X}_i\|}{\sqrt{T}}\left\|\mathbf{P}_{\hat{F}_x}-\mathbf{P}_{F_x^0}\right\|\frac{\left\|\mathbf{M}_{\hat{F}_y}\mathbf{M}_{\hat{F}_x}\mathbf{X}_i\right\|}{\sqrt{T}}\\
&+ \frac{\left\|\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{M}_{\hat{F}_y}\right\|}{\sqrt{T}}\left\|\mathbf{P}_{\hat{F}_x}-\mathbf{P}_{F_x^0}\right\|\frac{\|\mathbf{X}_i\|}{\sqrt{T}}\\
&+ \frac{\left\|\mathbf{X}_i'\mathbf{M}_{F_x^0}\right\|}{\sqrt{T}}\left\|\mathbf{P}_{\hat{F}_y}-\mathbf{P}_{F_y^0}\right\|\frac{\left\|\mathbf{M}_{F_x^0}\mathbf{X}_i\right\|}{\sqrt{T}} = O_p(\delta_{NT}^{-1}),
\end{aligned}$$

from using again $\left\|\mathbf{P}_{\hat{F}_x}-\mathbf{P}_{F_x^0}\right\| = O_p(\delta_{NT}^{-1})$ and $\left\|\mathbf{P}_{\hat{F}_y}-\mathbf{P}_{F_y^0}\right\| = O_p(\delta_{NT}^{-1})$ by Lemma B.6, $\frac{\|\mathbf{X}_i\|}{\sqrt{T}} = O_p(1)$

by Assumptions 2, 3, 4 as discussed above, and also

$$\frac{\left\|\mathbf{M}_{\hat{F}_x}\mathbf{X}_i\right\|}{\sqrt{T}} \leq \frac{\|\mathbf{X}_i\|}{\sqrt{T}} + \left\|\mathbf{P}_{\hat{F}_x}\right\| \frac{\|\mathbf{X}_i\|}{\sqrt{T}} = O_p(1), \tag{B.24}$$

$$\frac{\|\mathbf{M}_{F_x^0}\mathbf{X}_i\|}{\sqrt{T}} = \frac{\|\mathbf{M}_{F_x^0}\mathbf{V}_i\|}{\sqrt{T}} \leq \frac{\|\mathbf{V}_i\|}{\sqrt{T}} + \|\mathbf{P}_{F_x^0}\| \frac{\|\mathbf{V}_i\|}{\sqrt{T}} = O_p(1), \tag{B.25}$$

$$\frac{\left\|\mathbf{M}_{\hat{F}_y}\mathbf{M}_{\hat{F}_x}\mathbf{X}_i\right\|}{\sqrt{T}} \leq \frac{\left\|\mathbf{M}_{\hat{F}_x}\mathbf{X}_i\right\|}{\sqrt{T}} + \left\|\mathbf{P}_{\hat{F}_y}\right\| \frac{\left\|\mathbf{M}_{\hat{F}_x}\mathbf{X}_i\right\|}{\sqrt{T}} = O_p(1), \tag{B.26}$$

$$\frac{\left\|\mathbf{M}_{\hat{F}_y}\mathbf{M}_{F_x^0}\mathbf{X}_i\right\|}{\sqrt{T}} \leq \frac{\|\mathbf{M}_{F_x^0}\mathbf{X}_i\|}{\sqrt{T}} + \left\|\mathbf{P}_{\hat{F}_y}\right\| \frac{\|\mathbf{M}_{F_x^0}\mathbf{X}_i\|}{\sqrt{T}} = O_p(1), \tag{B.27}$$

since $\left\|\mathbf{P}_{\hat{F}_y}\right\| = \left\|\hat{\mathbf{F}}_y(\hat{\mathbf{F}}_y'\hat{\mathbf{F}}_y)^{-1}\hat{\mathbf{F}}_y'\right\| = \left\|\hat{\mathbf{F}}_y'\hat{\mathbf{F}}_y(\hat{\mathbf{F}}_y'\hat{\mathbf{F}}_y)^{-1}\right\| = \|\mathbf{I}_{m_y}\| = \text{tr}\left(\mathbf{I}_{m_y}\right) = m_y$. Hence, $\mathbf{Q}_{11} = O_p(\delta_{NT}^{-1})$.

By similar arguments, it can be shown that $\mathbf{Q}_{21} = O_p(\delta_{NT}^{-1})$, $\mathbf{Q}_{12} = O_p(\delta_{NT}^{-1})$ and $\mathbf{Q}_{22} = O_p(\delta_{NT}^{-1})$, which complete the proof of the result in (A.1). The result in (A.2) can be shown in an analogous way. ■

Proof of Lemma 2. The proof is analogous to that of Lemma 1 as the summation over i does not affect the results. It is therefore omitted. ■

Proof of Lemma 3. We start with (A.5). First note that by using $\mathbf{M}_{\hat{F}_x} = \mathbf{I}_T - T^{-1}\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x'$ the left-hand-side of (A.5) can be written as

$$\begin{aligned}
&\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x'\left(\mathbf{\Sigma}_{kNT}-\bar{\mathbf{\Sigma}}_{kNT}\right)\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&= \frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x'\left(\mathbf{\Sigma}_{kNT}-\bar{\mathbf{\Sigma}}_{kNT}\right)\mathbf{u}_i \\
&- \frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x'\left(\mathbf{\Sigma}_{kNT}-\bar{\mathbf{\Sigma}}_{kNT}\right)\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x'}{T}\mathbf{u}_i \\
&= \mathbf{e}_1 + \mathbf{e}_2.
\end{aligned}$$

$$\begin{aligned}
\mathbf{e}_1 &= \frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\left(\mathbf{\Sigma}_{kNT}-\bar{\mathbf{\Sigma}}_{kNT}\right)\mathbf{u}_i \\
&+ \frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\left(\mathbf{\Sigma}_{kNT}-\bar{\mathbf{\Sigma}}_{kNT}\right)\mathbf{u}_i \\
&= \mathbf{a}_1 + \mathbf{a}_2.
\end{aligned}$$

$$\begin{aligned}
|\mathbf{a}_1| &\leq \frac{1}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0\prime}\right\|\left\|\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\right\|\left\|\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\right\|\|\mathbf{G}_x\|\left\|\frac{\mathbf{F}_x^{0\prime}\sqrt{N}\left(\mathbf{\Sigma}_{kNT}-\bar{\mathbf{\Sigma}}_{kNT}\right)\mathbf{u}_i}{T}\right\| \\
&= O_p\left(T^{-1/2}\right),
\end{aligned}$$

because $\left\|\mathbf{\Gamma}_{xi}^{0\prime}\right\| = O_p(1)$ and $\left\|\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\right\| = O_p(1)$ by Assumption 4, $\left\|\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\right\| = O_p(1)$, $\|\mathbf{G}_x\| = O_p(1)$, and $\left\|\frac{\mathbf{F}_x^{0\prime}\sqrt{N}\left(\mathbf{\Sigma}_{kNT}-\bar{\mathbf{\Sigma}}_{kNT}\right)\mathbf{u}_i}{T}\right\| = O_p(1)$ by (B.1). Thus, we have $\mathbf{a}_1 = O_p\left(T^{-1/2}\right)$.

As for $\mathbf{a}_2$, we have the following

$$\begin{aligned}
|\mathbf{a}_2| &\leq \frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0\prime}\right\|\left\|\left(\mathbf{\Upsilon}_{xkN}^{0}\right)^{-1}\right\|\left\|\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\sqrt{N}\left(\mathbf{\Sigma}_{kNT}-\bar{\mathbf{\Sigma}}_{kNT}\right)\mathbf{u}_i}{T\sqrt{T}}\right\| \\
&= O_p\left(\delta_{NT}^{-1}\right),
\end{aligned}$$

by using Assumption 4, $\left\|\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\right\| = O_p(1)$ and

$$\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\sqrt{N}\left(\boldsymbol{\Sigma}_{kNT} - \bar{\boldsymbol{\Sigma}}_{kNT}\right)\mathbf{u}_i}{T\sqrt{T}}\right\|$$
$$\leq \left\|\frac{1}{T}\sum_{t=1}^{T}\left(\hat{\mathbf{f}}_{x,t} - \mathbf{G}_x^{*\prime}\mathbf{f}_{x,t}^0\right)\frac{1}{\sqrt{NT}}\sum_{l=1}^{k}\sum_{i=1}^{N}\sum_{s=1}^{T}u_{is}\left[v_{\ell it}v_{\ell is} - E\left(v_{\ell it}v_{\ell is}\right)\right]\right\|$$
$$\leq \left(\frac{1}{T}\sum_{t=1}^{T}\left\|\hat{\mathbf{f}}_{x,t} - \mathbf{G}_x^{*\prime}\mathbf{f}_{x,t}^0\right\|^2\right)^{1/2}\left\{\frac{1}{T}\sum_{t=1}^{T}\left\|\frac{1}{\sqrt{NT}}\sum_{l=1}^{k}\sum_{i=1}^{N}\sum_{s=1}^{T}u_{is}\left[v_{\ell it}v_{\ell is} - E\left(v_{\ell it}v_{\ell is}\right)\right]\right\|^2\right\}^{1/2}$$
$$= O_p\left(\delta_{NT}^{-1}\right),$$

where the second inequality is derived using the Cauchy-Schwarz inequality and the order is determined by using $\left\|\frac{1}{\sqrt{NT}}\sum_{l=1}^{k}\sum_{i=1}^{N}\sum_{s=1}^{T}u_{is}\left[v_{\ell it}v_{\ell is} - E\left(v_{\ell it}v_{\ell is}\right)\right]\right\| = O_p(1)$ and (B.3). Thus, $\mathbf{e}_1 = O_p\left(T^{-1/2}\right) + O_p\left(\delta_{NT}^{-1}\right)$. Next, consider $\mathbf{e}_2$ which is

$$|\mathbf{e}_2| \leq \sqrt{\frac{N}{T}}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\|\right)\left\|\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\right\|\left\|\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\hat{\mathbf{F}}_x\left(\boldsymbol{\Sigma}_{kNT} - \bar{\boldsymbol{\Sigma}}_{kNT}\right)\hat{\mathbf{F}}_x}{T}\right\|\left\|\frac{\hat{\mathbf{F}}_x'}{\sqrt{T}}\right\|$$
$$= \sqrt{\frac{N}{T}}O_p(1)\left\|\frac{\hat{\mathbf{F}}_x\left(\boldsymbol{\Sigma}_{kNT} - \bar{\boldsymbol{\Sigma}}_{kNT}\right)\hat{\mathbf{F}}_x}{T}\right\|,$$

because $\left\|T^{-1/2}\hat{\mathbf{F}}_x\right\| = \sqrt{\mathrm{tr}\left[T^{-1}\hat{\mathbf{F}}_x'\hat{\mathbf{F}}_x\right]} = \sqrt{\mathrm{tr}\left[\mathbf{I}_{m_x}\right]} = \sqrt{m_x} = O(1)$, $\left\|\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\right\| = O_p(1)$, $\left\|\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\right\| = O_p(1)$, and

$$\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\| \leq \frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\right\|\left\|\boldsymbol{\gamma}_i^0\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\| + \frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\right\|\left\|\frac{\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| = O_p(1),$$

by the same arguments as above and Assumptions 1, 3, 4. We also have

$$\left\|\sqrt{\frac{N}{T}}\frac{1}{T}\hat{\mathbf{F}}_{x,-1}'\left(\boldsymbol{\Sigma}_{kNT} - \bar{\boldsymbol{\Sigma}}_{kNT}\right)\hat{\mathbf{F}}_x\right\|$$
$$\leq \left\|\sqrt{\frac{N}{T}}\frac{1}{T}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\left(\boldsymbol{\Sigma}_{kNT} - \bar{\boldsymbol{\Sigma}}_{kNT}\right)\mathbf{F}_x^0\mathbf{G}_x\right\|$$
$$+ \left\|\sqrt{\frac{N}{T}}\frac{1}{T}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\left(\boldsymbol{\Sigma}_{kNT} - \bar{\boldsymbol{\Sigma}}_{kNT}\right)\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)\right\|$$
$$+ \left\|\sqrt{\frac{N}{T}}\frac{1}{T}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\left(\boldsymbol{\Sigma}_{kNT} - \bar{\boldsymbol{\Sigma}}_{kNT}\right)\mathbf{F}_x^0\mathbf{G}_x\right\|$$
$$+ \left\|\sqrt{\frac{N}{T}}\frac{1}{T}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\left(\boldsymbol{\Sigma}_{kNT} - \bar{\boldsymbol{\Sigma}}_{kNT}\right)\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)\right\|$$
$$= \|\mathbf{L}_1\| + \|\mathbf{L}_2\| + \|\mathbf{L}_3\| + \|\mathbf{L}_4\|.$$

$$\|\mathbf{L}_1\| \leq \sqrt{\frac{1}{T}}\left\|\mathbf{G}_x'\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\sqrt{N}\left(\boldsymbol{\Sigma}_{kNT} - \bar{\boldsymbol{\Sigma}}_{kNT}\right)\mathbf{F}_x^0}{T}\right\|\left\|\mathbf{G}_x\right\| = O_p\left(T^{-1/2}\right),$$

by (B.2), and

$$\|\mathbf{L}_2\| = \|\mathbf{L}_3\| \leq \left\|\mathbf{G}_x\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\sqrt{N}\left(\boldsymbol{\Sigma}_{kNT} - \bar{\boldsymbol{\Sigma}}_{kNT}\right)\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T\sqrt{T}}\right\| = O_p\left(\delta_{NT}^{-1}\right),$$

because $\|\mathbf{G}_x\| = O_p(1)$ and

$$\left\| \frac{\mathbf{F}_x^{0\prime}\sqrt{N}\left(\mathbf{\Sigma}_{kNT} - \bar{\mathbf{\Sigma}}_{kNT}\right)\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T\sqrt{T}} \right\|$$
$$\leq \left\| \frac{1}{T}\sum_{s=1}^{T}\left\{ \frac{1}{\sqrt{NT}}\sum_{l=1}^{k}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{f}_{x,t}^0\left[v_{\ell it}v_{\ell is} - E\left(v_{\ell it}v_{\ell is}\right)\right]\right\}\left(\hat{\mathbf{f}}_{x,s} - \mathbf{G}_x^{*\prime}\mathbf{f}_{x,s}^0\right)^{\prime} \right\|$$
$$\leq \left(\frac{1}{T}\sum_{s=1}^{T}\left\|\hat{\mathbf{f}}_{x,s} - \mathbf{G}_x^{*\prime}\mathbf{f}_{x,s}^0\right\|^2\right)^{1/2}\left\{\frac{1}{T}\sum_{s=1}^{T}\left\|\frac{1}{\sqrt{NT}}\sum_{l=1}^{k}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{f}_{x,t}^0\left[v_{\ell it}v_{\ell is} - E\left(v_{\ell it}v_{\ell is}\right)\right]\right\|^2\right\}^{1/2}$$
$$= O_p\left(\delta_{NT}^{-1}\right),$$

where the second inequality is derived using the Cauchy-Schwarz inequality and the order is determined by using $\left\|\frac{1}{\sqrt{NT}}\sum_{l=1}^{k}\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{f}_{x,t}^0\left[v_{\ell it}v_{\ell is} - E\left(v_{\ell it}v_{\ell is}\right)\right]\right\| = O_p(1)$ and (B.3). Thus, $\mathbf{e}_2 = O_p\left(T^{-1/2}\right) + O_p\left(\delta_{NT}^{-1}\right)$.

$$\|\mathbf{L}_4\| = \left\|\sqrt{\frac{1}{T}}\frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left(\hat{\mathbf{f}}_{x,t} - \mathbf{G}_x^{*\prime}\mathbf{f}_{x,t}^0\right)\left(\hat{\mathbf{f}}_{x,s} - \mathbf{G}_x^{\prime}\mathbf{f}_{x,s}^0\right)^{\prime}\frac{1}{\sqrt{N}}\sum_{l=1}^{k}\sum_{i=1}^{N}\left[v_{\ell it}v_{\ell is} - E\left(v_{\ell it}v_{\ell is}\right)\right]\right\|,$$

so that, by the Cauchy-Schwarz inequality we have

$$\|\mathbf{L}_4\| \leq \sqrt{T}\left(\frac{1}{T}\left\|\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right\|^2\right)\left\{\frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\left[v_{\ell it}v_{\ell is} - E\left(v_{\ell it}v_{\ell is}\right)\right]\right]^2\right\}^{1/2}$$
$$= \sqrt{T}O_p\left(\delta_{NT}^{-2}\right).$$

Thus, $\mathbf{e}_2 = O_p\left(T^{-1/2}\right) + \sqrt{T}O_p\left(\delta_{NT}^{-2}\right)$. Collecting all the results, the required expression is obtained. The result in (A.6) is proved in a similar way. ■

Proof of Lemma 4. We begin with (A.7). Following the discussion in Bai (2009a, p.1266), we have

$$\hat{\mathbf{F}}_x\mathbf{G}_x^{-1} - \mathbf{F}_x^0 = \left(\mathbf{E}_{x1} + \mathbf{E}_{x2} + \mathbf{E}_{x3}\right)\hat{\mathbf{Q}}_x$$
$$= \frac{1}{NT}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{F}_x^0\boldsymbol{\gamma}_{\ell i}^0\mathbf{v}_{\ell i}^{\prime}\hat{\mathbf{F}}_x\hat{\mathbf{Q}}_x + \frac{1}{NT}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{v}_{\ell i}\boldsymbol{\gamma}_{\ell i}^{0\prime}\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x\hat{\mathbf{Q}}_x$$
$$+ \frac{1}{NT}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{v}_{\ell i}\mathbf{v}_{\ell i}^{\prime}\hat{\mathbf{F}}_x\hat{\mathbf{Q}}_x, \tag{B.28}$$

where $\hat{\mathbf{Q}}_x = \left(\mathbf{\Upsilon}_{xkN}^0\mathbf{\Lambda}_{0\hat{F}_x}\right)^{-1}$ with $\mathbf{\Upsilon}_{xkN}^0 = \frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\boldsymbol{\gamma}_{\ell i}^0\boldsymbol{\gamma}_{\ell i}^{0\prime}$ and $\mathbf{\Lambda}_{0\hat{F}_x} = T^{-1}\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x$.
Then, from (B.28) we have

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\mathbf{F}_x^{0\prime}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i$$
$$= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left[\mathbf{G}_x^{-1\prime}\hat{\mathbf{F}}_x^{\prime} - \mathbf{F}_x^{0\prime}\right]\mathbf{M}_{\hat{F}_x}\mathbf{u}_i$$
$$= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\hat{\mathbf{F}}_x^{\prime}\mathbf{v}_{\ell j}\boldsymbol{\gamma}_{\ell j}^{0\prime}\mathbf{F}_x^{0\prime}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i$$
$$- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\hat{\mathbf{F}}_x^{\prime}\mathbf{F}_x^0\boldsymbol{\gamma}_{\ell j}^0\mathbf{v}_{\ell j}^{\prime}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i$$
$$- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\hat{\mathbf{F}}_x^{\prime}\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}^{\prime}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i$$
$$= -\left(\mathbf{d}_1 + \mathbf{d}_2 + \mathbf{d}_3\right).$$

Start with $\mathbf{d}_1$, which is given by

$$\frac{\mathbf{d}_1}{\sqrt{NT}} = \frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\mathbf{A}_{kNT}\left[\mathbf{G}_x^{-1\prime}\hat{\mathbf{F}}_x^{\prime} - \mathbf{F}_x^{0\prime}\right]\mathbf{M}_{\hat{F}_x}\mathbf{u}_i$$

which is a $k \times 1$ vector, where

$$\mathbf{A}_{kNT} = \frac{1}{N}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\frac{\hat{\mathbf{F}}_x^{\prime}\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime}.$$

We have

$$\begin{aligned}\frac{1}{N}\sum_{j=1}^{N}\frac{\hat{\mathbf{F}}_x^{\prime}\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime} &= \frac{1}{N}\sum_{j=1}^{N}\mathbf{G}_x^{\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime} + \frac{1}{N}\sum_{j=1}^{N}\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)^{\prime}\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime} \\ &= O_p\left(T^{-1/2}N^{-1/2}\right) + O_p\left(N^{-1}\right) + N^{-1/2}O_p\left(\delta_{NT}^{-2}\right),\end{aligned}$$

as the first term is $O_p\left(T^{-1/2}N^{-1/2}\right)$ by independence of $\mathbf{v}_{\ell j}$ and $\gamma_{\ell j}^0$ and the second term is $O_p\left(N^{-1}\right) + N^{-1/2}O_p\left(\delta_{NT}^{-2}\right)$ by (B.10) in Lemma B.4. This gives the following

$$\|\mathbf{A}_{kNT}\| = O_p\left(T^{-1/2}N^{-1/2}\right) + O_p\left(N^{-1}\right) + N^{-1/2}O_p\left(\delta_{NT}^{-2}\right).$$

Next,

$$\begin{aligned}\frac{|\mathbf{d}_1|}{\sqrt{NT}} &\leq \left\|\frac{1}{NT}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\mathbf{A}_{kNT}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)^{\prime}\mathbf{u}_i\right\| \\ &+ \left\|\frac{1}{NT^2}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\mathbf{A}_{kNT}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)^{\prime}\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x^{\prime}\mathbf{u}_i\right\| \\ &\leq \left(\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0\prime}\right\|\left\|\gamma_i^0\right\|\right)\left\|\hat{\mathbf{Q}}_x^{\prime}\right\|\left\|\mathbf{A}_{kNT}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)^{\prime}\mathbf{F}_y^0}{T}\right\| \\ &+ \left(\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0\prime}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)^{\prime}\varepsilon_i}{T}\right\|\right)\left\|\hat{\mathbf{Q}}_x^{\prime}\right\|\left\|\mathbf{A}_{kNT}\right\| \\ &+ \left(\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0\prime}\right\|\left\|\gamma_i^0\right\|\right)\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)^{\prime}\hat{\mathbf{F}}_x}{T}\right\|\left\|\frac{\hat{\mathbf{F}}_x}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\hat{\mathbf{Q}}_x^{\prime}\right\|\left\|\mathbf{A}_{kNT}\right\| \\ &+ \left(\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0\prime}\right\|\left\|\frac{\varepsilon_i}{\sqrt{T}}\right\|\right)\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)^{\prime}\hat{\mathbf{F}}_x}{T}\right\|\left\|\frac{\hat{\mathbf{F}}_x}{\sqrt{T}}\right\|\left\|\hat{\mathbf{Q}}_x^{\prime}\right\|\left\|\mathbf{A}_{kNT}\right\| \\ &= O_p\left(\delta_{NT}^{-2}\right)\left[O_p\left(T^{-1/2}N^{-1/2}\right) + O_p\left(N^{-1}\right) + N^{-1/2}O_p\left(\delta_{NT}^{-2}\right)\right],\end{aligned}$$

by (B.4), (B.6), (B.7), Assumptions 1, 3, 4 and $\|\mathbf{A}_{kNT}\| = O_p\left(T^{-1/2}N^{-1/2}\right) + O_p\left(N^{-1}\right) + N^{-1/2}O_p\left(\delta_{NT}^{-2}\right)$ as shown above. We therefore have

$$\mathbf{d}_1 = \sqrt{NT}O_p\left(\delta_{NT}^{-2}\right) \times \left[O_p\left(T^{-1/2}N^{-1/2}\right) + O_p\left(N^{-1}\right) + N^{-1/2}O_p\left(\delta_{NT}^{-2}\right)\right].$$

Now consider $\mathbf{d}_2$ which can be written as

$$\begin{aligned}\mathbf{d}_2 &= \frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell j}^0\mathbf{v}_{\ell j}^{\prime}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\ &= \frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^0\right)^{-1}\left(\sum_{\ell=1}^{k}\gamma_{\ell j}^0\mathbf{v}_{\ell j}^{\prime}\right)\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\ &= \frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{\Upsilon}_{xkN}^0\right)^{-1}\mathbf{\Gamma}_{xj}^0\mathbf{V}_j^{\prime}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i,\end{aligned}$$

Consider now $\mathbf{d}_3$. Defining $\mathbf{\Sigma}_{kNT} = N^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'$, we have

$$\begin{aligned}
\mathbf{d}_3 &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x'\hat{\mathbf{F}}_x'\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\frac{1}{T}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x'\hat{\mathbf{F}}_x'\left(\frac{1}{N}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'\right)\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&= \frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x'\hat{\mathbf{F}}_x'\mathbf{\Sigma}_{kNT}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&= \frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}(\mathbf{\Upsilon}_{xkN}^{0})^{-1}\mathbf{\Lambda}_{0\hat{F}_x}^{-1\prime}\hat{\mathbf{F}}_x'\mathbf{\Sigma}_{kNT}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&= \frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}(\mathbf{\Upsilon}_{xkN}^{0})^{-1}\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x'\mathbf{\Sigma}_{kNT}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i.
\end{aligned}$$

where the definitions of $\hat{\mathbf{Q}}_x$ and $\mathbf{\Lambda}_{0\hat{F}_x}$ are given above. Hence, the expressions for $\mathbf{d}_1$, $\mathbf{d}_2$ and $\mathbf{d}_3$ gives the required result in (A.7). The result in (A.8) is obtained in an analogous manner. ■

Proof of Lemma 5. First of all consider (A.9). By adding and subtracting terms and using $\mathbf{M}_{F_y^0}\mathbf{F}_y^0 = \mathbf{0}$, we get

$$\begin{aligned}
\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\mathbf{F}_x^{0\prime}\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_y}\mathbf{u}_i &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\mathbf{F}_x^{0\prime}\mathbf{M}_{\hat{F}_x}\mathbf{M}_{F_y^0}\mathbf{u}_i \\
&+ \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\mathbf{F}_x^{0\prime}\mathbf{M}_{\hat{F}_x}(\mathbf{M}_{\hat{F}_y} - \mathbf{M}_{F_y^0})\mathbf{u}_i, \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\mathbf{F}_x^{0\prime}(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0})\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\mathbf{F}_x^{0\prime}(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0})(\mathbf{M}_{\hat{F}_y} - \mathbf{M}_{F_y^0})\mathbf{u}_i, \qquad \text{(B.29)}
\end{aligned}$$

where the second equality is due to $\mathbf{M}_{F_x^0}\mathbf{F}_x^0 = \mathbf{0}$ and $\mathbf{M}_{F_y^0}\mathbf{u}_i = \mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i$. Let us now begin with the first term in (B.29). Since $\hat{\mathbf{F}}_x'\hat{\mathbf{F}}_x/T = \mathbf{I}_{m_x}$, we have $\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0} = \mathbf{P}_{F_x^0} - \mathbf{P}_{\hat{F}_x} = -\left(\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x'}{T} - \mathbf{P}_{F_x^0}\right)$. Using this result and by adding and subtracting terms, we get

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\mathbf{F}_x^{0\prime}(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0})\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\mathbf{F}_x^{0\prime}\left(\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x'}{T} - \mathbf{P}_{F_x^0}\right)\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\mathbf{G}_x\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\left[\mathbf{G}_x\mathbf{G}_x' - \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= -\left(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4\right).
\end{aligned}$$

$$
\begin{aligned}
|\mathbf{e}_1| &\leq \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_{xi}^{0\prime} \frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \mathbf{G}_x^{\prime} \mathbf{F}_x^{0\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \right\| \\
&\leq \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_{xi}^{0} \right\| \left\| \frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| \left\| \mathbf{G}_x \right\| \left\| \mathbf{F}_x^{0\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \right\| \\
&\leq \sqrt{N} \left(\frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_{xi}^{0} \right\| \left\| \frac{\mathbf{F}_x^{0\prime} \boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \right) \left\| \frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| \left\| \mathbf{G}_x \right\| \\
&+ \sqrt{N} \left(\frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_{xi}^{0} \right\| \left\| \frac{\mathbf{F}_y^{0\prime} \boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \right) \left\| \frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| \left\| \mathbf{G}_x \right\| \left\| \frac{\mathbf{F}_x^0}{\sqrt{T}} \right\| \left\| \frac{\mathbf{F}_y^0}{\sqrt{T}} \right\| \left\| \left(\frac{\mathbf{F}_y^{0\prime} \mathbf{F}_y^0}{T} \right)^{-1} \right\| \\
&= \sqrt{N} O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

by Assumptions 1, 3, 4 and (B.5).

$$
\begin{aligned}
|\mathbf{e}_2| &\leq \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_{xi}^{0\prime} \frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \right\| \\
&\leq \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_{xi}^{0\prime} \right\| \left\| \frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| \left\| \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \right\| \\
&\leq \sqrt{NT} \left(\frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_{xi}^{0\prime} \right\| \left\| \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)^{\prime} \boldsymbol{\varepsilon}_i}{T} \right\| \right) \left\| \frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| \\
&+ \sqrt{N} \left(\frac{1}{N} \sum_{i=1}^{N} \left\| \boldsymbol{\Gamma}_{xi}^{0} \right\| \left\| \frac{\mathbf{F}_y^{0\prime} \boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \right) \left\| \frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| \left\| \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)^{\prime} \mathbf{F}_y^0}{T} \right\| \left\| \left(\frac{\mathbf{F}_y^{0\prime} \mathbf{F}_y^0}{T} \right)^{-1} \right\| \\
&= \sqrt{NT} O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

again, by Assumptions 1, 3, 4, (B.5) and (B.6).

$$
\begin{aligned}
\mathbf{e}_3 &= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \mathbf{G}_x \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \mathbf{G}_x \mathbf{G}_x^{\prime} \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \left[\mathbf{G}_x \mathbf{G}_x^{\prime} - \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \right] \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \boldsymbol{\Gamma}_{xi}^{0\prime} \frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \left[\mathbf{G}_x \mathbf{G}_x^{\prime} - \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \right] \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \\
&= \mathbf{a}_1 + \mathbf{a}_2.
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{e}_4| &\leq \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\left[\mathbf{G}_x\mathbf{G}_x^{\prime}-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i\right\| \\
&\leq \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right\|\left\|\mathbf{G}_x\mathbf{G}_x^{\prime}-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&\leq \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|\right)\left\|\frac{\mathbf{F}_x^{0}}{\sqrt{T}}\right\|^2\left\|\mathbf{G}_x\mathbf{G}_x^{\prime}-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\| \\
&+ \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|\right)\left\|\frac{\mathbf{F}_x^{0}}{\sqrt{T}}\right\|^3\left\|\mathbf{G}_x\mathbf{G}_x^{\prime}-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\| \\
&= \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

by Assumptions 1, 3, 4, and (B.11). Now consider $\mathbf{a}_1$ and $\mathbf{a}_2$ in $\mathbf{e}_3$. We begin with $\mathbf{a}_2$ which is given by

$$
\begin{aligned}
|\mathbf{a}_2| &\leq \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0\prime}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\left[\mathbf{G}_x\mathbf{G}_x^{\prime}-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)^{\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i\right\| \\
&\leq \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right\|\left\|\mathbf{G}_x\mathbf{G}_x^{\prime}-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)^{\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i\right\| \\
&\leq \sqrt{NT}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)^{\prime}\boldsymbol{\varepsilon}_i}{T}\right\|\right)\left\|\frac{\mathbf{F}_x^{0}}{\sqrt{T}}\right\|^2\left\|\mathbf{G}_x\mathbf{G}_x^{\prime}-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\| \\
&+ \sqrt{NT}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{\Gamma}_{xi}^{0}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|\right)\left\|\frac{\mathbf{F}_x^{0}}{\sqrt{T}}\right\|^2\left\|\mathbf{G}_x\mathbf{G}_x^{\prime}-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)^{\prime}\mathbf{F}_y^0}{T}\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

by the same arguments as above and (B.6). As for $\mathbf{a}_1$, we have

$$
\begin{aligned}
\mathbf{a}_1 &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)^{\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\left(\mathbf{E}_{x1}^{\prime}+\mathbf{E}_{x2}^{\prime}+\mathbf{E}_{x3}^{\prime}\right)\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\frac{\hat{\mathbf{F}}_x^{\prime}\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T} \\
&+ \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\frac{\hat{\mathbf{F}}_x^{\prime}\mathbf{F}_x^0}{T}\gamma_{\ell j}^{0}\frac{\mathbf{v}_{\ell j}^{\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T} \\
&+ \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\frac{\hat{\mathbf{F}}_x^{\prime}\mathbf{v}_{\ell j}}{T}\frac{\mathbf{v}_{\ell j}^{\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T} \\
&= \mathbf{c}_1+\mathbf{c}_2+\mathbf{c}_3.
\end{aligned}
$$

$$\begin{aligned}
\mathbf{c}_1 &= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x'\frac{\hat{\mathbf{F}}_x'\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x'\mathbf{G}_x'\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T} \\
&+ \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x'\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T} \\
&= \mathbf{d}_1+\mathbf{d}_2
\end{aligned}$$

$$\begin{aligned}
|\mathbf{d}_1| &\le \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x'\mathbf{G}_x'\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T}\right\| \\
&\le T^{-1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|\right)\left\|\hat{\mathbf{Q}}_x\right\|\left\|\mathbf{G}_x\right\|\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell j}}{\sqrt{T}}\gamma_{\ell j}^{0\prime}\right\| \\
&+ T^{-1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|\right)\left\|\hat{\mathbf{Q}}_x\right\|\left\|\mathbf{G}_x\right\|\left\|\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell j}}{\sqrt{T}}\gamma_{\ell j}^{0\prime}\right\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\| \\
&= T^{-1/2}O_p(1)=O_p\left(T^{-1/2}\right).
\end{aligned}$$

and

$$\begin{aligned}
|\mathbf{d}_2| &\le \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x'\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T}\right\| \\
&\le T^{-1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|\right)\left\|\hat{\mathbf{Q}}_x\right\|\left(\sum_{\ell=1}^{k}\left\|\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{v}_{\ell j}}{\sqrt{T}}\gamma_{\ell j}^{0\prime}\right\|\right) \\
&+ T^{-1/2}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|\right)\left\|\hat{\mathbf{Q}}_x\right\|\left(\sum_{\ell=1}^{k}\left\|\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{v}_{\ell j}}{\sqrt{T}}\gamma_{\ell j}^{0\prime}\right\|\right) \\
&\times \left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\| \\
&= T^{-1/2}\left[O\left(N^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)\right]=O_p\left(N^{-1/2}T^{-1/2}\right)+T^{-1/2}O_p\left(\delta_{NT}^{-2}\right).
\end{aligned}$$

due to Assumptions 1, 3, 4 and (B.10).

$$\begin{aligned}
\mathbf{c}_2 &= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x'\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\gamma_{\ell j}^{0}\frac{\mathbf{v}_{\ell j}'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T} \\
&= \sqrt{\frac{1}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell j}^{0}\frac{\mathbf{v}_{\ell j}'\boldsymbol{\varepsilon}_i}{\sqrt{T}} \\
&- \sqrt{\frac{1}{NT}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell j}^{0}\frac{\mathbf{v}_{\ell j}'\mathbf{F}_y}{\sqrt{T}}\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\frac{\mathbf{F}_y'\boldsymbol{\varepsilon}_i}{\sqrt{T}} \\
&= \mathbf{d}_1+\mathbf{d}_2.
\end{aligned}$$

Now consider $\mathbf{d}_1$, which is

$$
\begin{aligned}
\mathbf{d}_1 &= \sqrt{\frac{1}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\gamma_{\ell j}^{0}\frac{\mathbf{v}_{\ell j}^{\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \\
&= \sqrt{\frac{1}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\gamma_{\ell i}^{0}\frac{\mathbf{v}_{\ell i}^{\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \\
&+ \sqrt{\frac{1}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\gamma_{\ell j}^{0}\frac{\mathbf{v}_{\ell j}^{\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \\
&= \mathbf{e}_1 + \mathbf{e}_2.
\end{aligned}
$$

As for $\mathbf{e}_1$, we have

$$
\begin{aligned}
|\mathbf{e}_1| &= \left\|\sqrt{\frac{1}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\gamma_{\ell i}^{0}\frac{\mathbf{v}_{\ell i}^{\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&\leq \sqrt{\frac{1}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\right\|\left\|\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\|\left\|\gamma_{\ell i}^{0}\right\|\left\|\frac{\mathbf{v}_{\ell i}^{\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= O_p(N^{-1/2}),
\end{aligned}
$$

by Assumptions 1, 2 and 4. Next is $\mathbf{e}_2$, which is

$$
\begin{aligned}
\mathbf{e}_2 &= \sqrt{\frac{1}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\gamma_{\ell j}^{0}\frac{\mathbf{v}_{\ell j}^{\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \\
&= \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{j\neq i}^{N}\sum_{t=1}^{T}\gamma_{\ell j}^{0}v_{\ell jt}\varepsilon_{it}\right) \\
&= \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\mathbf{A}_{kNT,i},
\end{aligned}
$$

where $\mathbf{A}_{kNT,i} = \frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{j\neq i}^{N}\sum_{t=1}^{T}\gamma_{\ell j}^{0}v_{\ell jt}\varepsilon_{it} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\frac{1}{\sqrt{N}}\sum_{\ell=1}^{k}\sum_{j\neq i}^{N}\gamma_{\ell j}^{0}v_{\ell jt}\right)\varepsilon_{it} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{B}_{kN,t}\varepsilon_{it}$, with $\mathbf{B}_{kN,t} = \frac{1}{\sqrt{N}}\sum_{\ell=1}^{k}\sum_{j\neq i}^{N}\gamma_{\ell j}^{0}v_{\ell jt}$. Clearly, $\mathbf{B}_{kN,t} = O_p(1)$ for each t, k and N as $\boldsymbol{\gamma}_{\ell j}^{0}$ is independent from $v_{\ell jt}$ by Assumption 4. By the same argument, $\mathbf{A}_{kNT,i} = O_p(1)$ for each i, k, N and and T because of independency of ε_{it} by Assumptions 2 and 4. And so, due to independency of $\boldsymbol{\Gamma}_{xi}^{0}$ from $\mathbf{A}_{kNT,i}$, we have

$$
|\mathbf{e}_2| = \frac{1}{\sqrt{N}}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\mathbf{A}_{kNT,i}\right\| = O_p(N^{-1/2}).
$$

Now consider $\mathbf{c}_3$, which is given by

$$
\begin{aligned}
\mathbf{c}_3 &= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\frac{\hat{\mathbf{F}}_x^{\prime}\mathbf{v}_{\ell j}}{T}\frac{\mathbf{v}_{\ell j}^{\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\mathbf{G}_x^{\prime}\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell j}}{T}\frac{\mathbf{v}_{\ell j}^{\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T} \\
&+ \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\hat{\mathbf{Q}}_x^{\prime}\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^{0}\mathbf{G}_x\right)^{\prime}\mathbf{v}_{\ell j}}{T}\frac{\mathbf{v}_{\ell j}^{\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T} \\
&= \mathbf{d}_1 + \mathbf{d}_2,
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{d}_1| &= \left\| \sqrt{\frac{T}{N}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{\Gamma}_{xi}^{0\prime} \hat{\mathbf{Q}}_x^{\prime} \mathbf{G}_x^{\prime} \frac{\mathbf{F}_x^{0\prime} \mathbf{v}_{\ell j}}{T} \frac{\mathbf{v}_{\ell j}^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i}{T} \right\| \\
&= \left\| \sqrt{\frac{T}{N}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\mathbf{v}_{\ell j}^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i}{T} \mathbf{\Gamma}_{xi}^{0\prime} \hat{\mathbf{Q}}_x^{\prime} \mathbf{G}_x^{\prime} \frac{\mathbf{F}_x^{0\prime} \mathbf{v}_{\ell j}}{T} \right\| \\
&\leq T^{-1/2} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{j=1}^{N} \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{v}_{\ell j}^{\prime} \boldsymbol{\varepsilon}_i \mathbf{\Gamma}_{xi}^{0\prime} \right\| \left\| \hat{\mathbf{Q}}_x \right\| \left\| \mathbf{G}_x \right\| \left\| \frac{\mathbf{F}_x^{0\prime} \mathbf{v}_{\ell j}}{\sqrt{T}} \right\| \\
&+ T^{-1} \left(\frac{1}{N} \sum_{\ell=1}^{k} \sum_{j=1}^{N} \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\mathbf{F}_y^{0\prime} \boldsymbol{\varepsilon}_i}{\sqrt{T}} \mathbf{\Gamma}_{xi}^{0\prime} \right\| \left\| \frac{\mathbf{v}_{\ell j}^{\prime} \mathbf{F}_y^0}{\sqrt{T}} \right\| \left\| \frac{\mathbf{F}_x^{0\prime} \mathbf{v}_{\ell j}}{\sqrt{T}} \right\| \right) \left\| \hat{\mathbf{Q}}_x \right\| \left\| \mathbf{G}_x \right\| \left\| \left(\frac{\mathbf{F}_y^{0\prime} \mathbf{F}_y^0}{T} \right)^{-1} \right\| \\
&= O_p\left(T^{-1/2}\right).
\end{aligned}
$$

where the second equality is because $\mathbf{v}_{\ell j}^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i$ is a scalar and therefore commutable, and the order is determined by using Assumptions 1, 2, 3 and 4.

$$
\begin{aligned}
|\mathbf{d}_2| &= \left\| \sqrt{\frac{T}{N}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j=1}^{N} \mathbf{\Gamma}_{xi}^{0\prime} \hat{\mathbf{Q}}_x^{\prime} \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)^{\prime} \mathbf{v}_{\ell j}}{T} \frac{\mathbf{v}_{\ell j}^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i}{T} \right\| \\
&= \left\| \sqrt{\frac{T}{N}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\mathbf{v}_{\ell j}^{\prime} \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i}{T} \mathbf{\Gamma}_{xi}^{0\prime} \hat{\mathbf{Q}}_x^{\prime} \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)^{\prime} \mathbf{v}_{\ell j}}{T} \right\| \\
&\leq \frac{1}{N} \sum_{\ell=1}^{k} \sum_{j=1}^{N} \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{v}_{\ell j}^{\prime} \boldsymbol{\varepsilon}_i \mathbf{\Gamma}_{xi}^{0\prime} \right\| \left\| \hat{\mathbf{Q}}_x \right\| \left\| \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)^{\prime} \mathbf{v}_{\ell j}}{T} \right\| \\
&+ T^{-1/2} \left(\frac{1}{N} \sum_{\ell=1}^{k} \sum_{j=1}^{N} \left\| \frac{\mathbf{v}_{\ell j}^{\prime} \mathbf{F}_y^0}{\sqrt{T}} \right\| \left\| \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)^{\prime} \mathbf{v}_{\ell j}}{T} \right\| \right) \\
&\times \left\| \hat{\mathbf{Q}}_x \right\| \left\| \left(\frac{\mathbf{F}_y^{0\prime} \mathbf{F}_y^0}{T} \right)^{-1} \right\| \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \frac{\mathbf{F}_y^{0\prime} \boldsymbol{\varepsilon}_i}{\sqrt{T}} \mathbf{\Gamma}_{xi}^{0\prime} \right\| \\
&= O_p\left(T^{-1/2}\right),
\end{aligned}
$$

which holds by the same arguments as above and (B.8). By putting the results together, we therefore get

$$
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{\Gamma}_{xi}^{0\prime} \mathbf{F}_x^{0\prime} (\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0}) \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i = o_p(1).
$$

By following similar steps it is also shown that the second term in (B.29) is

$$
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{\Gamma}_{xi}^{0\prime} \mathbf{F}_x^{0\prime} (\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0})(\mathbf{M}_{\hat{F}_y} - \mathbf{M}_{F_y^0}) \mathbf{u}_i = o_p(1).
$$

Hence, we have

$$
\begin{aligned}
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{\Gamma}_{xi}^{0\prime} \mathbf{F}_x^{0\prime} \mathbf{M}_{\hat{F}_x} \mathbf{M}_{\hat{F}_y} \mathbf{u}_i &= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{\Gamma}_{xi}^{0\prime} \mathbf{F}_x^{0\prime} (\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0}) \mathbf{M}_{F_y^0} \boldsymbol{\varepsilon}_i \\
& \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{\Gamma}_{xi}^{0\prime} \mathbf{F}_x^{0\prime} (\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0})(\mathbf{M}_{\hat{F}_y} - \mathbf{M}_{F_y^0}) \mathbf{u}_i \\
&= o_p(1). \qquad \text{(B.30)}
\end{aligned}
$$

The result in (A.10) is shown in the same manner. ■

Proof of Lemma 6. First consider (A.11). The left-hand-side of (A.11) can be written as

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\Gamma}_{xj}^{0}\mathbf{V}_{j}^{\prime}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&=\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\sum_{\ell=1}^{k}\boldsymbol{\gamma}_{\ell j}^{0}\mathbf{v}_{\ell j}^{\prime}\right)\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&=\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\gamma}_{\ell j}^{0}\mathbf{v}_{\ell j}^{\prime}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&=\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\gamma}_{\ell j}^{0}\mathbf{u}_{i}^{\prime}\mathbf{M}_{\hat{F}_x}\mathbf{v}_{\ell j} \\
&=\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xj}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\gamma}_{\ell i}^{0}\mathbf{u}_{j}^{\prime}\mathbf{M}_{\hat{F}_x}\mathbf{v}_{\ell i} \\
&=\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\left(\frac{1}{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xj}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\gamma}_{\ell i}^{0}\mathbf{u}_{j}^{\prime}\right)\mathbf{M}_{\hat{F}_x}\mathbf{v}_{\ell i} \\
&=\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{H}_{\ell,i}^{\prime}\mathbf{M}_{\hat{F}_x}\mathbf{v}_{\ell i},
\end{aligned}$$

where $\mathbf{H}_{\ell,i}=\frac{1}{N}\sum_{j=1}^{N}\mathbf{u}_j\boldsymbol{\gamma}_{\ell i}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\boldsymbol{\Gamma}_{xj}^{0}$. By adding and subtracting terms we get

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{H}_{\ell,i}^{\prime}\mathbf{M}_{\hat{F}_x}\mathbf{v}_{\ell i} \\
&=\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{H}_{\ell,i}^{\prime}\mathbf{M}_{F_x^0}\mathbf{v}_{\ell i}+\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{H}_{\ell,i}^{\prime}(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0})\mathbf{v}_{\ell i}. \qquad \text{(B.31)}
\end{aligned}$$

Now, consider the second term in (B.31). Since $\hat{\mathbf{F}}_x^{\prime}\hat{\mathbf{F}}_x/T=\mathbf{I}_{m_x}$, we have $\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0}=\mathbf{P}_{F_x^0}-\mathbf{P}_{\hat{F}_x}=-\left(\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x^{\prime}}{T}-\mathbf{P}_{F_x^0}\right)$. Using this result and by adding and subtracting terms, we get

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{H}_{\ell,i}^{\prime}(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0})\mathbf{v}_{\ell i} \\
&=-\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{H}_{\ell,i}^{\prime}\left(\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x^{\prime}}{T}-\mathbf{P}_{F_x^0}\right)\mathbf{v}_{\ell i} \\
&=-\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\frac{\mathbf{H}_{\ell,i}^{\prime}\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)}{T}\mathbf{G}_x^{\prime}\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell i} \\
&-\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\frac{\mathbf{H}_{\ell,i}^{\prime}\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)}{T}\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)^{\prime}\mathbf{v}_{\ell i} \\
&-\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\frac{\mathbf{H}_{\ell,i}^{\prime}\mathbf{F}_x^0}{T}\mathbf{G}_x\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)^{\prime}\mathbf{v}_{\ell i} \\
&-\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\frac{\mathbf{H}_{\ell,i}^{\prime}\mathbf{F}_x^0}{T}\left[\mathbf{G}_x\mathbf{G}_x^{\prime}-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell i} \\
&=-\left(\mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3+\mathbf{e}_4\right).
\end{aligned}$$

$$
\begin{aligned}
|\mathbf{e}_1| &\leq \frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \left\| \frac{\mathbf{H}'_{\ell,i}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \mathbf{G}'_x \mathbf{F}_x^{0\prime} \mathbf{v}_{\ell i} \right\| \\
&\leq \frac{1}{\sqrt{N}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\| \boldsymbol{\Gamma}_{xj}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \boldsymbol{\gamma}_{\ell i}^0 \frac{\mathbf{u}'_j\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \mathbf{G}'_x \frac{\mathbf{F}_x^{0\prime} \mathbf{v}_{\ell i}}{\sqrt{T}} \right\| \\
&\leq \sqrt{N} \left(\frac{1}{N} \sum_{j=1}^{N} \left\|\boldsymbol{\Gamma}_{xj}^0\right\| \left\| \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \right\| \left\| \frac{\mathbf{u}'_j\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| \right) \\
&\times \left(\frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \left\|\boldsymbol{\gamma}_{\ell i}^0\right\| \left\|\mathbf{G}_x\right\| \left\| \frac{\mathbf{F}_x^{0\prime} \mathbf{v}_{\ell i}}{\sqrt{T}} \right\| \right) = \sqrt{N} O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

as $\left\| \frac{\mathbf{u}'_j\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| \leq \left\|\boldsymbol{\gamma}_j^{0\prime}\right\| \left\| \frac{\mathbf{F}_y^{0\prime}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| + \left\| \frac{\boldsymbol{\varepsilon}'_j\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| = O_p\left(\delta_{NT}^{-2}\right)$ by (B.6) and (B.7).

$$
\begin{aligned}
|\mathbf{e}_2| &\leq \frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \left\| \frac{\mathbf{H}'_{\ell,i}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)' \mathbf{v}_{\ell i} \right\| \\
&\leq \frac{1}{\sqrt{NT}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\| \boldsymbol{\Gamma}_{xj}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \boldsymbol{\gamma}_{\ell i}^0 \frac{\mathbf{u}'_j\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)' \mathbf{v}_{\ell i} \right\| \\
&\leq \sqrt{NT} \left(\frac{1}{N} \sum_{j=1}^{N} \left\|\boldsymbol{\Gamma}_{xj}^0\right\| \left\| \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \right\| \left\| \frac{\mathbf{u}'_j\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| \right) \\
&\times \left(\frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \left\|\boldsymbol{\gamma}_{\ell i}^0\right\| \left\| \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)' \mathbf{v}_{\ell i}}{T} \right\| \right) = \sqrt{NT} O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

by using $\left\| \frac{\mathbf{u}'_j\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| = O_p\left(\delta_{NT}^{-2}\right)$ and $\left\| \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)' \mathbf{v}_{\ell i}}{T} \right\| = O_p\left(\delta_{NT}^{-2}\right)$ due to (B.8).

$$
\begin{aligned}
\mathbf{e}_3 &= \frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \frac{\mathbf{H}'_{\ell,i} \mathbf{F}_x^0}{T} \mathbf{G}_x \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)' \mathbf{v}_{\ell i} \\
&= \frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \frac{\mathbf{H}'_{\ell,i} \mathbf{F}_x^0}{T} \mathbf{G}_x \mathbf{G}'_x \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)' \mathbf{v}_{\ell i} \\
&= \frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \frac{\mathbf{H}'_{\ell,i} \mathbf{F}_x^0}{T} \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T}\right)^{-1} \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)' \mathbf{v}_{\ell i} \\
&+ \frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \frac{\mathbf{H}'_{\ell,i} \mathbf{F}_x^0}{T} \left[\mathbf{G}_x \mathbf{G}'_x - \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T}\right)^{-1}\right] \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)' \mathbf{v}_{\ell i} \\
&= \mathbf{a}_1 + \mathbf{a}_2.
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{e}_4| &\leq \frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \left\| \frac{\mathbf{H}'_{\ell,i} \mathbf{F}_x^0}{T} \left[\mathbf{G}_x \mathbf{G}'_x - \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T}\right)^{-1}\right] \mathbf{F}_x^{0\prime} \mathbf{v}_{\ell i} \right\| \\
&\leq \sqrt{N} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \left\| \frac{\mathbf{H}'_{\ell,i} \mathbf{F}_x^0}{T} \right\| \left\| \mathbf{G}_x \mathbf{G}'_x - \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T}\right)^{-1} \right\| \left\| \frac{\mathbf{F}_x^{0\prime} \mathbf{v}_{\ell i}}{\sqrt{T}} \right\| \\
&= \sqrt{N} O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

by (B.11) and because

$$\left\| \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \right\| \leq \frac{1}{N}\sum_{j=1}^{N} \left\|\mathbf{\Gamma}^{0\prime}_{xj}\right\| \left\| \left(\mathbf{\Upsilon}^0_{xkN}\right)^{-1} \right\| \left\|\boldsymbol{\gamma}^0_{\ell i}\right\| \left\| \frac{\mathbf{u}'_j}{\sqrt{T}} \right\| \left\| \frac{\mathbf{F}^0_x}{\sqrt{T}} \right\| = O_p\left(1\right),$$

by Assumptions 1, 2, 3, 4. Next, consider $\mathbf{a}_1$ and $\mathbf{a}_2$ in the expression of $\mathbf{e}_3$. Start with $\mathbf{a}_2$ which is as follows

$$\begin{aligned}
\left|\mathbf{a}_2\right| &\leq \frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k}\sum_{i=1}^{N} \left\| \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \left[\mathbf{G}_x\mathbf{G}'_x - \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \right] \left(\hat{\mathbf{F}}_x\mathbf{G}^{-1}_x - \mathbf{F}^0_x \right)' \mathbf{v}_{\ell i} \right\| \\
&\leq \sqrt{NT}\frac{1}{N} \sum_{\ell=1}^{k}\sum_{i=1}^{N} \left\| \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \right\| \left\| \mathbf{G}_x\mathbf{G}'_x - \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \right\| \left\|\mathbf{G}^{-1}_x\right\| \left\| \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}^0_x\mathbf{G}_x\right)' \mathbf{v}_{\ell i}}{T} \right\| \\
&= \sqrt{NT}O_p\left(\delta^{-4}_{NT}\right),
\end{aligned}$$

by (B.8), (B.11) and $\left\| \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \right\| = O_p\left(1\right)$, which is shown above. As for $\mathbf{a}_1$, we have

$$\begin{aligned}
\mathbf{a}_1 &= \frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k}\sum_{i=1}^{N} \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \left(\hat{\mathbf{F}}_x\mathbf{G}^{-1}_x - \mathbf{F}^0_x \right)' \mathbf{v}_{\ell i} \\
&= \frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k}\sum_{i=1}^{N} \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \left(\mathbf{E}'_{x1} + \mathbf{E}'_{x2} + \mathbf{E}'_{x3} \right) \mathbf{v}_{\ell i} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N} \sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \hat{\mathbf{Q}}'_x \sum_{h=1}^{k} \frac{\hat{\mathbf{F}}'_x\mathbf{v}_{hj}}{T} \boldsymbol{\gamma}^{0\prime}_{hj} \frac{\mathbf{F}^{0\prime}_x\mathbf{v}_{\ell i}}{T} \\
&+ \sqrt{\frac{T}{N}}\frac{1}{N} \sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \hat{\mathbf{Q}}'_x \frac{\hat{\mathbf{F}}'_x\mathbf{F}^0_x}{T} \sum_{h=1}^{k} \boldsymbol{\gamma}^0_{hj} \frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&+ \sqrt{\frac{T}{N}}\frac{1}{N} \sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \hat{\mathbf{Q}}'_x \sum_{h=1}^{k} \frac{\hat{\mathbf{F}}'_x\mathbf{v}_{hj}}{T} \frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&= \mathbf{c}_1 + \mathbf{c}_2 + \mathbf{c}_3.
\end{aligned}$$

$$\begin{aligned}
\mathbf{c}_1 &= \sqrt{\frac{T}{N}}\frac{1}{N} \sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \hat{\mathbf{Q}}'_x \sum_{h=1}^{k} \frac{\hat{\mathbf{F}}'_x\mathbf{v}_{hj}}{T} \boldsymbol{\gamma}^{0\prime}_{hj} \frac{\mathbf{F}^{0\prime}_x\mathbf{v}_{\ell i}}{T} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N} \sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \hat{\mathbf{Q}}'_x\mathbf{G}'_x \sum_{h=1}^{k} \frac{\mathbf{F}^{0\prime}_x\mathbf{v}_{hj}}{T} \boldsymbol{\gamma}^{0\prime}_{hj} \frac{\mathbf{F}^{0\prime}_x\mathbf{v}_{\ell i}}{T} \\
&+ \sqrt{\frac{T}{N}}\frac{1}{N} \sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \hat{\mathbf{Q}}'_x \sum_{h=1}^{k} \frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}^0_x\mathbf{G}_x\right)' \mathbf{v}_{hj}}{T} \boldsymbol{\gamma}^{0\prime}_{hj} \frac{\mathbf{F}^{0\prime}_x\mathbf{v}_{\ell i}}{T} \\
&= \mathbf{d}_1 + \mathbf{d}_2
\end{aligned}$$

$$\begin{aligned}
\left|\mathbf{d}_1\right| &\leq \frac{1}{N} \sum_{\ell=1}^{k}\sum_{i=1}^{N} \left\| \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \hat{\mathbf{Q}}'_x\mathbf{G}'_x \left(\sum_{h=1}^{k} \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \frac{\mathbf{F}^{0\prime}_x\mathbf{v}_{hj}}{\sqrt{T}} \boldsymbol{\gamma}^{0\prime}_{hj} \right) \frac{\mathbf{F}^{0\prime}_x\mathbf{v}_{\ell i}}{T} \right\| \\
&\leq T^{-1/2} \left(\frac{1}{N} \sum_{\ell=1}^{k}\sum_{i=1}^{N} \left\| \frac{\mathbf{H}'_{\ell,i}\mathbf{F}^0_x}{T} \right\| \left\| \left(\frac{\mathbf{F}^{0\prime}_x\mathbf{F}^0_x}{T} \right)^{-1} \right\| \left\|\hat{\mathbf{Q}}'_x\right\| \left\|\mathbf{G}'_x\right\| \left\| \frac{\mathbf{F}^{0\prime}_x\mathbf{v}_{\ell i}}{\sqrt{T}} \right\| \right) \\
&\times \left\| \frac{1}{\sqrt{N}} \sum_{h=1}^{k}\sum_{j=1}^{N} \frac{\mathbf{F}^{0\prime}_x\mathbf{v}_{hj}}{\sqrt{T}} \boldsymbol{\gamma}^{0\prime}_{hj} \right\| \\
&= T^{-1/2}O_p\left(1\right) = O_p\left(T^{-1/2}\right).
\end{aligned}$$

$$
\begin{aligned}
|\mathbf{d}_2| &\le \sqrt{N}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\left\|\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\left(\frac{1}{N}\sum_{h=1}^{k}\sum_{j=1}^{N}\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{v}_{hj}}{T}\gamma_{hj}^{0\prime}\right)\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell i}}{\sqrt{T}}\right\| \\
&\le \sqrt{N}\left(\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\left\|\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\right\|\left\|\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\hat{\mathbf{Q}}'_x\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell i}}{\sqrt{T}}\right\|\right) \\
&\times\left(\frac{1}{N}\sum_{h=1}^{k}\sum_{j=1}^{N}\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{v}_{hj}}{T}\right\|\left\|\gamma_{hj}^{0\prime}\right\|\right) \\
&= \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

by (B.8).

$$
\begin{aligned}
\mathbf{c}_2 &= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\frac{\hat{\mathbf{F}}'_x\mathbf{F}_x^0}{T}\sum_{h=1}^{k}\gamma_{hj}^{0}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\sum_{h=1}^{k}\gamma_{hj}^{0}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T}.
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{c}_3 &= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\sum_{h=1}^{k}\frac{\hat{\mathbf{F}}'_x\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\mathbf{G}'_x\sum_{h=1}^{k}\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&+ \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\mathbf{G}'_x\sum_{h=1}^{k}\frac{\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)'\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&= \mathbf{d}_1+\mathbf{d}_2,
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{d}_1 &= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\mathbf{G}'_x\sum_{h=1}^{k}\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&= \frac{1}{\sqrt{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\mathbf{G}'_x\sum_{h=1}^{k}\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{hi}}{\sqrt{T}}\frac{\mathbf{v}'_{hi}\mathbf{v}_{\ell i}}{T} \\
&+ \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\mathbf{G}'_x\sum_{h=1}^{k}\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&= \mathbf{b}_1+\mathbf{b}_2.
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{b}_1| &\le \frac{1}{\sqrt{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{h=1}^{k}\sum_{i=1}^{N}\left\|\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\right\|\left\|\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\hat{\mathbf{Q}}_x\right\|\left\|\mathbf{G}_x\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{hi}}{\sqrt{T}}\right\|\left\|\frac{\mathbf{v}'_{hi}\mathbf{v}_{\ell i}}{T}\right\| \\
&= O_p\left(N^{-1/2}\right),
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{b}_2| &\leq \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\left\|\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\mathbf{G}'_x\sum_{h=1}^{k}\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T}\right\| \\
&\leq \sqrt{\frac{T}{N}}\frac{1}{N^2}\sum_{\ell=1}^{k}\sum_{h=1}^{k}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\sum_{n=1}^{N}\left\|\boldsymbol{\Gamma}_{xn}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell i}^0\frac{\mathbf{u}'_n\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\mathbf{G}'_x\frac{\mathbf{F}_x^{0\prime}\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T}\right\| \\
&= \sqrt{\frac{T}{N}}\frac{1}{N^2}\sum_{\ell=1}^{k}\sum_{h=1}^{k}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\sum_{n=1}^{N}\left\|\boldsymbol{\Gamma}_{xn}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell i}^0\frac{\mathbf{v}'_{\ell i}\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\mathbf{F}_x^0}{T}\mathbf{G}_x\hat{\mathbf{Q}}_x\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_n}{T}\right\| \\
&\leq \frac{1}{\sqrt{T}}\sum_{\ell=1}^{k}\sum_{h=1}^{k}\left(\frac{1}{N}\sum_{n=1}^{N}\left\|\boldsymbol{\Gamma}_{xn}^{0\prime}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_n}{T}\right\|\right)\left(\frac{1}{N}\sum_{j\neq i}^{N}\left\|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\gamma_{\ell i}^0\frac{\mathbf{v}'_{\ell i}\mathbf{v}_{hj}}{\sqrt{T}}\right\|\left\|\frac{\mathbf{v}'_{hj}\mathbf{F}_x^0}{\sqrt{T}}\right\|\right) \\
&\times\|\mathbf{G}_x\|\left\|\hat{\mathbf{Q}}_x\right\|\left\|\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\right\| \\
&= O_p\left(T^{-1/2}\right).
\end{aligned}
$$

Also, we have

$$
\begin{aligned}
\mathbf{d}_2 &= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\mathbf{G}'_x\sum_{h=1}^{k}\frac{\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)'\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N^2}\sum_{\ell=1}^{k}\sum_{h=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_{xn}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell i}^0\frac{\mathbf{u}'_n\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{Q}}'_x\mathbf{G}'_x \\
&\times\frac{\left(\hat{\mathbf{F}}_x\mathbf{G}^{-1}-\mathbf{F}_x^0\right)'\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&= \sqrt{\frac{T}{N}}\frac{1}{N^2}\sum_{\ell=1}^{k}\sum_{h=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_{xn}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell i}^0\frac{\mathbf{v}'_{\ell i}\mathbf{v}_{hj}}{T}\frac{\mathbf{v}'_{hj}\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)}{T} \\
&\times\mathbf{G}_x\hat{\mathbf{Q}}_x\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_n}{T},
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{d}_2| &\leq \sqrt{N}\sqrt{\frac{T}{N}}\left(\frac{1}{N}\sum_{n=1}^{N}\left\|\boldsymbol{\Gamma}_{xn}^{0\prime}\right\|\left\|\frac{\mathbf{u}_n}{\sqrt{T}}\right\|\right)\left\|\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\right\|\left\|\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\gamma_{\ell i}^0\mathbf{v}'_{\ell i}\right\| \\
&\times\left(\frac{1}{N}\sum_{h=1}^{k}\sum_{j=1}^{N}\left\|\frac{\mathbf{v}_{hj}}{\sqrt{T}}\right\|\left\|\frac{\mathbf{v}'_{hj}\left(\hat{\mathbf{F}}'_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)}{T}\right\|\right)\|\mathbf{G}_x\|\left\|\hat{\mathbf{Q}}_x\right\|\left\|\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}}{\sqrt{T}}\right\| \\
&= \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

by (B.8). By putting the results together, we get

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{H}'_{\ell,i}(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0})\mathbf{v}_{\ell i} \\
&= -\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\sum_{h=1}^{k}\gamma_{hj}^0\frac{\mathbf{v}'_{hj}\mathbf{v}_{\ell i}}{T} \\
&+ \sqrt{N}O_p\left(\delta_{NT}^{-2}\right)+O_p\left(\delta_{NT}^{-1}\right) \\
&= -\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{H}'_{\ell,i}\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\boldsymbol{\Gamma}_{xj}^0\frac{\mathbf{V}'_j\mathbf{v}_{\ell i}}{T} \\
&+ \sqrt{N}O_p\left(\delta_{NT}^{-2}\right)+O_p\left(\delta_{NT}^{-1}\right).
\end{aligned}
$$

And so, we have

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{H}_{\ell,i}'\mathbf{M}_{\hat{F}_x}\mathbf{v}_{\ell i}\\
&=\frac{1}{\sqrt{NT}}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{H}_{\ell,i}'\mathbf{M}_{F_x^0}\mathbf{v}_{\ell i}\\
&-\sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{H}_{\ell,i}'\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\boldsymbol{\Gamma}_{xj}^0\frac{\mathbf{V}_j'\mathbf{v}_{\ell i}}{T}\\
&+\sqrt{N}O_p\left(\delta_{NT}^{-2}\right)+O_p\left(\delta_{NT}^{-1}\right)\\
&=\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xj}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell i}^0\mathbf{u}_j'\mathbf{M}_{F_x^0}\mathbf{v}_{\ell i}\\
&-\sqrt{\frac{T}{N}}\frac{1}{N^2}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_{xn}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell i}^0\frac{\mathbf{u}_n'\mathbf{F}_x^0}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\boldsymbol{\Gamma}_{xj}^0\frac{\mathbf{V}_j'\mathbf{v}_{\ell i}}{T}\\
&+\sqrt{N}O_p\left(\delta_{NT}^{-2}\right)+O_p\left(\delta_{NT}^{-1}\right)\\
&=\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\boldsymbol{\Gamma}_{xj}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell i}^0\mathbf{v}_{\ell i}'\mathbf{M}_{F_x^0}\mathbf{u}_j\\
&-\sqrt{\frac{T}{N}}\frac{1}{N^2}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_{xn}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\gamma_{\ell i}^0\frac{\mathbf{v}_{\ell i}'\mathbf{V}_j}{T}\boldsymbol{\Gamma}_{xj}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_n}{T}\\
&+\sqrt{N}O_p\left(\delta_{NT}^{-2}\right)+O_p\left(\delta_{NT}^{-1}\right)\\
&=\frac{1}{\sqrt{NT}}\frac{1}{N}\sum_{i=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\boldsymbol{\Gamma}_{xn}^0\mathbf{V}_{n,-1}'\mathbf{M}_{F_{x,-1}^0}\mathbf{M}_{F_x^0}\mathbf{u}_i\\
&-\sqrt{\frac{T}{N}}\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\boldsymbol{\Gamma}_{xn}^0\frac{\mathbf{V}_n'\mathbf{V}_j}{T}\boldsymbol{\Gamma}_{xj}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T}\\
&+\sqrt{N}O_p\left(\delta_{NT}^{-2}\right)+O_p\left(\delta_{NT}^{-1}\right),
\end{aligned}\tag{B.32}
$$

which provides the required result in (A.11). The result in (A.12) is proved in the same way. ■

Proof of Lemma 7. We first prove (A.13). By noting that $\hat{\mathbf{A}}\hat{\mathbf{B}}=\mathbf{A}\mathbf{B}+\left(\hat{\mathbf{A}}-\mathbf{A}\right)\hat{\mathbf{B}}+\mathbf{A}\left(\hat{\mathbf{B}}-\mathbf{B}\right)$, the left-hand-side of (A.13) can be written as

$$
\begin{aligned}
&\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x'\bar{\boldsymbol{\Sigma}}_{kNT}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i\\
&=\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\bar{\boldsymbol{\Sigma}}_{kNT}\mathbf{M}_{F_x^0}\mathbf{u}_i\\
&+\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x'\bar{\boldsymbol{\Sigma}}_{kNT}\left(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0}\right)\mathbf{u}_i\\
&+\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\left[\left(\frac{\hat{\mathbf{F}}_x'\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\right]\bar{\boldsymbol{\Sigma}}_{kNT}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i\\
&=\sqrt{\frac{N}{T}}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\bar{\boldsymbol{\Sigma}}_{kNT}\mathbf{M}_{F_x^0}\mathbf{u}_i\\
&+\sqrt{\frac{N}{T}}\left(\mathbf{a}_1+\mathbf{a}_2\right).
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{a}_1| &\leq \frac{1}{NT}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x^{\prime}\bar{\boldsymbol{\Sigma}}_{kNT}\left(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0}\right)\mathbf{u}_i\right\| \\
&\leq \left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\|\right)\left\|\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\|\left\|\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\hat{\mathbf{F}}_x}{\sqrt{T}}\right\|\lambda_{\max}\left(\bar{\boldsymbol{\Sigma}}_{kNT}\right)\left\|\mathbf{P}_{\hat{F}_x}-\mathbf{P}_{F_x^0}\right\| \\
&= O_p\left(\delta_{NT}^{-1}\right),
\end{aligned}
$$

by Lemma B.6 and $T^{-1/2}\left\|\bar{\boldsymbol{\Sigma}}_{kNT}\hat{\mathbf{F}}_x\right\| \leq \lambda_{\max}\left(\bar{\boldsymbol{\Sigma}}_{kNT}\right)T^{-1/2}\left\|\hat{\mathbf{F}}_x\right\| = O_p(1)$ which holds by Assumptions 2 and 3. Next, we have

$$
\begin{aligned}
|\mathbf{a}_2| &\leq \frac{1}{NT}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\left[\left(\frac{\hat{\mathbf{F}}_x^{\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x^{\prime}-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\right]\bar{\boldsymbol{\Sigma}}_{kNT}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i\right\| \\
&\leq \frac{1}{\sqrt{T}}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\|\right)\left\|\left[\hat{\mathbf{F}}_x\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}-\mathbf{F}_x^0\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\| \\
&\times\lambda_{\max}\left(\bar{\boldsymbol{\Sigma}}_{kNT}\right) \\
&+\frac{1}{\sqrt{T}}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\boldsymbol{\Gamma}_{xi}^{0\prime}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\|\right)\left\|\left[\hat{\mathbf{F}}_x\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}-\mathbf{F}_x^0\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\| \\
&\times\lambda_{\max}\left(\bar{\boldsymbol{\Sigma}}_{kNT}\right)\left\|\mathbf{P}_{\hat{F}_x}\right\| \\
&= O_p(1)\frac{1}{\sqrt{T}}\left\|\left[\hat{\mathbf{F}}_x\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}-\mathbf{F}_x^0\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\|,
\end{aligned}
$$

by Assumptions 1, 2, 3, 4, using $\left\|\mathbf{P}_{\hat{F}_x}\right\| = \left\|\hat{\mathbf{F}}_x\left(\hat{\mathbf{F}}_x^{\prime}\hat{\mathbf{F}}_x\right)^{-1}\hat{\mathbf{F}}_x^{\prime}\right\| = \left\|\hat{\mathbf{F}}_x^{\prime}\hat{\mathbf{F}}_x\left(\hat{\mathbf{F}}_x^{\prime}\hat{\mathbf{F}}_x\right)^{-1}\right\| = \left\|\mathbf{I}_{m_x}\right\| = \sqrt{m_x}$, which holds by properties of norms and definition of the matrix $\mathbf{P}_{\hat{F}_x}$, and because

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\left\|\left[\hat{\mathbf{F}}_x\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}-\mathbf{F}_x^0\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\| \\
&=\frac{1}{\sqrt{T}}\left\|\left[\hat{\mathbf{F}}_x\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}-\mathbf{F}_x^0\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}\right]\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\| \\
&=\frac{1}{\sqrt{T}}\left\|\left[\hat{\mathbf{F}}_x\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}-\mathbf{P}_{F_x^0}\hat{\mathbf{F}}_x\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}\right]\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\| \\
&=\frac{1}{\sqrt{T}}\left\|\mathbf{M}_{F_x^0}\hat{\mathbf{F}}_x\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\| \leq \frac{1}{\sqrt{T}}\left\|\mathbf{M}_{F_x^0}\hat{\mathbf{F}}_x\right\|\left\|\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\| \\
&= O_p\left(\delta_{NT}^{-1}\right),
\end{aligned}
$$

since $T^{-1/2}\left\|\mathbf{M}_{F_x^0}\hat{\mathbf{F}}_x\right\| = T^{-1/2}\left\|(\mathbf{M}_{F_x^0}-\mathbf{M}_{\hat{F}_x})\hat{\mathbf{F}}_x\right\| \leq \left\|\mathbf{M}_{F_x^0}-\mathbf{M}_{\hat{F}_x}\right\|T^{-1/2}\left\|\hat{\mathbf{F}}_x\right\| = \left\|\mathbf{P}_{F_x^0}-\mathbf{P}_{\hat{F}_x}\right\|T^{-1/2}\left\|\hat{\mathbf{F}}_x\right\| = O_p\left(\delta_{NT}^{-1}\right)$ by Lemma B.6, $\left\|\left(\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\right)^{-1}\right\| = O_p(1)$ and $\left\|\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\right\| = O_p(1)$.

Hence, by putting the results together we get the required result in (A.13), which is

$$
\begin{aligned}
&\frac{1}{\sqrt{N}T^{3/2}}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\hat{\mathbf{F}}_x^{\prime}\mathbf{F}_x^0}{T}\right)^{-1}\hat{\mathbf{F}}_x^{\prime}\bar{\boldsymbol{\Sigma}}_{kNT}\mathbf{M}_{\hat{F}_x}\mathbf{u}_i \\
&=\sqrt{\frac{N}{T}}\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\Gamma}_{xi}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^{0}\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\bar{\boldsymbol{\Sigma}}_{kNT}\mathbf{M}_{F_x^0}\mathbf{u}_i+o_p(1).
\end{aligned}
$$

The result in (A.14) is proved by following the same steps. ■

Proof of Lemma 8. We start with (A.16), which is given by

$$\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i + \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0}\right)\mathbf{u}_i. \tag{B.33}$$

Now consider the second term in (B.33). By using $\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0} = -\left(\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x'}{T} - \mathbf{P}_{F_x^0}\right)$, and adding and subtracting terms, we get

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0}\right)\mathbf{u}_i \\
&= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left(\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x'}{T} - \mathbf{P}_{F_x^0}\right)\mathbf{u}_i \\
&= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{u}_i \\
&- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_x^0}{T}\mathbf{G}_x\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{u}_i \\
&- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_x^0}{T}\left[\mathbf{G}_x\mathbf{G}_x' - \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&= -\left(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4\right),
\end{aligned}$$

$$|\mathbf{e}_2| \le \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_{i,-1}'\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{u}_i}{T}\right\| = \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right),$$

by $\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{u}_i}{T}\right\| = O_p\left(\delta_{NT}^{-2}\right)$ and (B.8).

$$|\mathbf{e}_3| \le \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\mathbf{G}_x\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{u}_i}{T}\right\| = \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),$$

by using again $\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{u}_i}{T}\right\| = O_p\left(\delta_{NT}^{-2}\right)$.

$$|\mathbf{e}_4| \le \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\mathbf{G}_x\mathbf{G}_x' - \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\| = \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),$$

by (B.11). Consider now $\mathbf{e}_1$ which can be written as

$$
\begin{aligned}
\mathbf{e}_1 &= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_{i,-1}' \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \mathbf{G}_x' \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T} \mathbf{G}_x \mathbf{G}_x' \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T} \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T} \left[\mathbf{G}_x \mathbf{G}_x' - \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right] \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&= \mathbf{a}_1 + \mathbf{a}_2.
\end{aligned}
$$

Start with $\mathbf{a}_2$ which is given by

$$
\begin{aligned}
|\mathbf{a}_2| &\leq \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T} \right\| \left\| \mathbf{G}_x \mathbf{G}_x' - \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \right\| \left\| \frac{\mathbf{F}_x^0}{\sqrt{T}} \right\| \left\| \frac{\mathbf{u}_i}{\sqrt{T}} \right\| \\
&= \sqrt{NT} O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

by (B.8) and (B.11). As for $\mathbf{a}_1$ we have

$$
\begin{aligned}
\mathbf{a}_1 &= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_x \mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T} \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{1}{NT} \sum_{\ell=1}^{k} \sum_{j=1}^{N} \frac{\mathbf{V}_i' \mathbf{F}_x^0}{T} \gamma_{\ell j}^0 \mathbf{v}_{\ell j}' \hat{\mathbf{F}}_x \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{1}{NT} \sum_{\ell=1}^{k} \sum_{j=1}^{N} \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{T} \gamma_{\ell j}^{0\prime} \mathbf{F}_x^{0\prime} \hat{\mathbf{F}}_x \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{1}{NT} \sum_{\ell=1}^{k} \sum_{j=1}^{N} \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{T} \mathbf{v}_{\ell j}' \hat{\mathbf{F}}_x \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&= \mathbf{d}_1 + \mathbf{d}_2 + \mathbf{d}_3,
\end{aligned}
$$

$$
\begin{aligned}
&|\mathbf{d}_1| \\
&\leq \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{\mathbf{V}_i' \mathbf{F}_x^0}{\sqrt{T}} \left(\frac{1}{\sqrt{N}} \sum_{\ell=1}^{k} \sum_{j=1}^{N} \gamma_{\ell j}^0 \frac{\mathbf{v}_{\ell j}' \hat{\mathbf{F}}_x}{T} \right) \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T} \right\| \\
&\leq \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{\mathbf{V}_i' \mathbf{F}_x^0}{\sqrt{T}} \right\| \left\| \frac{1}{\sqrt{N}} \sum_{\ell=1}^{k} \sum_{j=1}^{N} \gamma_{\ell j}^0 \frac{\mathbf{v}_{\ell j}' \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x\right)}{T} \right\| \left\| \hat{\mathbf{Q}}_x \right\| \left\| \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \right\| \left\| \frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{\sqrt{T}} \right\| \\
&+ \frac{1}{\sqrt{T}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \left\| \frac{\mathbf{V}_i' \mathbf{F}_x^0}{\sqrt{T}} \right\| \left\| \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \gamma_{\ell j}^0 \frac{\mathbf{v}_{\ell j}' \mathbf{F}_x^0}{\sqrt{T}} \right\| \left\| \mathbf{G}_x \right\| \left\| \hat{\mathbf{Q}}_x \right\| \left\| \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1} \right\| \left\| \frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{\sqrt{T}} \right\| \\
&= O_p\left(\delta_{NT}^{-2}\right) + O_p\left(N^{-1/2}\right) + O_p\left(T^{-1/2}\right) = O_p\left(\delta_{NT}^{-1}\right).
\end{aligned}
$$

by (B.10).

$$\begin{aligned}
\mathbf{d}_2 &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime}\frac{\mathbf{F}_x^{0\prime}\hat{\mathbf{F}}_x}{T}\hat{\mathbf{Q}}_x\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{T}\gamma_{\ell j}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&= \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{V}_j}{T}\boldsymbol{\Gamma}_{xj}^{0\prime}\left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T}.
\end{aligned}$$

$$\begin{aligned}
\mathbf{d}_3 &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{T}\mathbf{v}_{\ell j}'\hat{\mathbf{F}}_x\hat{\mathbf{Q}}_x\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{T}\mathbf{v}_{\ell j}'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)\hat{\mathbf{Q}}_x\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&+ \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{T}\mathbf{v}_{\ell j}'\mathbf{F}_x^0\mathbf{G}_x\hat{\mathbf{Q}}_x\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&= \mathbf{c}_1+\mathbf{c}_2,
\end{aligned}$$

$$\begin{aligned}
\mathbf{c}_1 &= \frac{1}{\sqrt{NT}}\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{T}\mathbf{v}_{\ell j}'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)\hat{\mathbf{Q}}_x\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{T}\mathbf{v}_{\ell j}'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)\hat{\mathbf{Q}}_x\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&+ \frac{1}{\sqrt{NT}}\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\sum_{j\neq i}^{N}\frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{T}\mathbf{v}_{\ell j}'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)\hat{\mathbf{Q}}_x\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&= \mathbf{b}_1+\mathbf{b}_2.
\end{aligned}$$

$$\begin{aligned}
|\mathbf{b}_1| &\leq \sqrt{\frac{T}{N}}\frac{1}{N}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i}{\sqrt{T}}\right\|\left\|\frac{\mathbf{v}_{\ell i}}{\sqrt{T}}\right\|\left\|\frac{\mathbf{v}_{\ell i}'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)}{T}\right\|\left\|\hat{\mathbf{Q}}_x\right\|\left\|\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\| \\
&= \sqrt{\frac{T}{N}}O_p\left(\delta_{NT}^{-2}\right) = O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}$$

by (B.8).

$$
\begin{aligned}
|\mathbf{b}_2| &\leq \left\| \frac{1}{\sqrt{NT}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{\sqrt{T}} \frac{\mathbf{v}_{\ell j}' \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)}{\sqrt{T}} \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \frac{\mathbf{F}_x^{0\prime} \mathbf{F}_y^0}{T} \boldsymbol{\gamma}_{yi} \right\| \\
&+ \left\| \frac{1}{\sqrt{N}} \frac{1}{NT} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{\sqrt{T}} \frac{\mathbf{v}_{\ell j}' \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)}{\sqrt{T}} \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \frac{\mathbf{F}_x^{0\prime} \boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \\
&\leq \left\| \frac{1}{\sqrt{NT}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{\mathbf{v}_{\ell j}' \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)}{\sqrt{T}} \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \frac{\mathbf{F}_x^{0\prime} \mathbf{F}_y^0}{T} \boldsymbol{\gamma}_{yi} \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{\sqrt{T}} \right\| \\
&+ \left\| \frac{1}{\sqrt{N}} \frac{1}{NT} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{\sqrt{T}} \frac{\mathbf{v}_{\ell j}' \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)}{\sqrt{T}} \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \frac{\mathbf{F}_x^{0\prime} \boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \\
&\leq \frac{1}{N} \sum_{\ell=1}^{k} \sum_{j\neq i}^{N} \left\| \frac{\mathbf{v}_{\ell j}' \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)}{T} \right\| \left\| \hat{\mathbf{Q}}_x \right\| \left\| \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \right\| \left\| \frac{\mathbf{F}_x^0}{\sqrt{T}} \right\| \left\| \frac{\mathbf{F}_y^0}{\sqrt{T}} \right\| \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\gamma}_{yi} \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{\sqrt{T}} \right\| \\
&+ \sqrt{\frac{N}{T}} \frac{1}{N^2} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \left\| \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{\sqrt{T}} \right\| \left\| \frac{\mathbf{v}_{\ell j}' \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)}{T} \right\| \left\| \hat{\mathbf{Q}}_x \right\| \left\| \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \right\| \left\| \frac{\mathbf{F}_x^{0\prime} \boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \\
&= O_p \left(\delta_{NT}^{-2} \right).
\end{aligned}
$$

by using independency of $\mathbf{v}_{\ell i}$ from $\mathbf{v}_{\ell j}$ when $i \neq j$ as well as (B.8).

$$
\begin{aligned}
\mathbf{c}_2 &= \frac{1}{\sqrt{NT}} \frac{1}{NT} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{T} \mathbf{v}_{\ell j}' \mathbf{F}_x^0 \mathbf{G}_x \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}} \frac{1}{NT} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{T} \mathbf{v}_{\ell j}' \mathbf{F}_x^0 \mathbf{G}_x \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&+ \frac{1}{\sqrt{NT}} \frac{1}{NT} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{\mathbf{V}_i' \mathbf{v}_{\ell j}}{T} \mathbf{v}_{\ell j}' \mathbf{F}_x^0 \mathbf{G}_x \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \mathbf{F}_x^{0\prime} \mathbf{u}_i \\
&= \mathbf{b}_1 + \mathbf{b}_2.
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{b}_1| &\leq \sqrt{\frac{1}{N}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \left\| \frac{\mathbf{V}_i}{\sqrt{T}} \right\| \left\| \frac{\mathbf{v}_{\ell i}}{\sqrt{T}} \right\| \left\| \frac{\mathbf{v}_{\ell i}' \mathbf{F}_x^0}{\sqrt{T}} \right\| \left\| \hat{\mathbf{Q}}_x \right\| \left\| \mathbf{G}_x \right\| \left\| \left(\frac{\mathbf{F}_x^{0\prime} \mathbf{F}_x^0}{T} \right)^{-1} \right\| \left\| \frac{\mathbf{F}_x^0}{\sqrt{T}} \right\| \left\| \frac{\mathbf{u}_i}{\sqrt{T}} \right\| \\
&= O_p \left(N^{-1/2} \right),
\end{aligned}
$$

and also

$$
\begin{aligned}
|\mathbf{b}_2| &\leq \left\| \frac{1}{\sqrt{NT}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{\sqrt{T}} \frac{\mathbf{v}_{\ell j}'\mathbf{F}_x^0}{\sqrt{T}} \mathbf{G}_x \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T} \right)^{-1} \frac{\mathbf{F}_x^{0\prime}\mathbf{F}_y^0}{T} \boldsymbol{\gamma}_{yi} \right\| \\
&+ \left\| \frac{1}{\sqrt{N}} \frac{1}{NT} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{\sqrt{T}} \frac{\mathbf{v}_{\ell j}'\mathbf{F}_x^0}{\sqrt{T}} \mathbf{G}_x \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T} \right)^{-1} \frac{\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \\
&\leq \left\| \frac{1}{\sqrt{NT}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{\mathbf{v}_{\ell j}'\mathbf{F}_x^0}{\sqrt{T}} \mathbf{G}_x \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T} \right)^{-1} \frac{\mathbf{F}_x^{0\prime}\mathbf{F}_y^0}{T} \boldsymbol{\gamma}_{yi} \frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{\sqrt{T}} \right\| \\
&+ \left\| \frac{1}{\sqrt{N}} \frac{1}{NT} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{\sqrt{T}} \frac{\mathbf{v}_{\ell j}'\mathbf{F}_x^0}{\sqrt{T}} \mathbf{G}_x \hat{\mathbf{Q}}_x \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T} \right)^{-1} \frac{\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \\
&\leq \frac{1}{\sqrt{T}} \frac{1}{N} \sum_{\ell=1}^{k} \sum_{j\neq i}^{N} \left\| \frac{\mathbf{v}_{\ell j}'\mathbf{F}_x^0}{\sqrt{T}} \right\| \|\mathbf{G}_x\| \left\| \hat{\mathbf{Q}}_x \right\| \left\| \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T} \right)^{-1} \right\| \left\| \frac{\mathbf{F}_x^0}{\sqrt{T}} \right\| \left\| \frac{\mathbf{F}_y^0}{\sqrt{T}} \right\| \left\| \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{\gamma}_{yi} \frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{\sqrt{T}} \right\| \\
&+ \frac{1}{\sqrt{N}} \frac{N}{T} \frac{1}{N^2} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \sum_{j\neq i}^{N} \left\| \frac{\mathbf{V}_i'\mathbf{v}_{\ell j}}{\sqrt{T}} \right\| \left\| \frac{\mathbf{v}_{\ell j}'\mathbf{F}_x^0}{\sqrt{T}} \right\| \|\mathbf{G}_x\| \left\| \hat{\mathbf{Q}}_x \right\| \left\| \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T} \right)^{-1} \right\| \left\| \frac{\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \\
&= O_p\left(N^{-1/2}\right) + O_p\left(T^{-1/2}\right) = O_p\left(\delta_{NT}^{-1}\right).
\end{aligned}
$$

by making use of independency of $\mathbf{v}_{\ell i}$ from $\mathbf{v}_{\ell j}$ when $i \neq j$. By adding everything together we therefore have

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}} \sum_{\ell=1}^{k} \sum_{i=1}^{N} \mathbf{V}_i'(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0})\mathbf{u}_i \\
&= -\sqrt{\frac{T}{N}} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\mathbf{V}_i'\mathbf{V}_j}{T} \boldsymbol{\Gamma}_{xj}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T} \right)^{-1} \frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T} \\
&+ \sqrt{N} O_p\left(\delta_{NT}^{-2}\right) + O_p\left(\delta_{NT}^{-1}\right).
\end{aligned}
$$

Collecting the results together, we obtain

$$
\begin{aligned}
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i = &\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i \\
&- \sqrt{\frac{T}{N}} \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\mathbf{V}_i'\mathbf{V}_j}{T} \boldsymbol{\Gamma}_{xj}^{0\prime} \left(\boldsymbol{\Upsilon}_{xkN}^0\right)^{-1} \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T} \right)^{-1} \frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T} \\
&+ \sqrt{N} O_p\left(\delta_{NT}^{-2}\right) + O_p\left(\delta_{NT}^{-1}\right),
\end{aligned}
$$

which is the required result in (A.15). The result in (A.16) can be shown by following the similar steps as discussed above. ■

Proof of Lemma 9. We begin with (A.17), which is

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_y}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i'\mathbf{M}_{F_x^0}\mathbf{M}_{F_y^0}\mathbf{u}_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i' \left(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0}\right) \mathbf{M}_{F_y^0}\mathbf{u}_i \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i'\mathbf{M}_{\hat{F}_x} \left(\mathbf{M}_{\hat{F}_y} - \mathbf{M}_{F_y^0}\right) \mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i'\mathbf{M}_{F_x^0}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i' \left(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0}\right) \mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}_i'\mathbf{M}_{\hat{F}_x} \left(\mathbf{M}_{\hat{F}_y} - \mathbf{M}_{F_y^0}\right) \mathbf{u}_i, \qquad \text{(B.34)}
\end{aligned}
$$

where the second equality is because of $\mathbf{M}_{F_x^0}\mathbf{u}_i = \mathbf{M}_{F_x^0}\boldsymbol{\varepsilon}_i$. Now consider the second term in (B.34). By using $\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0} = -\left(\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x'}{T} - \mathbf{P}_{F_x^0}\right)$, and adding and subtracting terms, we get

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0}\right)\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left(\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x'}{T} - \mathbf{P}_{F_x^0}\right)\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&\quad - \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&\quad - \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_x^0}{T}\mathbf{G}_x\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&\quad - \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_x^0}{T}\left[\mathbf{G}_x\mathbf{G}_x' - \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= -\left(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4\right),
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{e}_1 &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T}\mathbf{G}_x\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T}\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&\quad + \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T}\left[\mathbf{G}_x\mathbf{G}_x' - \left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i \\
&= \mathbf{a}_1 + \mathbf{a}_2,
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{a}_1| &\le \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T}\right\|\left\|\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&\le \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T}\right\|\left\|\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&\quad + \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1} - \mathbf{F}_x^0\right)}{T}\right\|\left\|\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

by (B.8) and

$$
\begin{aligned}
|\mathbf{a}_2| &\le \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)}{T}\right\|\left\|\mathbf{G}_x\mathbf{G}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&\le \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)}{T}\right\|\left\|\mathbf{G}_x\mathbf{G}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&+ \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0\right)}{T}\right\|\left\|\mathbf{G}_x\mathbf{G}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= \sqrt{N}O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

by (B.8) and (B.11).

$$
\begin{aligned}
|\mathbf{e}_2| &\le \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)}{T}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T}\right\| \\
&\le \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)}{T}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\boldsymbol{\varepsilon}_i}{T}\right\| \\
&+ \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)}{T}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{F}_y^0}{T}\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

by (B.6), (B.7) and (B.8).

$$
\begin{aligned}
|\mathbf{e}_3| &\le \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_x^0}{\sqrt{T}}\right\|\|\mathbf{G}_x\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{T}\right\| \\
&\le \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_x^0}{\sqrt{T}}\right\|\|\mathbf{G}_x\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\boldsymbol{\varepsilon}_i}{T}\right\| \\
&+ \sqrt{\frac{N}{T}}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_x^0}{\sqrt{T}}\right\|\|\mathbf{G}_x\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\mathbf{F}_y^0}{T}\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

by (B.6) and (B.7).

$$
\begin{aligned}
|\mathbf{e}_4| &\le \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\mathbf{G}_x\mathbf{G}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&\le \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\mathbf{G}_x\mathbf{G}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&+ \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\mathbf{G}_x\mathbf{G}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\| \\
&= \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

by using (B.11). And so, $\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\left(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0}\right)\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i = o_p(1)$.

It remains to consider the second term in (B.34). By using $\mathbf{M}_{\hat{F}_y} - \mathbf{M}_{F_y^0} = -\left(\frac{\hat{\mathbf{F}}_y\hat{\mathbf{F}}_y'}{T} - \mathbf{P}_{F_y^0}\right)$, and adding and subtracting terms, we get

$$
\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\left(\mathbf{M}_{\hat{F}_y} - \mathbf{M}_{F_y^0}\right)\mathbf{u}_i \\
&= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\left(\frac{\hat{\mathbf{F}}_y\hat{\mathbf{F}}_y'}{T} - \mathbf{P}_{F_y^0}\right)\mathbf{u}_i \\
&= -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0\mathbf{G}_y\right)}{T}\mathbf{G}_y'\mathbf{F}_y^{0\prime}\mathbf{u}_i \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0\mathbf{G}_y\right)}{T}\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{F}_y^0}{T}\mathbf{G}_y\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{F}_y^0}{T}\left[\mathbf{G}_y\mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right]\mathbf{F}_y^{0\prime}\mathbf{u}_i \\
&= -\left(\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 + \mathbf{e}_4\right).
\end{aligned}
$$

We start with $\mathbf{e}_2$, which is

$$
\begin{aligned}
|\mathbf{e}_2| &\le \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{u}_i\right\| \\
&\le \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{u}_i}{T}\right\| \\
&+ \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i}{\sqrt{T}}\right\|\left\|\frac{\hat{\mathbf{F}}_x}{\sqrt{T}}\right\|\left\|\frac{\hat{\mathbf{F}}_x'\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)'\mathbf{u}_i}{T}\right\| \\
&= \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

by using (B.4), (B.8) and $\left\|\frac{(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x)'\mathbf{u}_i}{T}\right\| = O_p\left(\delta_{NT}^{-2}\right)$.

$$
\begin{aligned}
\mathbf{e}_3 &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{F}_y^0}{T}\mathbf{G}_y\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_y^0}{T}\mathbf{G}_y\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\hat{\mathbf{F}}_x}{T}\frac{\hat{\mathbf{F}}_x'\mathbf{F}_y^0}{T}\mathbf{G}_y\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_y^0}{T}\mathbf{G}_y\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_x^0\mathbf{G}_x}{T}\frac{\hat{\mathbf{F}}_x'\mathbf{F}_y^0}{T}\mathbf{G}_y\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i \\
&\quad -\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\right)}{T}\frac{\hat{\mathbf{F}}_x'\mathbf{F}_y^0}{T}\mathbf{G}_y\left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i \\
&= \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3,
\end{aligned}
$$

where

$$|\mathbf{a}_1| \leq \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_y^0}{\sqrt{T}}\right\|\|\mathbf{G}_y\|\left\|\frac{\left(\hat{\mathbf{F}}_y-\mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i}{T}\right\| = \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),$$

$$|\mathbf{a}_2| \leq \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_x^0}{\sqrt{T}}\right\|\|\mathbf{G}_x\|\left\|\frac{\hat{\mathbf{F}}_x}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\|\mathbf{G}_y\|\left\|\frac{\left(\hat{\mathbf{F}}_y-\mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i}{T}\right\| = \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),$$

$$|\mathbf{a}_3| \leq \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)}{T}\right\|\left\|\frac{\hat{\mathbf{F}}_x}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\|\mathbf{G}_y\|\left\|\frac{\left(\hat{\mathbf{F}}_y-\mathbf{F}_y^0\mathbf{G}_y\right)'\mathbf{u}_i}{T}\right\| = \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right),$$

which hold by (B.8) and $\left\|\frac{(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{u}_i}{T}\right\| = O_p\left(\delta_{NT}^{-2}\right)$ as shown above.

$$\begin{aligned}
\mathbf{e}_4 &= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{F}_y^0}{T}\left[\mathbf{G}_y\mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right]\mathbf{F}_y^{0\prime}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_y^0}{T}\left[\mathbf{G}_y\mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right]\mathbf{F}_y^{0\prime}\mathbf{u}_i \\
&- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\hat{\mathbf{F}}_x}{T}\frac{\hat{\mathbf{F}}_x'\mathbf{F}_y^0}{T}\left[\mathbf{G}_y\mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right]\mathbf{F}_y^{0\prime}\mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_y^0}{T}\left[\mathbf{G}_y\mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right]\mathbf{F}_y^{0\prime}\mathbf{u}_i \\
&- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_x^0\mathbf{G}_x}{T}\frac{\hat{\mathbf{F}}_x'\mathbf{F}_y^0}{T}\left[\mathbf{G}_y\mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right]\mathbf{F}_y^{0\prime}\mathbf{u}_i \\
&- \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)}{T}\frac{\hat{\mathbf{F}}_x'\mathbf{F}_y^0}{T}\left[\mathbf{G}_y\mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right]\mathbf{F}_y^{0\prime}\mathbf{u}_i \\
&= \mathbf{a}_1+\mathbf{a}_2+\mathbf{a}_3,
\end{aligned}$$

$$|\mathbf{a}_1| \leq \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\mathbf{G}_y\mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\mathbf{u}_i}{T}\right\| = \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),$$

$$|\mathbf{a}_2| \leq \sqrt{N}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_x^0}{\sqrt{T}}\right\|\|\mathbf{G}_x\|\left\|\frac{\hat{\mathbf{F}}_x}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\mathbf{G}_y\mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\mathbf{u}_i}{T}\right\| = \sqrt{N}O_p\left(\delta_{NT}^{-2}\right),$$

$$|\mathbf{a}_3| \leq \sqrt{NT}\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x\right)}{T}\right\|\left\|\frac{\hat{\mathbf{F}}_x}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\mathbf{G}_y\mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\mathbf{u}_i}{T}\right\| = \sqrt{NT}O_p\left(\delta_{NT}^{-4}\right),$$

due to (B.8) and (B.21). Now consider $\mathbf{e}_1$ which is

$$
\begin{aligned}
\mathbf{e}_1 &= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \mathbf{M}_{\hat{F}_x} \left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y \right)}{T} \mathbf{G}_y' \mathbf{F}_y^{0\prime} \mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y \right)}{T} \mathbf{G}_y' \mathbf{F}_y^{0\prime} \mathbf{u}_i \\
&- \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \hat{\mathbf{F}}_x}{T} \frac{\hat{\mathbf{F}}_x' \left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y \right)}{T} \mathbf{G}_y' \mathbf{F}_y^{0\prime} \mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y \right)}{T} \mathbf{G}_y' \mathbf{F}_y^{0\prime} \mathbf{u}_i \\
&- \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \mathbf{F}_x^0 \mathbf{G}_x}{T} \frac{\hat{\mathbf{F}}_x' \left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y \right)}{T} \mathbf{G}_y' \mathbf{F}_y^{0\prime} \mathbf{u}_i \\
&- \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)}{T} \frac{\hat{\mathbf{F}}_x' \left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y \right)}{T} \mathbf{G}_y' \mathbf{F}_y^{0\prime} \mathbf{u}_i \\
&= \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_3.
\end{aligned}
$$

We begin with $\mathbf{a}_2$, which is given by

$$
\begin{aligned}
|\mathbf{a}_2| &\leq \sqrt{N} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{\mathbf{V}_i' \mathbf{F}_x^0}{\sqrt{T}} \right\| \|\mathbf{G}_x\| \left\| \frac{\hat{\mathbf{F}}_x' \left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y \right)}{T} \right\| \|\mathbf{G}_y\| \left\| \frac{\mathbf{F}_y^{0\prime} \mathbf{u}_i}{T} \right\| \\
&= \sqrt{N} O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{a}_3| &\leq \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_x - \mathbf{F}_x^0 \mathbf{G}_x \right)}{T} \right\| \left\| \frac{\hat{\mathbf{F}}_x' \left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y \right)}{T} \right\| \|\mathbf{G}_y\| \left\| \frac{\mathbf{F}_y^{0\prime} \mathbf{u}_i}{T} \right\| \\
&= \sqrt{NT} O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

which holds by (B.8) and (B.14). Now consider $\mathbf{a}_1$. By adding and subtracting we get

$$
\begin{aligned}
\mathbf{a}_1 &= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_y - \mathbf{F}_y^0 \mathbf{G}_y \right)}{T} \mathbf{G}_y' \mathbf{F}_y^{0\prime} \mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_y \mathbf{G}_y^{-1} - \mathbf{F}_y^0 \right)}{T} \mathbf{G}_y \mathbf{G}_y' \mathbf{F}_y^{0\prime} \mathbf{u}_i \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_y \mathbf{G}_y^{-1} - \mathbf{F}_y^0 \right)}{T} \left(\frac{\mathbf{F}_y^{0\prime} \mathbf{F}_y^0}{T} \right)^{-1} \mathbf{F}_y^{0\prime} \mathbf{u}_i \\
&+ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_y \mathbf{G}_y^{-1} - \mathbf{F}_y^0 \right)}{T} \left[\mathbf{G}_y \mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime} \mathbf{F}_y^0}{T} \right)^{-1} \right] \mathbf{F}_y^{0\prime} \mathbf{u}_i \\
&= \mathbf{b}_1 + \mathbf{b}_2.
\end{aligned}
$$

Consider next $\mathbf{b}_2$, which is

$$
\begin{aligned}
|\mathbf{b}_2| &\leq \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{\mathbf{V}_i' \left(\hat{\mathbf{F}}_y \mathbf{G}_y^{-1} - \mathbf{F}_y^0 \right)}{T} \right\| \left\| \mathbf{G}_y \mathbf{G}_y' - \left(\frac{\mathbf{F}_y^{0\prime} \mathbf{F}_y^0}{T} \right)^{-1} \right\| \left\| \frac{\mathbf{F}_y^{0\prime} \mathbf{u}_i}{T} \right\| \\
&= \sqrt{NT} O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

by (B.18) and (B.21). Next is $\mathbf{b}_1$. In line with the discussion in Bai (2009a, p.1266), we have

$$\begin{aligned}\hat{\mathbf{F}}_y\mathbf{G}_y^{-1}-\mathbf{F}_y^0 &= \frac{1}{NT}\sum_{i=1}^{N}\mathbf{W}_i\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)'\mathbf{W}_i'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y \\ &+\frac{1}{NT}\sum_{i=1}^{N}\mathbf{W}_i\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)\mathbf{u}_i'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y+\frac{1}{NT}\sum_{i=1}^{N}\mathbf{u}_i\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)'\mathbf{W}_i'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y \\ &+\frac{1}{NT}\sum_{i=1}^{N}\mathbf{F}_y^0\boldsymbol{\gamma}_{yi}^0\boldsymbol{\varepsilon}_i'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y+\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\boldsymbol{\gamma}_{\ell i}^{0\prime}\mathbf{F}_y^{0\prime}\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y+\frac{1}{NT}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y. \qquad \text{(B.35)}\end{aligned}$$

with $\mathbf{G}_y=\left(\boldsymbol{\Xi}_{yNT}\hat{\mathbf{Q}}_y\right)^{-1}$ where $\boldsymbol{\Xi}_{yNT}$ is assumed to be an invertible matrix, $\hat{\mathbf{Q}}_y=\left(\boldsymbol{\Upsilon}_{yN}^0\boldsymbol{\Lambda}_{0\hat{F}_y}\right)^{-1}$ and $\boldsymbol{\Lambda}_{0\hat{F}_y}=T^{-1}\mathbf{F}_y^{0\prime}\hat{\mathbf{F}}_y$ and $\boldsymbol{\Upsilon}_{yN}^0=\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\gamma}_{yi}^0\boldsymbol{\gamma}_{yi}^{0\prime}$.

By making use of (B.35), this term can be written as follows

$$\begin{aligned}\mathbf{b}_1 &= \frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{W}_j}{T}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)'\mathbf{W}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i \\ &+\frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{i=1}^{N}\frac{\mathbf{V}_i'\mathbf{W}_j}{T}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)\mathbf{u}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i \\ &+\frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{u}_j}{T}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)'\mathbf{W}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i \\ &+\frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_y^0}{T}\boldsymbol{\gamma}_{yj}^0\boldsymbol{\varepsilon}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i \\ &+\frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\boldsymbol{\varepsilon}_j}{T}\boldsymbol{\gamma}_{yj}^{0\prime}\mathbf{F}_y^{0\prime}\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i \\ &+\frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\boldsymbol{\varepsilon}_j}{T}\boldsymbol{\varepsilon}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i \\ &=\mathbf{c}_1+\mathbf{c}_2+\mathbf{c}_3+\mathbf{c}_4+\mathbf{c}_5+\mathbf{c}_6. \qquad \text{(B.36)}\end{aligned}$$

$$\begin{aligned}|\mathbf{c}_1| &= \left\|\frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{W}_j}{T}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)'\mathbf{W}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i\right\| \\ &\leq \frac{1}{\sqrt{NT}}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i}{\sqrt{T}}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\|\right)\left(\frac{1}{N}\sum_{j=1}^{N}\left\|\frac{\mathbf{W}_j}{\sqrt{T}}\right\|^2\right)\left\|\sqrt{NT}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)\right\|^2\left\|\frac{\hat{\mathbf{F}}_y}{\sqrt{T}}\right\| \\ &\times\left\|\hat{\mathbf{Q}}_y\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\| \\ &= O_p\left(N^{-1/2}T^{-1/2}\right),\end{aligned}$$

where the order is determined by making use of $\sqrt{NT}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV1}\right)=O_p(1)$, which holds by Proposition 3.

$$
\begin{aligned}
|\mathbf{c}_2| &= \left\| \frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{W}_j}{T}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)\mathbf{u}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i \right\| \\
&\leq \left\| \frac{1}{N}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{W}_j}{T}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)\mathbf{u}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\gamma_{yi}^0 \right\| \\
&+ \left\| \frac{1}{NT}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{W}_j}{T}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)\mathbf{u}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \\
&= \left\| \frac{1}{N}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{u}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\gamma_{yi}^0\frac{\mathbf{V}_i'\mathbf{W}_j}{T}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right) \right\| \\
&+ \left\| \frac{1}{NT}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{W}_j}{T}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV}\right)\mathbf{u}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}} \right\| \\
&\leq \frac{1}{\sqrt{N}}\left\| \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\gamma_{yi}^0\mathbf{V}_i \right\| \left(\frac{1}{N}\sum_{j=1}^{N}\left\|\frac{\mathbf{W}_j}{\sqrt{T}}\right\|\left\|\frac{\mathbf{u}_j}{\sqrt{T}}\right\|\right)\left\|\sqrt{NT}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV1}\right)\right\|\left\|\frac{\hat{\mathbf{F}}_y}{\sqrt{T}}\right\| \\
&\times \left\|\hat{\mathbf{Q}}_y\right\| \\
&+ \frac{1}{\sqrt{T}}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right\|\right)\left(\frac{1}{N}\sum_{j=1}^{N}\left\|\frac{\mathbf{W}_j}{\sqrt{T}}\right\|\left\|\frac{\mathbf{u}_j}{\sqrt{T}}\right\|\right)\left\|\sqrt{NT}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV1}\right)\right\|\left\|\frac{\hat{\mathbf{F}}_y}{\sqrt{T}}\right\| \\
&\times \left\|\hat{\mathbf{Q}}_y\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\| \\
&= O_p\left(N^{-1/2}\right)+O_p\left(T^{-1/2}\right)=O_p\left(\delta_{NT}^{-1}\right),
\end{aligned}
$$

where the second equality holds because $\mathbf{u}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\boldsymbol{\gamma}_{yi}^0$ is a scalar, whereas the order is determined by using again $\sqrt{NT}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV1}\right)=O_p(1)$ by Proposition 3. As for $\mathbf{c}_3$, we have

$$
\begin{aligned}
|\mathbf{c}_3| &\leq \frac{1}{\sqrt{T}}\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{N}\sum_{j=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{u}_j}{\sqrt{T}}\right\|\left\|\frac{\mathbf{W}_j}{\sqrt{T}}\right\|\right)\left\|\sqrt{NT}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{IV1}\right)\right\|\left\|\frac{\hat{\mathbf{F}}_y}{\sqrt{T}}\right\|\left\|\hat{\mathbf{Q}}_y\right\| \\
&\times \left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\| \\
&= O_p\left(T^{-1/2}\right),
\end{aligned}
$$

by similar arguments as above.

$$
\begin{aligned}
|\mathbf{c}_4| &= \left\| \frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_y^0}{T}\gamma_{yj}^0\boldsymbol{\varepsilon}_j'\hat{\mathbf{F}}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i\right\| \\
&\leq \left\| \frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_y^0}{T}\gamma_{yj}^0\boldsymbol{\varepsilon}_j'\mathbf{F}_y^0\mathbf{G}_y\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i\right\| \\
&+ \left\| \frac{1}{NT}\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{\mathbf{V}_i'\mathbf{F}_y^0}{T}\gamma_{yj}^0\boldsymbol{\varepsilon}_j'\left(\hat{\mathbf{F}}_y-\mathbf{F}_y^0\mathbf{G}_y\right)\hat{\mathbf{Q}}_y\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\mathbf{F}_y^{0\prime}\mathbf{u}_i\right\| \\
&\leq \frac{1}{\sqrt{T}}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\mathbf{u}_i}{T}\right\|\right)\left\|\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\gamma_{yj}^0\frac{\boldsymbol{\varepsilon}_j'\mathbf{F}_y^0}{\sqrt{T}}\right\|\|\mathbf{G}_y\|\left\|\hat{\mathbf{Q}}_y\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\| \\
&+ \left(\frac{1}{N}\sum_{i=1}^{N}\left\|\frac{\mathbf{V}_i'\mathbf{F}_y^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_y^{0\prime}\mathbf{u}_i}{T}\right\|\right)\left\|\frac{1}{\sqrt{N}}\sum_{j=1}^{N}\gamma_{yj}^0\frac{\boldsymbol{\varepsilon}_j'\left(\hat{\mathbf{F}}_y-\mathbf{F}_y^0\mathbf{G}_y\right)}{T}\right\|\left\|\hat{\mathbf{Q}}_y\right\|\left\|\left(\frac{\mathbf{F}_y^{0\prime}\mathbf{F}_y^0}{T}\right)^{-1}\right\| \\
&= O_p\left(T^{-1/2}\right)+O_p\left(\delta_{NT}^{-2}\right)+O_p\left(N^{-1/2}\right)=O_p\left(\delta_{NT}^{-1}\right),
\end{aligned}
$$

by (B.20). Finally, by similar arguments, we can show that $\mathbf{c}_5 = O_p\left(\delta_{NT}^{-1}\right)$ and also $\mathbf{c}_6 = O_p\left(T^{-1/2}\right) + O_p\left(\delta_{NT}^{-2}\right)$. Thus, we have

$$
\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_y}\mathbf{u}_i = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\mathbf{V}_i'\mathbf{M}_{F_x^0}\mathbf{M}_{F_y^0}\mathbf{u}_i + o_p(1) \tag{B.37}
$$

The result in (A.18) can be shown as above. ■

Proof of Lemma 10. We begin with (A.28). By using $\hat{\mathbf{F}}_x'\hat{\mathbf{F}}_x/T = \mathbf{I}_{m_x}$, we have $\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F_x^0} = \mathbf{P}_{F_x^0} - \mathbf{P}_{\hat{F}_x} = -\left(\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x'}{T} - \mathbf{P}_{F_x^0}\right)$. And so, we get

$$
\begin{aligned}
& T^{-1/2}\mathbf{X}_i'(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0})\mathbf{u}_i \\
&= T^{-1/2}\mathbf{X}_i'(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0})\mathbf{u}_i \\
&= -\frac{1}{\sqrt{T}}\mathbf{X}_i'\left(\frac{\hat{\mathbf{F}}_x\hat{\mathbf{F}}_x'}{T}-\mathbf{P}_{F_x^0}\right)\mathbf{u}_i \\
&= -\frac{1}{\sqrt{T}}\frac{\mathbf{X}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\mathbf{G}'\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&- \frac{1}{\sqrt{T}}\frac{\mathbf{X}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\mathbf{u}_i \\
&- \frac{1}{\sqrt{T}}\frac{\mathbf{X}_i'\mathbf{F}_x^0}{T}\mathbf{G}\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\mathbf{u}_i \\
&- \frac{1}{\sqrt{T}}\frac{\mathbf{X}_i'\mathbf{F}_x^0}{T}\left[\mathbf{G}\mathbf{G}'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right]\mathbf{F}_x^{0\prime}\mathbf{u}_i \\
&= -\left(\mathbf{a}_1+\mathbf{a}_2+\mathbf{a}_3+\mathbf{a}_4\right),
\end{aligned}
$$

$$
\begin{aligned}
|\mathbf{a}_1| &\le \sqrt{T}\left\|\frac{\mathbf{X}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\|\|\mathbf{G}\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T}\right\| \\
&\le \sqrt{T}\|\boldsymbol{\Gamma}_{xi}\|\left\|\frac{\mathbf{F}_x'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\|\|\mathbf{G}\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\| \\
&+ \sqrt{T}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\|\|\mathbf{G}\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\| \\
&= \sqrt{T}O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

from $\left\|\frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\| = O_p\left(\delta_{NT}^{-2}\right)$ and $\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\| = O_p\left(\delta_{NT}^{-2}\right)$ by Lemma B.4, $\|\mathbf{G}\| = O_p(1)$, $\|\boldsymbol{\Gamma}_{xi}\| = O_p(1)$ by Assumption 4, $\frac{\|\mathbf{F}_x^0\|}{\sqrt{T}} = O_p(1)$, and $\frac{\|\mathbf{u}_i\|}{\sqrt{T}} \le \|\boldsymbol{\gamma}_i^0\|\frac{\|\mathbf{F}_y^0\|}{\sqrt{T}} + \|\boldsymbol{\lambda}_i^0\|\frac{\|\mathbf{F}_y^0\|}{\sqrt{T}} + \frac{\|\boldsymbol{\varepsilon}_i\|}{\sqrt{T}} = O_p(1)$ by Assumptions 1, 3 and 4.

$$
\begin{aligned}
|\mathbf{a}_2| &\le \sqrt{T}\left\|\frac{\mathbf{X}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\mathbf{u}_i}{T}\right\| \\
&\le \sqrt{T}\|\boldsymbol{\Gamma}_{xi}\|\left\|\frac{\mathbf{F}_x'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\mathbf{u}_i}{T}\right\| \\
&+ \sqrt{T}\left\|\frac{\mathbf{V}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\mathbf{u}_i}{T}\right\| \\
&= \sqrt{T}O_p\left(\delta_{NT}^{-4}\right),
\end{aligned}
$$

by similar arguments as above and $\left\|\frac{\mathbf{u}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\| \le \|\boldsymbol{\gamma}_i^0\|\left\|\frac{\mathbf{F}_x^{0\prime}\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\| + \|\boldsymbol{\lambda}_i^0\|\left\|\frac{\mathbf{F}_y^{0\prime}\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\| + \left\|\frac{\boldsymbol{\varepsilon}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\| = O_p\left(\delta_{NT}^{-2}\right)$ by Lemma B.4 and Assumption 4.

$$
\begin{aligned}
|\mathbf{a}_3| &\le \sqrt{T}\left\|\frac{\mathbf{X}_i'\mathbf{F}_x^0}{T}\right\|\|\mathbf{G}\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\mathbf{u}_i}{T}\right\| \\
&\le \sqrt{T}\|\boldsymbol{\Gamma}_{xi}\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|^2\|\mathbf{G}\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\mathbf{u}_i}{T}\right\| \\
&+ \sqrt{T}\left\|\frac{\mathbf{V}_i}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\|\mathbf{G}\|\left\|\frac{\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)'\mathbf{u}_i}{T}\right\| \\
&= \sqrt{T}O_p\left(\delta_{NT}^{-2}\right),
\end{aligned}
$$

from $\left\|\frac{\mathbf{u}_i'\left(\hat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}\right)}{T}\right\| = O_p\left(\delta_{NT}^{-2}\right)$ as shown above by using Lemma B.4 and Assumption 4, $\|\mathbf{G}_x\| = O_p(1)$, $\|\boldsymbol{\Gamma}_{xi}\| = O_p(1)$ by Assumption 4 as well as $\frac{\|\mathbf{F}_x^0\|}{\sqrt{T}} = O_p(1)$ by Assumption 3.

$$
\begin{aligned}
|\mathbf{a}_4| &\leq \sqrt{T}\left\|\frac{\mathbf{X}_i'\mathbf{F}_x^0}{T}\right\|\left\|\mathbf{G}\mathbf{G}'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{F}_x^{0\prime}\mathbf{u}_i}{T}\right\| \\
&\leq \sqrt{T}\left\|\mathbf{\Gamma}_{xi}\right\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|^2\left\|\mathbf{G}\mathbf{G}'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\| \\
&+ \sqrt{T}\left\|\frac{\mathbf{V}_i}{\sqrt{T}}\right\|\left\|\frac{\mathbf{F}_x^0}{\sqrt{T}}\right\|\left\|\mathbf{G}\mathbf{G}'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right\|\left\|\frac{\mathbf{u}_i}{\sqrt{T}}\right\| \\
&= \sqrt{T}O_p\left(\delta_{NT}^{-2}\right).
\end{aligned}
$$

from making use of $\left\|\mathbf{G}\mathbf{G}'-\left(T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\right)^{-1}\right\| = O_p\left(\delta_{NT}^{-2}\right)$ by Lemma B.4, $\frac{\|\mathbf{V}_i\|}{\sqrt{T}} = O_p(1)$ by Assumption 2, $\frac{\|\mathbf{F}_x^0\|}{\sqrt{T}} = O_p(1)$ by Assumption 3, $\|\mathbf{\Gamma}_{xi}\| = O_p(1)$ by Assumption 4 and $\frac{\|\mathbf{u}_i\|}{\sqrt{T}} = O_p(1)$ by Assumptions 1, 3 and 4.

By putting the results together, we therefore have

$$
\left\|T^{-1/2}\mathbf{X}_i'(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0})\mathbf{u}_i\right\| = \sqrt{T}O_p\left(\delta_{NT}^{-2}\right). \tag{B.38}
$$

Thus,

$$
T^{-1/2}\mathbf{X}_i'\mathbf{M}_{\hat{F}_x}\mathbf{u}_i = T^{-1/2}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{u}_i + \sqrt{T}O_p\left(\delta_{NT}^{-2}\right). \tag{B.39}
$$

The results in (A.29) and (A.30) can be shown in a similar manner. ■

Proof of Lemma 11. Using the identity $\hat{\mathbf{u}}_i = \mathbf{u}_i - \mathbf{W}_i\left(\hat{\boldsymbol{\theta}}_{IV}-\boldsymbol{\theta}\right)$ we have

$$
\begin{aligned}
\frac{1}{NT}\sum_{i=1}^N \hat{\boldsymbol{\xi}}_{\hat{F}iT}\hat{\boldsymbol{\xi}}_{\hat{F}iT}' &= \frac{1}{NT}\sum_{i=1}^N \boldsymbol{\xi}_{\hat{F}iT}\boldsymbol{\xi}_{\hat{F}iT}' \\
&- \frac{1}{NT}\sum_{i=1}^N \boldsymbol{\xi}_{\hat{F}iT}\left(\hat{\boldsymbol{\theta}}_{IV}-\boldsymbol{\theta}\right)'\mathbf{W}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{Z}}_i \\
&- \frac{1}{NT}\sum_{i=1}^N \hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{W}_i\left(\hat{\boldsymbol{\theta}}_{IV}-\boldsymbol{\theta}\right)\boldsymbol{\xi}_{\hat{F}iT}' \\
&+ \frac{1}{NT}\sum_{i=1}^N \hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{W}_i\left(\hat{\boldsymbol{\theta}}_{IV}-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}_{IV}-\boldsymbol{\theta}\right)'\mathbf{W}_i'\mathbf{M}_{\hat{F}_y}\hat{\mathbf{Z}}_i \\
&= \frac{1}{NT}\sum_{i=1}^N \boldsymbol{\xi}_{\hat{F}iT}\boldsymbol{\xi}_{\hat{F}iT}' - \mathbf{E}_1 - \mathbf{E}_2 + \mathbf{E}_3.
\end{aligned}
$$

We have

$$
\|\mathbf{E}_1\| \leq \sqrt{T}\left\|\hat{\boldsymbol{\theta}}_{IV}-\boldsymbol{\theta}\right\|\frac{1}{N}\sum_{i=1}^N\left\|\frac{\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{W}_i}{T}\right\|\frac{\left\|\boldsymbol{\xi}_{\hat{F}iT}'\right\|}{\sqrt{T}} = O_p\left(\frac{1}{\sqrt{N}}\right),
$$

since $\left\|\frac{\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{W}_i}{T}\right\| \leq \left\|\frac{\hat{\mathbf{Z}}_i'\mathbf{W}_i}{T}\right\| + \left\|\frac{\hat{\mathbf{Z}}_i'\hat{\mathbf{F}}_y}{T}\right\|\left\|\frac{\hat{\mathbf{F}}_y'\mathbf{W}_i}{T}\right\| = O_p(1)$. Similarly, $\|\mathbf{E}_2\| = O_p\left(\frac{1}{\sqrt{N}}\right)$. Also

$$
\|\mathbf{E}_3\| \leq T\left\|\hat{\boldsymbol{\theta}}_{IV}-\boldsymbol{\theta}\right\|^2\frac{1}{N}\sum_{i=1}^N\left\|\frac{\hat{\mathbf{Z}}_i'\mathbf{M}_{\hat{F}_y}\mathbf{W}_i}{T}\right\|^2 = O_p\left(\frac{1}{N}\right).
$$

Thus,

$$
\frac{1}{NT}\sum_{i=1}^N \hat{\boldsymbol{\xi}}_{\hat{F}iT}\hat{\boldsymbol{\xi}}_{\hat{F}iT}' = \frac{1}{NT}\sum_{i=1}^N \boldsymbol{\xi}_{\hat{F}iT}\boldsymbol{\xi}_{\hat{F}iT}' + o_p(1).
$$

as required. ■

Proof of Lemma 12. By Lemma 11 and Proposition 2 we have

$$
\frac{1}{NT}\sum_{i=1}^N \boldsymbol{\xi}_{\hat{F}iT}\boldsymbol{\xi}_{\hat{F}iT}' = \frac{1}{NT}\sum_{i=1}^N \mathbf{Z}_i'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_i + o_p(1).
$$

Noting that $E\left(\mathbf{Z}_j'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_i\right) = \mathbf{0}$ for all $i \neq j$ and using Lemma B.1, $\frac{1}{NT}\sum_{i=1}^N \mathbf{Z}_i'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_i \xrightarrow{p}$ $\lim_{N,T\to\infty}\frac{1}{NT}\sum_{i=1}^N E\left(\mathbf{Z}_i'\mathbf{M}_{F_y^0}\boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i'\mathbf{M}_{F_y^0}\mathbf{Z}_i\right)$, which yields $\frac{1}{NT}\sum_{i=1}^N \boldsymbol{\xi}_{\hat{F}iT}\boldsymbol{\xi}_{\hat{F}iT}' - \boldsymbol{\Omega} = o_p(1)$ when $(N,T)\to\infty$ jointly, as required.
■

Proof of Lemma 13. First of all, by Lemma 2, $\hat{\mathbf{A}}_{NT} - \frac{1}{N}\sum_{i=1}^N \mathbf{A}_{i,T} = o_p(1)$ and $\hat{\mathbf{B}}_{NT} - \frac{1}{N}\sum_{i=1}^N \mathbf{B}_{i,T} = o_p(1)$, then applying Lemma B.1 yields the required results. ■

Proof of Lemma 14. See Van Der Vaart and Wellner (1996). ■

Lemma B.7 *Under Assumptions 2 to 4, and Assumption 7, we have*

$$(a)\quad \sup_{1\le i\le N}\|T^{-1}\mathbf{V}_i'(\mathbf{F}_x^0 - \widehat{\mathbf{F}}_x\mathbf{G}_x^{-1})\| = O_p(N^{1/4}\delta_{NT}^{-2}) + O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})$$

$$(b)\quad \sup_{1\le i\le N}\|T^{-1}\mathbf{V}_{i,-1}'(\mathbf{F}_{x,-1}^0 - \widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\| = O_p(N^{1/4}\delta_{NT}^{-2}) + O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})$$

$$(c)\quad \sup_{1\le i\le N}\|T^{-1}\mathbf{V}_i'(\mathbf{F}_{x,-1}^0 - \widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\| = O_p(N^{1/4}\delta_{NT}^{-2}) + O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})$$

$$(d)\quad \sup_{1\le i\le N}\|T^{-1}\mathbf{V}_{i,-1}'(\mathbf{F}_x^0 - \widehat{\mathbf{F}}_x\mathbf{G}_x^{-1})\| = O_p(N^{1/4}\delta_{NT}^{-2}) + O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})$$

Proof of Lemma B.7. Consider (a). With the equation (B.28), we have

$$\begin{aligned}
&\sup_{1\le i\le N}\|T^{-1}\mathbf{V}_i'(\mathbf{F}_x^0 - \widehat{\mathbf{F}}_x\mathbf{G}_x^{-1})\| \\
\le &\sup_{1\le i\le N} N^{-1}T^{-2}\|\sum_{\ell=1}^k\sum_{j=1}^N \mathbf{V}_i'\mathbf{F}_x^0\boldsymbol{\gamma}_{\ell j}^0\mathbf{v}_{\ell j}'\widehat{\mathbf{F}}_x\|\|\hat{\mathbf{Q}}_x\| + \sup_{1\le i\le N} N^{-1}T^{-2}\|\sum_{\ell=1}^k\sum_{j=1}^N \mathbf{V}_i'\mathbf{v}_{\ell j}\boldsymbol{\gamma}_{\ell j}^{0\prime}\mathbf{F}_x^{0\prime}\widehat{\mathbf{F}}_x\|\|\hat{\mathbf{Q}}_x\| \\
&+ \sup_{1\le i\le N} N^{-1}T^{-2}\|\sum_{\ell=1}^k\sum_{j=1}^N \mathbf{V}_i'\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'\widehat{\mathbf{F}}_x\|\|\hat{\mathbf{Q}}_x\|
\end{aligned}$$

Since $\hat{\mathbf{Q}}_x = O_p(1)$, we omit $\hat{\mathbf{Q}}_x$ in the following analysis. The first term is bounded in norm by

$$T^{-1/2}\cdot\sup_{1\le i\le N}\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|\cdot\left\|N^{-1}T^{-1}\sum_{\ell=1}^k\sum_{j=1}^N\boldsymbol{\gamma}_{\ell j}^0\mathbf{v}_{\ell j}'\widehat{\mathbf{F}}_x\right\|$$

Since $E\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|^4 \le \Delta$ by Assumption 7 (ii), we have

$$\sup_{1\le i\le N}\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\| = O_p(N^{1/4}) \tag{B.40}$$

Note that $N^{-1}T^{-1}\sum_{\ell=1}^k\sum_{j=1}^N\boldsymbol{\gamma}_{\ell j}^0\mathbf{v}_{\ell j}'\widehat{\mathbf{F}}_x = \mathbf{A}_{kNT}' = O_p(N^{-1/2}T^{-1/2}) + O_p(N^{-1}) + O_p(N^{-1/2}\delta_{NT}^{-2})$ as shown in the proof of Lemma 4, and with (B.40), the first term is $O_p(N^{-1/4}T^{-1}) + O_p(N^{-3/4}T^{-1/2}) + O_p(N^{-1/4}T^{-1/2}\delta_{NT}^{-2})$. The second term is bounded in norm by

$$\begin{aligned}
&\sup_{1\le i\le N} N^{-1}T^{-1}\|\sum_{\ell=1}^N\sum_{\ell=1}^k\sum_{j=1}^N \mathbf{V}_i'\mathbf{v}_{\ell j}\boldsymbol{\gamma}_{\ell j}^{0\prime}\|\|T^{-1/2}\mathbf{F}_x^0\|\|T^{-1/2}\widehat{\mathbf{F}}_x\| \\
=&\sup_{1\le i\le N} N^{-1}T^{-1}\|\sum_{\ell=1}^k\sum_{j=1}^N \mathbf{V}_i'\mathbf{v}_{\ell j}\boldsymbol{\gamma}_{\ell j}^{0\prime}\| \times O_p(1) \\
\le&\sup_{1\le i\le N} N^{-1}T^{-1}\sum_{\ell=1}^k\sum_{j=1}^N\|E(\mathbf{V}_i'\mathbf{v}_{\ell j})\|\|\boldsymbol{\gamma}_{\ell j}^0\| + \sup_{1\le i\le N} N^{-1}T^{-1}\|\sum_{\ell=1}^k\sum_{j=1}^N(\mathbf{V}_i'\mathbf{v}_{\ell j} - E(\mathbf{V}_i'\mathbf{v}_{\ell j}))\boldsymbol{\gamma}_{\ell j}^{0\prime}\| \\
\le& N^{-1}\cdot\sup_{1\le i\le N}\sum_{\ell=1}^k\|T^{-1}E(\mathbf{V}_i'\mathbf{v}_{\ell i})\|\cdot\sup_{1\le\ell\le k,1\le i\le N}\|\boldsymbol{\gamma}_{\ell i}^0\| + \sup_{1\le i\le N} N^{-1}T^{-1}\|\sum_{\ell=1}^k\sum_{j=1}^N(\mathbf{V}_i'\mathbf{v}_{\ell j} - E(\mathbf{V}_i'\mathbf{v}_{\ell j}))\boldsymbol{\gamma}_{\ell j}^{0\prime}\| \\
=&O_p(N^{-3/4}) + O_p(N^{-1/4}T^{-1/2}),
\end{aligned}$$

by Assumption 4 and Assumption 7(iii). Consider the third term. We have

$$
\begin{aligned}
&\sup_{1\le i\le N}\|T^{-1/2}\mathbf{V}_i\|^2=\sup_{1\le i\le N}T^{-1}\sum_{\ell=1}^{k}\sum_{t=1}^{T}v_{\ell it}^2\\
\le&\sup_{1\le i\le N}T^{-1}\sum_{\ell=1}^{k}\sum_{t=1}^{T}Ev_{\ell it}^2+T^{-1/2}\cdot\sup_{1\le i\le N}T^{-1/2}\sum_{\ell=1}^{k}\sum_{t=1}^{T}(v_{\ell it}^2-Ev_{\ell it}^2)\le k\Delta+O_p(N^{1/2}T^{-1/2})\\
=&O_p(1)+O_p(N^{1/2}T^{-1/2})
\end{aligned}
$$

since $E\|T^{-1/2}\sum_{\ell=1}^{k}\sum_{t=1}^{T}(v_{\ell it}^2-Ev_{\ell it}^2)\|^2\le\Delta$ by Assumption 7(iv). With Assumption 2 and 3, we can show that $E\|N^{-1/2}T^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}[\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'-E(\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}')]\mathbf{F}_x^0\|^2\le\Delta$. Thus, the third term is bounded in norm by

$$
\begin{aligned}
&\sup_{1\le i\le N}T^{-1/2}\|\mathbf{V}_i\|\cdot\|N^{-1}T^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'\|\|T^{-1/2}(\mathbf{F}_x^0-\widehat{\mathbf{F}}_x\mathbf{G}_x^{-1})\|\\
&+T^{-1}\cdot\sup_{1\le i\le N}\|N^{-1}T^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{V}_i'E(\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}')\mathbf{F}_x^0\|\cdot\|\mathbf{G}_x^{-1}\|\\
&+N^{-1/2}T^{-1/2}\cdot\sup_{1\le i\le N}T^{-1/2}\|\mathbf{V}_i\|\cdot\|N^{-1/2}T^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}[\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'-E(\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}')]\mathbf{F}_x^0\|\|\mathbf{G}_x^{-1}\|\\
=&O_p(\delta_{NT}^{-2})+O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})+O_p(N^{1/4}T^{-1})
\end{aligned}
$$

where $\|N^{-1}T^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'\|=O_p(\delta_{NT}^{-1})$, which suggests from

$$
\begin{aligned}
&\|N^{-1}T^{-1/2}\sum_{\ell=1}^{k}\sum_{j=1}^{N}E(\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}')\|=T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}\Big|N^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}E\left(v_{\ell js}v_{\ell jt}\right)\Big|^2\\
\le&\Delta N^{-1}T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\big|E\left(v_{\ell js}v_{\ell jt}\right)\big|=\Delta N^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\Big[T^{-1}\sum_{s=1}^{T}\sum_{t=1}^{T}\big|E\left(v_{\ell js}v_{\ell jt}\right)\big|\Big]\le\Delta\,,
\end{aligned}
\tag{B.41}
$$

and

$$
\begin{aligned}
&E\|N^{-1/2}T^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}[\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}'-E(\mathbf{v}_{\ell j}\mathbf{v}_{\ell j}')]\|^2\\
=&T^{-2}\sum_{s=1}^{T}\sum_{t=1}^{T}E\big[N^{-1/2}\sum_{\ell=1}^{k}\sum_{j=1}^{N}(v_{\ell js}v_{\ell jt}-E(v_{\ell js}v_{\ell jt})\big]^2\le\Delta\,,
\end{aligned}
\tag{B.42}
$$

given $|N^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}E\left(v_{\ell js}v_{\ell jt}\right)|\le N^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\sqrt{Ev_{\ell js}^2Ev_{\ell jt}^2}\le\Delta$ and Assumption 2(iv). Collecting the above three terms, the claim (a) holds. Closely follow the proof of (a), we can derive (b), (c) and (d), thus details are omitted. This completes the proof. ■

Proof of Lemma 15. Consider (a). The left hand is bounded in norm by

$$
N^{-1}T^{-1}\sum_{i=1}^{N}\|\mathbf{\Gamma}_{xi}'\mathbf{F}_x^{0\prime}\mathbf{M}_{\widehat{\mathbf{F}}_x}\mathbf{u}_i\|+N^{-1}T^{-1}\sum_{i=1}^{N}\|\mathbf{V}_i'\mathbf{M}_{\widehat{\mathbf{F}}_x^0}\mathbf{u}_i-\mathbf{V}_i'\mathbf{M}_{\mathbf{F}_x}\mathbf{u}_i\|
$$

Consider $N^{-1}T^{-1}\sum_{i=1}^{N}\|\mathbf{\Gamma}_{xi}'\mathbf{F}_x^{0\prime}\mathbf{M}_{\widehat{\mathbf{F}}_x}\mathbf{u}_i\|$, which is bounded by $N^{-1}T^{-1}\sum_{i=1}^{N}\|\mathbf{\Gamma}_{xi}\|\|\mathbf{F}_x^{0\prime}\mathbf{M}_{\widehat{\mathbf{F}}_x}\mathbf{u}_i\|$. Note that $\mathbf{M}_{\mathbf{F}_x^0}\mathbf{F}_x^0=\mathbf{0}$, we have $\mathbf{M}_{\widehat{\mathbf{F}}_x}\mathbf{F}_x^0=(\mathbf{M}_{\widehat{\mathbf{F}}_x}-\mathbf{M}_{\mathbf{F}_x^0})\mathbf{F}_x^0$. We expand $T\mathbf{M}_{\widehat{\mathbf{F}}_x}-T\mathbf{M}_{\mathbf{F}_x^0}$ as following

$$
-(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\mathbf{G}_x'\mathbf{F}_x^{0\prime}-\mathbf{F}_x^0\mathbf{G}_x(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'-(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'-\mathbf{F}_x^0\left(\mathbf{G}_x\mathbf{G}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right)\mathbf{F}_x^{0\prime}\,,
$$

then

$$
\begin{aligned}
&N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{M}_{\widehat{\mathbf{F}}_x}\mathbf{u}_i\| \\
\leq &N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-2}\mathbf{F}_x^{0\prime}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{u}_i\|+N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-2}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\mathbf{G}_x(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{u}_i\| \\
&+N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-2}\mathbf{F}_x^{0\prime}(\widehat{\mathbf{F}}_x-\mathbf{F}\mathbf{G}_x)(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{u}_i\| \\
&+N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-2}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\left(\mathbf{G}_x\mathbf{G}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right)\mathbf{F}_x^{0\prime}\mathbf{u}_i\| \\
= &\mathbb{B}_1+\mathbb{B}_2+\mathbb{B}_3+\mathbb{B}_4\,.
\end{aligned}
$$

Consider $\mathbb{B}_1$. Given $\mathbf{u}_i=\mathbf{F}_y^0\boldsymbol{\gamma}_{yi}+\boldsymbol{\varepsilon}_i$, we have

$$
\begin{aligned}
&N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i\| \\
\leq &N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|\boldsymbol{\gamma}_{yi}\|\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_y^0\|+N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-1}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i\|=O_p(1)
\end{aligned}
$$

Thus, $\mathbb{B}_1$ is bounded in norm by

$$
N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i\|\times\|T^{-1}\mathbf{F}_x^{0\prime}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|\|\mathbf{G}_x\|=O_p(\delta_{NT}^{-2})
$$

Similarly, we can show that $\mathbb{B}_4=O_p(\delta_{NT}^{-2})$.

Consider $\mathbb{B}_2$. The term is bounded in norm by $N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-1}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{u}_i\|\cdot\|T^{-1/2}\mathbf{F}_x^0\|^2\|\mathbf{G}_x\|$, which is $O_p(1)\cdot N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-1}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{u}_i\|$. Furthermore, the term $N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-1}(\widehat{\mathbf{F}}_x-\mathbf{F}\mathbf{G}_x)'\mathbf{u}_i\|$ is bounded in norm by

$$
\begin{aligned}
&N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\boldsymbol{\gamma}_{yi}\|\|T^{-1}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{F}_y^0\|+N^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-1}(\widehat{\mathbf{F}}_x\mathbf{G}_x^{-1}-\mathbf{F}_x^0)'\boldsymbol{\varepsilon}_i\|\|\mathbf{G}_x\| \\
= &O_p(\delta_{NT}^{-2})
\end{aligned}
$$

by Lemmas B.5. Thus, $\mathbb{B}_2=O_p(\delta_{NT}^{-2})$. Analogously, we have $\mathbb{B}_3=O_p(\delta_{NT}^{-4})$. Collecting the above four terms, we have $N^{-1}T^{-1}\sum_{i=1}^{N}\|\boldsymbol{\Gamma}_{xi}'\mathbf{F}_x^{0\prime}\mathbf{M}_{\widehat{\mathbf{F}}_x}\mathbf{u}_i\|=O_p(\delta_{NT}^{-2})$.

Next, we tend to prove that $N^{-1}T^{-1}\sum_{i=1}^{N}\|\mathbf{V}_i'(\mathbf{M}_{\widehat{\mathbf{F}}_x}-\mathbf{M}_{\mathbf{F}_x^0})\mathbf{u}_i\|=O_p(\delta_{NT}^{-2})$. Since $\mathbf{M}_{\widehat{\mathbf{F}}_x}-\mathbf{M}_{\mathbf{F}_x^0}=-T^{-1}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\mathbf{G}_x'\mathbf{F}_x^{0\prime}-T^{-1}\mathbf{F}_x^0\mathbf{G}_x(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'-T^{-1}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'-T^{-1}\mathbf{F}_x^0\left(\mathbf{G}_x\mathbf{G}_x'-\left(T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}\right)^{-1}\right)\mathbf{F}_x^{0\prime}$, we have

$$
\begin{aligned}
&N^{-1}\sum_{i=1}^{N}\|T^{-1}\mathbf{V}_i'(\mathbf{M}_{\widehat{\mathbf{F}}_x}-\mathbf{M}_{\mathbf{F}_x^0})\mathbf{u}_i\| \\
= &N^{-1}\sum_{i=1}^{N}\|T^{-2}\mathbf{V}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{u}_i\|+N^{-1}\sum_{i=1}^{N}\|T^{-2}\mathbf{V}_i'\mathbf{F}_x^0\mathbf{G}_x(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{u}_i\| \\
&+N^{-1}\sum_{i=1}^{N}\|T^{-2}\mathbf{V}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{u}_i\|+N^{-1}\sum_{i=1}^{N}\|T^{-2}\mathbf{V}_i'\mathbf{F}_x^0(\mathbf{G}_x\mathbf{G}_x'-\left(T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}\right)^{-1})\mathbf{F}_x^{0\prime}\mathbf{u}_i\| \\
= &\mathbb{B}_5+\mathbb{B}_6+\mathbb{B}_7+\mathbb{B}_8
\end{aligned}
$$

We bound $\mathbb{B}_5$ in norm by

$$
\begin{aligned}
&N^{-1}\sum_{i=1}^{N}\|T^{-1}\mathbf{V}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{u}_i\|\|\mathbf{G}_x\| \\
\leq &N^{-1}\sum_{i=1}^{N}\|T^{-1}\mathbf{V}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|\|\boldsymbol{\gamma}_{yi}\|\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_y^0\|\|\mathbf{G}_x\|+N^{-1}\sum_{i=1}^{N}\|T^{-1}\mathbf{V}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|\|T^{-1}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_i\|\|\mathbf{G}_x\| \\
= &O_p(\delta_{NT}^{-2})
\end{aligned}
$$

by Lemma B.4. With Lemmas B.4 and B.5, $\mathbb{B}_6$ is bounded in norm by

$$N^{-1}\sum_{i=1}^{N}\|T^{-1}\mathbf{V}_i'\mathbf{F}_x^0\|\|T^{-1}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{u}_i\|\|\mathbf{G}_x\|=O_p(T^{-1/2}\delta_{NT}^{-2})$$

Similarly, we can show that $\mathbb{B}_7=O_p(\delta_{NT}^{-4})$ and $\mathbb{B}_8=O_p(T^{-1/2}\delta_{NT}^{-2})$. With the stochastic orders of the above eight terms, we obtain (a). Following the argument in the proof of (a), we can derive (b).

Consider (c). The term is bounded in norm by

$$\begin{aligned}
&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{X}_j'\mathbf{M}_{\widehat{\mathbf{F}}_x}\mathbf{X}_i-T^{-1}\mathbf{X}_j'\mathbf{M}_{\mathbf{F}_x^0}\mathbf{X}_i\|\\
\le&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{\Gamma}_{xj}'\mathbf{F}_x^{0\prime}\mathbf{M}_{\widehat{\mathbf{F}}_x}\mathbf{F}_x^0\mathbf{\Gamma}_{xi}\|+2\sup_{1\le i,j\le N}\|T^{-1}\mathbf{V}_j'\mathbf{M}_{\widehat{\mathbf{F}}_x}\mathbf{F}_x^0\mathbf{\Gamma}_{xi}\|+\sup_{1\le i,j\le N}\|T^{-1}\mathbf{V}_j'(\mathbf{M}_{\widehat{\mathbf{F}}_x}-\mathbf{M}_{\mathbf{F}_x^0})\mathbf{V}_i\|\\
=&\mathbb{C}_1+\mathbb{C}_2+\mathbb{C}_3
\end{aligned}$$

$\mathbb{C}_1$ is bounded in norm by

$$\begin{aligned}
&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{\Gamma}_{xj}'\mathbf{F}_x^{0\prime}\mathbf{M}_{\widehat{\mathbf{F}}_x}\mathbf{F}_x^0\mathbf{\Gamma}_{xi}\|=\sup_{1\le i,j\le N}\|T^{-1}\mathbf{\Gamma}_{xj}'(\mathbf{F}_x^0-\widehat{\mathbf{F}}_x\mathbf{G}_x^{-1})'\mathbf{M}_{\widehat{\mathbf{F}}_x}(\mathbf{F}_x^0-\widehat{\mathbf{F}}_x\mathbf{G}_x^{-1})\mathbf{\Gamma}_{xi}\|\\
\le&(\sup_{1\le i\le N}\|\mathbf{\Gamma}_{xi}\|)^2\cdot\|T^{-1/2}(\mathbf{F}_x^0-\widehat{\mathbf{F}}_x\mathbf{G}_x^{-1})\|^2=O_p(N^{1/2}\delta_{NT}^{-2})
\end{aligned}$$

Ignoring the scale 2, $\mathbb{C}_2$ is bounded in norm by

$$\begin{aligned}
&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{V}_j'(\mathbf{M}_{\widehat{\mathbf{F}}_x}-\mathbf{M}_{\mathbf{F}})\mathbf{F}_x^0\mathbf{\Gamma}_{xi}\|\\
=&\sup_{1\le i,j\le N}\|T^{-2}\mathbf{V}_j'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\mathbf{\Gamma}_{xi}\|+\sup_{1\le i,j\le N}\|T^{-2}\mathbf{V}_j'\mathbf{F}_x^0\mathbf{G}_x(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{F}_x^0\mathbf{\Gamma}_{xi}\|\\
&+\sup_{1\le i,j\le N}\|T^{-2}\mathbf{V}_j'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{F}_x^0\mathbf{\Gamma}_{xi}\|+\sup_{1\le i,j\le N}\|T^{-2}\mathbf{V}_j'\mathbf{F}_x^0(\mathbf{G}_x\mathbf{G}_x'-\big(T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}\big)^{-1})\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\mathbf{\Gamma}_{xi}\|
\end{aligned}$$

We bound the first term in norm by

$$\sup_{1\le j\le N}\|T^{-1}\mathbf{V}_j'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|\cdot\sup_{1\le i\le N}\|\mathbf{\Gamma}_{xi}\|\|\mathbf{G}_x\|\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\|=O_p(N^{1/2}\delta_{NT}^{-2})$$

With Lemma B.4, the second term is bounded in norm by

$$\sup_{1\le j\le N}\|T^{-1}\mathbf{V}_j'\mathbf{F}_x^0\|\cdot\sup_{1\le i\le N}\|\mathbf{\Gamma}_{xi}\|\|\mathbf{G}_x\|\|T^{-1}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{F}_x^0\|=O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})$$

Given Lemma B.4 and Lemma B.7(a), the third is bounded in norm by

$$\sup_{1\le j\le N}\|T^{-1}\mathbf{V}_j'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|\cdot\sup_{1\le i\le N}\|T^{-1}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{F}_x^0\mathbf{\Gamma}_{xi}\|=O_p(N^{1/2}\delta_{NT}^{-4})$$

the forth term is bounded in norm by

$$\sup_{1\le j\le N}\|T^{-1}\mathbf{V}_j'\mathbf{F}_x^0\|\cdot\sup_{1\le i\le N}\|\mathbf{\Gamma}_{xi}\|\|\mathbf{G}_x\mathbf{G}_x'-(T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0)^{-1}\|\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\|=O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})$$

by (B.40) and (B.11). Thus $\mathbb{C}_2$ is $O_p(N^{1/2}\delta_{NT}^{-2})$. $\mathbb{C}_3$ is bounded in norm by

$$\begin{aligned}
&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{V}_j'(\mathbf{M}_{\widehat{\mathbf{F}}_x}-\mathbf{M}_{\mathbf{F}_x^0})\mathbf{V}_i\|\\
=&\sup_{1\le i,j\le N}\|T^{-2}\mathbf{V}_j'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{V}_i\|+\sup_{1\le i,j\le N}\|T^{-2}\mathbf{V}_j'\mathbf{F}_x^0\mathbf{G}_x(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{V}_i\|\\
&+\sup_{1\le i,j\le N}\|T^{-2}\mathbf{V}_j'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{V}_i\|+\sup_{1\le i,j\le N}\|T^{-2}\mathbf{V}_j'\mathbf{F}_x^0(\mathbf{G}_x\mathbf{G}_x'-\big(T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\big)^{-1})\mathbf{F}_x^{0\prime}\mathbf{V}_i\|
\end{aligned}$$

The first term is bounded in norm by

$$\sup_{1\le j\le N}\|T^{-1}\mathbf{V}_j'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|\cdot\sup_{1\le i\le N}\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{V}_i\|\cdot\|\mathbf{G}_x\|=O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})$$

Similarly, the second term is $O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})$. The third term is bounded in norm by

$$\Big(\sup_{1\le i\le N}\|T^{-1}\mathbf{V}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|\Big)^2=O_p(N^{1/2}\delta_{NT}^{-4})$$

The fourth term is bounded in norm by

$$T^{-1}\cdot\sup_{1\le i\le N}\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|^2\cdot\|\mathbf{G}_x\mathbf{G}_x'-\left(T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0\right)^{-1}\|=O_p(N^{1/2}T^{-1}\delta_{NT}^{-2})$$

With the above terms, we have $\mathbb{C}_3=O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})+O_p(N^{1/2}\delta_{NT}^{-4})$. Then, we have (c).

Consider (d). The term is bounded in norm by

$$\begin{aligned}
&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{X}_{j,-1}'\mathbf{M}_{\hat F_{x,-1}}\mathbf{M}_{\hat F_x}\mathbf{M}_{\hat F_{x,-1}}\mathbf{X}_{i,-1}-T^{-1}\mathbf{X}_{j,-1}'\mathbf{M}_{F^0_{x,-1}}\mathbf{M}_{F^0_x}\mathbf{M}_{F^0_{x,-1}}\mathbf{X}_{i,-1}\|\\
\le&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{\Gamma}_{xj}'\mathbf{F}_{x,-1}^{0\prime}\mathbf{M}_{\hat F_{x,-1}}\mathbf{M}_{\hat F_x}\mathbf{M}_{\hat F_{x,-1}}\mathbf{F}^0_{x,-1}\mathbf{\Gamma}_{xi}\|\\
&+2\sup_{1\le i,j\le N}\|T^{-1}\mathbf{V}_{j,-1}'\mathbf{M}_{\hat F_{x,-1}}\mathbf{M}_{\hat F_x}\mathbf{M}_{\hat F_{x,-1}}\mathbf{F}^0_{x,-1}\mathbf{\Gamma}_{xi}\|\\
&+\sup_{1\le i,j\le N}\|T^{-1}\mathbf{V}_{j,-1}'(\mathbf{M}_{\hat F_{x,-1}}\mathbf{M}_{\hat F_x}\mathbf{M}_{\hat F_{x,-1}}-\mathbf{M}_{F^0_{x,-1}}\mathbf{M}_{F^0_x}\mathbf{M}_{F^0_{x,-1}})\mathbf{V}_{i,-1}\|\\
=&\mathbb{C}_4+\mathbb{C}_5+\mathbb{C}_6
\end{aligned}$$

For $\mathbb{C}_4$, we have

$$\begin{aligned}
\mathbb{C}_4=&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{\Gamma}_{xj}'(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})'\mathbf{M}_{\hat F_{x,-1}}\mathbf{M}_{\hat F_x}\mathbf{M}_{\hat F_{x,-1}}(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})'\mathbf{\Gamma}_{xi}\|\\
\le&(\sup_{1\le i\le N}\|\mathbf{\Gamma}_{xi}\|)^2\cdot\|T^{-1/2}(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\|^2=O_p(N^{1/2}\delta_{NT}^{-2})
\end{aligned}$$

Ignoring the scale 2, $\mathbb{C}_5$ is bounded in norm by

$$\begin{aligned}
&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{V}_{j,-1}'\mathbf{M}_{\hat F_{x,-1}}\mathbf{M}_{\hat F_x}\mathbf{M}_{\hat F_{x,-1}}(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\mathbf{\Gamma}_{xi}\|\\
\le&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{V}_{j,-1}'(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\mathbf{\Gamma}_{xi}\|\\
&+3\sup_{1\le i,j\le N}\|T^{-2}\mathbf{V}_{j,-1}'\widehat{F}_{x,-1}\widehat{F}_{x,-1}'(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\mathbf{\Gamma}_{xi}\|\\
&+\sup_{1\le i,j\le N}\|T^{-2}\mathbf{V}_{j,-1}'\widehat{F}_x\widehat{F}_x'(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\mathbf{\Gamma}_{xi}\|\\
&+\sup_{1\le i,j\le N}\|T^{-3}\mathbf{V}_{j,-1}'\widehat{F}_{x,-1}\widehat{F}_{x,-1}'\widehat{F}_x\widehat{F}_x'(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\mathbf{\Gamma}_{xi}\|\\
&+\sup_{1\le i,j\le N}\|T^{-3}\mathbf{V}_{j,-1}'\widehat{F}_x\widehat{F}_x'\widehat{F}_{x,-1}\widehat{F}_{x,-1}'(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\mathbf{\Gamma}_{xi}\|\\
&+\sup_{1\le i,j\le N}\|T^{-4}\mathbf{V}_{j,-1}'\widehat{F}_{x,-1}\widehat{F}_{x,-1}'\widehat{F}_x\widehat{F}_x'\widehat{F}_{x,-1}\widehat{F}_{x,-1}'(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\mathbf{\Gamma}_{xi}\|
\end{aligned}$$

With Lemma 15 (b), the first term is bounded in norm by

$$\sup_{1\le j\le N}\|T^{-1}\mathbf{V}_{j,-1}'(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\|\sup_{1\le i\le N}\|\mathbf{\Gamma}_{xi}\|=O_p(N^{1/2}\delta_{NT}^{-2})$$

With Lemma 15 (b), the second term is bounded in norm by

$$\begin{aligned}
&\sup_{1\le j\le N}\|T^{-1}\mathbf{V}_{j,-1}'\widehat{F}_{x,-1}\|\sup_{1\le i\le N}\|\mathbf{\Gamma}_{xi}\|\cdot\|T^{-1}\widehat{F}_{x,-1}'(\mathbf{F}^0_{x,-1}-\widehat{\mathbf{F}}_{x,-1}\mathbf{G}_x^{*-1})\|\\
=&O_p(\delta_{NT}^{-2})\cdot\sup_{1\le j\le N}\|T^{-1}\mathbf{V}_{j,-1}'\widehat{F}_{x,-1}\|\sup_{1\le i\le N}\|\mathbf{\Gamma}_{xi}\|\\
\le&O_p(\delta_{NT}^{-2})\|\mathbf{G}_x^*\|\cdot\sup_{1\le j\le N}\|T^{-1}\mathbf{V}_{j,-1}'\mathbf{F}^0_{x,-1}\|\sup_{1\le i\le N}\|\mathbf{\Gamma}_{xi}\|\\
&+O_p(\delta_{NT}^{-2})\cdot\sup_{1\le j\le N}\|T^{-1}\mathbf{V}_{j,-1}'(\widehat{\mathbf{F}}_{x,-1}-\mathbf{F}^0_{x,-1}\mathbf{G}_x^*)\|\sup_{1\le i\le N}\|\mathbf{\Gamma}_{xi}\|\\
=&O_p(N^{1/2}\delta_{NT}^{-2})+O_p(N^{1/2}\delta_{NT}^{-4})=O_p(N^{1/2}\delta_{NT}^{-2})
\end{aligned}$$

Analogously, other terms can be proved to be $O_p(N^{1/2}\delta_{NT}^{-2})$. Then $\mathbb{C}_5 = O_p(N^{1/2}\delta_{NT}^{-2})$. Consider $\mathbb{C}_6$, we have

$$
\begin{aligned}
&\sup_{1\le i,j\le N} \|T^{-1}\mathbf{V}'_{j,-1}(\mathbf{M}_{\hat{F}_{x,-1}}\mathbf{M}_{\hat{F}_x}\mathbf{M}_{\hat{F}_{x,-1}} - \mathbf{M}_{F^0_{x,-1}}\mathbf{M}_{F^0_x}\mathbf{M}_{F^0_{x,-1}})\mathbf{V}_{i,-1}\| \\
\le &\sup_{1\le i,j\le N} \|T^{-1}\mathbf{V}'_{j,-1}(\mathbf{M}_{\hat{F}_{x,-1}} - \mathbf{M}_{F^0_{x,-1}})\mathbf{M}_{F^0_x}\mathbf{M}_{F^0_{x,-1}}\mathbf{V}_{i,-1}\| \\
&+ \sup_{1\le i,j\le N} \|T^{-1}\mathbf{V}'_{j,-1}\mathbf{M}_{F^0_{x,-1}}(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F^0_x})\mathbf{M}_{F^0_{x,-1}}\mathbf{V}_{i,-1}\| \\
&+ \sup_{1\le i,j\le N} \|T^{-1}\mathbf{V}'_{j,-1}\mathbf{M}_{F^0_{x,-1}}\mathbf{M}_{F^0_x}(\mathbf{M}_{\hat{F}_{x,-1}} - \mathbf{M}_{F^0_{x,-1}})\mathbf{V}_{i,-1}\| \\
&+ \sup_{1\le i,j\le N} \|T^{-1}\mathbf{V}'_{j,-1}(\mathbf{M}_{\hat{F}_{x,-1}} - \mathbf{M}_{F^0_{x,-1}})(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F^0_x})\mathbf{M}_{F^0_{x,-1}}\mathbf{V}_{i,-1}\| \\
&+ \sup_{1\le i,j\le N} \|T^{-1}\mathbf{V}'_{j,-1}(\mathbf{M}_{\hat{F}_{x,-1}} - \mathbf{M}_{F^0_{x,-1}})\mathbf{M}_{F^0_x}(\mathbf{M}_{\hat{F}_{x,-1}} - \mathbf{M}_{F^0_{x,-1}})\mathbf{V}_{i,-1}\| \\
&+ \sup_{1\le i,j\le N} \|T^{-1}\mathbf{V}'_{j,-1}\mathbf{M}_{F^0_{x,-1}}(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F^0_x})(\mathbf{M}_{\hat{F}_{x,-1}} - \mathbf{M}_{F^0_{x,-1}})\mathbf{V}_{i,-1}\| \\
&+ \sup_{1\le i,j\le N} \|T^{-1}\mathbf{V}'_{j,-1}(\mathbf{M}_{\hat{F}_{x,-1}} - \mathbf{M}_{F^0_{x,-1}})(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F^0_x})(\mathbf{M}_{\hat{F}_{x,-1}} - \mathbf{M}_{F^0_{x,-1}})\mathbf{V}_{i,-1}\|
\end{aligned}
$$

Following the argument in the proof of $\mathbb{C}_3$, the first three terms are shown to be $O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2}) + O_p(N^{1/2}\delta_{NT}^{-4})$. Note that $\mathbf{M}_{\widehat{\mathbf{F}}_x} - \mathbf{M}_{\mathbf{F}^0_x} = -T^{-1}(\widehat{\mathbf{F}}_x - \mathbf{F}^0_x\mathbf{G}_x)\mathbf{G}'_x\mathbf{F}^{0\prime}_x - T^{-1}\mathbf{F}^0_x\mathbf{G}_x(\widehat{\mathbf{F}}_x - \mathbf{F}^0_x\mathbf{G}_x)'$ $- T^{-1}(\widehat{\mathbf{F}}_x - \mathbf{F}^0_x\mathbf{G}_x)(\widehat{\mathbf{F}}_x - \mathbf{F}^0_x\mathbf{G}_x)' - T^{-1}\mathbf{F}^0_x\left(\mathbf{G}_x\mathbf{G}'_x - \left(T^{-1}\mathbf{F}^{0\prime}_x\mathbf{F}\right)^{-1}\right)\mathbf{F}^{0\prime}_x$ and $\mathbf{M}_{\widehat{\mathbf{F}}_{x,-1}} - \mathbf{M}_{\mathbf{F}^0_{x,-1}} =$ $- T^{-1}(\widehat{\mathbf{F}}_{x,-1} - \mathbf{F}^0_{x,-1}\mathbf{G}_{x,-1})\mathbf{G}'_{x,-1}\mathbf{F}^{0\prime}_{x,-1} - T^{-1}\mathbf{F}^0_{x,-1}\mathbf{G}_{x,-1}(\widehat{\mathbf{F}}_{x,-1} - \mathbf{F}^0_{x,-1}\mathbf{G}_{x,-1})'$ $- T^{-1}(\widehat{\mathbf{F}}_{x,-1} - \mathbf{F}^0_{x,-1}\mathbf{G}_{x,-1})(\widehat{\mathbf{F}}_{x,-1} - \mathbf{F}^0_{x,-1}\mathbf{G}_{x,-1})' - T^{-1}\mathbf{F}^0_{x,-1}\left(\mathbf{G}_{x,-1}\mathbf{G}'_{x,-1} - \left(T^{-1}\mathbf{F}^{0\prime}_{x,-1}\mathbf{F}\right)^{-1}\right)\mathbf{F}^{0\prime}_{x,-1}$. We can substitute $\mathbf{M}_{\widehat{\mathbf{F}}_x} - \mathbf{M}_{\mathbf{F}^0_x}$ and $\mathbf{M}_{\widehat{\mathbf{F}}_{x,-1}} - \mathbf{M}_{\mathbf{F}^0_{x,-1}}$ with the above two equations, then follow the argument in the proof of $\mathbb{C}_3$ to derive that the other terms are $o_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2}) + o_p(N^{1/2}\delta_{NT}^{-4})$. Thus $\mathbb{C}_6 = O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2}) + O_p(N^{1/2}\delta_{NT}^{-4})$. Collecting the above terms, we have (d). Analogously, we can prove (e), Thus the details are omitted. Thus, we complete the proof. ■

Proof of Lemma 16. Consider (a). Define $\rho_{\max} = \sup_{1\le i\le N}|\rho_i|$. Note that $\mathbf{y}_i = \rho_i\mathbf{y}_{i,-1} + \mathbf{X}_i\boldsymbol{\beta}_i + \mathbf{F}_y\boldsymbol{\gamma}_{yi} + \boldsymbol{\varepsilon}_i$, then

$$
\mathbf{y}_i = \sum_{s=0}^{\infty}\mathbf{X}_{i,-s}\boldsymbol{\beta}_i\rho_i^s + \sum_{s=0}^{\infty}\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^s + \sum_{s=0}^{\infty}\boldsymbol{\varepsilon}_{i,-s}\rho_i^s\,, \tag{B.43}
$$

and

$$
\mathbf{y}_{i,-1} = \sum_{s=1}^{\infty}\mathbf{X}_{i,-s}\boldsymbol{\beta}_i\rho_i^{s-1} + \sum_{s=1}^{\infty}\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1} + \sum_{s=1}^{\infty}\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\,, \tag{B.44}
$$

With (B.44), we can derive that

$$
\begin{aligned}
&\sup_{1\le i,j\le N}\|T^{-1}\mathbf{X}'_i\mathbf{M}_{\hat{F}_x}\mathbf{y}_{i,-1} - T^{-1}\mathbf{X}'_i\mathbf{M}_{F^0_x}\mathbf{y}_{i,-1}\| \\
\le &\sup_{1\le i,j\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{X}'_i(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F^0_x})\mathbf{X}_{i,-s}\boldsymbol{\beta}_i\rho_i^{s-1}\| \\
&+ \sup_{1\le i,j\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{X}'_i(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F^0_x})\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\| + \sup_{1\le i,j\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{X}'_i(\mathbf{M}_{\hat{F}_x} - \mathbf{M}_{F^0_x})\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
= &\mathbb{D}_1 + \mathbb{D}_2 + \mathbb{D}_3
\end{aligned}
$$

Consider $\mathbb{D}_2$, which is bounded in norm by

$$\begin{aligned}
&\sup_{1\leq i,j\leq N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{X}_i'(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0})\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\| \\
\leq&\sup_{1\leq i,j\leq N}\|T^{-2}\sum_{s=1}^{\infty}\mathbf{X}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\| \\
&+\sup_{1\leq i,j\leq N}\|T^{-2}\sum_{s=1}^{\infty}\mathbf{X}_i'\mathbf{F}_x^0\mathbf{G}_x(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\| \\
&+\sup_{1\leq i,j\leq N}\|T^{-2}\sum_{s=1}^{\infty}\mathbf{X}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\| \\
&+\sup_{1\leq i,j\leq N}\|T^{-2}\sum_{s=1}^{\infty}\mathbf{X}_i'\mathbf{F}_x^0\left(\mathbf{G}_x\mathbf{G}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right)\mathbf{F}_x^{0\prime}\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\| \\
=&\mathbb{D}_{2.1}+\mathbb{D}_{2.2}+\mathbb{D}_{2.3}+\mathbb{D}_{2.4}
\end{aligned}$$

Note that

$$\begin{aligned}
&\sup_{1\leq i\leq N}\|T^{-1}\mathbf{X}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\| \\
\leq&\sup_{1\leq i\leq N}\|\boldsymbol{\Gamma}_{xi}\|\|T^{-1}\mathbf{F}_x'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|+\sup_{1\leq i\leq N}\|T^{-1}\mathbf{V}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\| \\
=&O_p(N^{1/4}\delta_{NT}^{-2})+O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})
\end{aligned} \tag{B.45}$$

The term $\mathbb{D}_{2.1}$ is bounded in norm by

$$\begin{aligned}
&\sup_{1\leq i\leq N}\|T^{-1}\mathbf{X}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|\cdot\sup_{1\leq i\leq N}\|\boldsymbol{\gamma}_{yi}\|\cdot\sup_{1\leq i\leq N}\sum_{s=1}^{\infty}\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_{y,-s}\||\rho_i|^{s-1}\cdot\|\mathbf{G}_x\| \\
=&\sup_{1\leq i\leq N}\sum_{s=1}^{\infty}\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_{y,-s}\||\rho_i|^{s-1}\cdot[O_p(N^{1/2}\delta_{NT}^{-2})+O_p(N^{3/4}T^{-1/2}\delta_{NT}^{-2})] \\
=&O_p(N^{1/2}\delta_{NT}^{-2})+O_p(N^{3/4}T^{-1/2}\delta_{NT}^{-2})
\end{aligned}$$

since $\sup_{1\leq i\leq N}\sum_{s=1}^{\infty}\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_{y,-s}\||\rho_i|^{s-1}\leq\sum_{s=1}^{\infty}\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_{y,-s}\|\rho_{\max}^{s-1}=O_p(1)$, which is implied from

$$\begin{aligned}
&E(\sum_{s=1}^{\infty}\|T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_{y,-s}\|\rho_{\max}^{s-1}) \\
\leq&\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\sqrt{E\|T^{-1/2}\mathbf{F}_x^0\|^2E\|T^{-1/2}\mathbf{F}_{y,-s}\|^2}\leq\Delta\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\leq\Delta
\end{aligned} \tag{B.46}$$

with $\rho_{\max}=\sup_{1\leq i\leq N}|\rho_i|<1$. Similarly, we have $\mathbb{D}_{2.4}=O_p(N^{1/2}\delta_{NT}^{-2})$. Because

$$\begin{aligned}
\hat{\mathbf{F}}_x'-\mathbf{G}_x'\mathbf{F}_x^{0\prime}=&\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{G}_x'\hat{\mathbf{Q}}_x'\hat{\mathbf{F}}_x'\mathbf{v}_{\ell i}\gamma_{\ell i}^{0\prime}\mathbf{F}_x^{0\prime}+\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{G}_x'\hat{\mathbf{Q}}_x'\hat{\mathbf{F}}_x'\mathbf{F}_x^0\gamma_{\ell i}^0\mathbf{v}_{\ell i}' \\
&+\frac{1}{NT}\sum_{\ell=1}^{k}\sum_{i=1}^{N}\mathbf{G}_x'\hat{\mathbf{Q}}_x'\hat{\mathbf{F}}_x'\mathbf{v}_{\ell i}\mathbf{v}_{\ell i}'.
\end{aligned} \tag{B.47}$$

the term $\mathbb{D}_{2.2}$ is bounded in norm by

$$
\begin{aligned}
&\sup_{1\le i,j\le N}\|T^{-2}\sum_{s=1}^{\infty}\mathbf{X}_i'\mathbf{F}_x^0\mathbf{G}_x(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\|\\
\le&\sup_{1\le i\le N}\|T^{-1}\mathbf{X}_i'\mathbf{F}_x^0\|\|\mathbf{G}_x\|\cdot\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\|\\
=&O_p(N^{1/4})\cdot\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\|\\
\le&O_p(N^{1/4})\cdot\|\mathbf{G}_x\|\|\hat{\mathbf{Q}}_x\|\cdot\sup_{1\le i,j\le N}\|N^{-1}T^{-2}\sum_{s=1}^{\infty}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\boldsymbol{\gamma}_{\ell h}^{0\prime}\mathbf{F}_x^{0\prime}\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\|\\
&+O_p(N^{1/4})\cdot\|\mathbf{G}_x\|\|\hat{\mathbf{Q}}_x\|\cdot\sup_{1\le i,j\le N}\|N^{-1}T^{-2}\sum_{s=1}^{\infty}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{F}_x^0\boldsymbol{\gamma}_{\ell h}^{0}\mathbf{v}_{\ell h}'\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\|\\
&+O_p(N^{1/4})\cdot\|\mathbf{G}_x\|\|\hat{\mathbf{Q}}_x\|\cdot\sup_{1\le i,j\le N}\|N^{-1}T^{-2}\sum_{s=1}^{\infty}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}'\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\|
\end{aligned}
$$

Note that $N^{-1}T^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\boldsymbol{\gamma}_{\ell h}^{0\prime}=\mathbf{A}_{kNT}=O_p(N^{-1/2}T^{-1/2})+O_p(N^{-1})$ as shown in the proof of Lemma 4. With (B.46), the first term is bounded in norm by

$$
\begin{aligned}
&O_p(N^{1/4})\cdot\|N^{-1}T^{-1}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\boldsymbol{\gamma}_{\ell h}^{0\prime}\|\cdot\sup_{1\le i\le N}\|\boldsymbol{\gamma}_{yi}\|\cdot T^{-1}\sum_{s=1}^{\infty}\|\mathbf{F}_x^{0\prime}\mathbf{F}_{y,-s}\|\rho_{\max}^{s-1}\\
=&O_p(T^{-1/2})+O_p(N^{-1/2})
\end{aligned}
$$

Similar to the argument in (B.46), we show $\sum_{s=1}^{\infty}\|N^{-1/2}T^{-1/2}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\boldsymbol{\gamma}_{\ell h}^{0}\mathbf{v}_{\ell h}'\mathbf{F}_{y,-s}\|\rho_{\max}^{s-1}=O_p(1)$. Then the second term is bounded in norm by

$$
\begin{aligned}
&O_p(N^{-1/4}T^{-1/2})\cdot\|T^{-1}\hat{\mathbf{F}}_x'\mathbf{F}_x^0\|\cdot\sup_{1\le i\le N}\|\boldsymbol{\gamma}_{yi}\|\cdot\sum_{s=1}^{\infty}\|N^{-1/2}T^{-1/2}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\boldsymbol{\gamma}_{\ell h}^{0}\mathbf{v}_{\ell h}'\mathbf{F}_{y,-s}\|\rho_{\max}^{s-1}\\
=&O_p(T^{-1/2})
\end{aligned}
$$

The third term is bounded in norm by

$$
O_p(N^{1/4})\cdot\sup_{1\le i\le N}\|\boldsymbol{\gamma}_{yi}\|\cdot\sum_{s=1}^{\infty}\|N^{-1}T^{-2}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}'\mathbf{F}_{y,-s}\|\rho_{\max}^{s-1}
$$

Note that

$$
\begin{aligned}
&N^{-1}T^{-2}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}'\mathbf{F}_{y,-s}\\
\le&N^{-1}T^{-2}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\mathbf{G}_x'\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}'\mathbf{F}_{y,-s}+N^{-1}T^{-2}\sum_{\ell=1}^{k}\sum_{h=1}^{N}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'E(\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}')\mathbf{F}_{y,-s}\\
&+N^{-1}T^{-2}\sum_{\ell=1}^{k}\sum_{h=1}^{N}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'(\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}'-E(\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}'))\mathbf{F}_{y,-s}
\end{aligned}
$$

then following the argument in (B.46), we can prove that

$$
\sum_{s=1}^{\infty}\|N^{-1}T^{-2}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}'\mathbf{F}_{y,-s}\|\rho_{\max}^{s-1}=O_p(\delta_{NT}^{-2})
$$

then the third term is $O_p(N^{1/2}\delta_{NT}^{-2})$. Thus, $\mathbb{D}_{2.2}=O_p(N^{1/2}\delta_{NT}^{-2})$. Analogously, $\mathbb{D}_{2.3}=O_p(N^{1/2}\delta_{NT}^{-4})+O_p(N^{3/4}T^{-1/2}\delta_{NT}^{-4})$. With the above four terms, $\mathbb{D}_2=O_p(N^{1/2}\delta_{NT}^{-2})+O_p(N^{3/4}T^{-1/2}\delta_{NT}^{-2})$.

Consider $\mathbb{D}_3$, which is bounded in norm by

$$
\begin{aligned}
&\sup_{1\le i,j\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{X}_i'(\mathbf{M}_{\hat{F}_x}-\mathbf{M}_{F_x^0})\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
\le&\sup_{1\le i,j\le N}\|T^{-2}\sum_{s=1}^{\infty}\mathbf{X}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\mathbf{G}_x'\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
&+\sup_{1\le i,j\le N}\|T^{-2}\sum_{s=1}^{\infty}\mathbf{X}_i'\mathbf{F}_x^0\mathbf{G}_x(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
&+\sup_{1\le i,j\le N}\|T^{-2}\sum_{s=1}^{\infty}\mathbf{X}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
&+\sup_{1\le i,j\le N}\|T^{-2}\sum_{s=1}^{\infty}\mathbf{X}_i'\mathbf{F}_x^0\left(\mathbf{G}_x\mathbf{G}_x'-\left(\frac{\mathbf{F}_x^{0\prime}\mathbf{F}_x^0}{T}\right)^{-1}\right)\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
=&\mathbb{D}_{3.1}+\mathbb{D}_{3.2}+\mathbb{D}_{3.3}+\mathbb{D}_{3.4}
\end{aligned}
$$

Note that $\sup_{1\le i\le N}\|T^{-1}\mathbf{X}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|=O_p(N^{1/4}\delta_{NT}^{-2})+O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-2})$, the term $\mathbb{D}_{3.1}$ is bounded in norm by

$$
\begin{aligned}
&\sup_{1\le i\le N}\|T^{-1}\mathbf{X}_i'(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)\|\cdot\sup_{1\le i\le N}\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|T^{-1}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\|\cdot\|\mathbf{G}_x\| \\
=&\sup_{1\le i\le N}\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|T^{-1/2}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\|\cdot[O_p(N^{1/4}T^{-1/2}\delta_{NT}^{-2})+O_p(N^{1/2}T^{-1}\delta_{NT}^{-2})] \\
=&O_p(N^{3/4}T^{-1/2}\delta_{NT}^{-2})+O_p(NT^{-1}\delta_{NT}^{-2})
\end{aligned}
$$

because $\sup\limits_{1\le i\le N}\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|T^{-1/2}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\|=O_p(N^{1/2})$, which is given by

$$
\begin{aligned}
&E\Big(\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|T^{-1/2}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\|\Big)^2=\sum_{s=1}^{\infty}\sum_{t=1}^{\infty}\rho_{\max}^{s-1}\rho_{\max}^{t-1}E\|T^{-1/2}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\|\|T^{-1/2}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-t}\| \\
\le&\sum_{s=1}^{\infty}\sum_{t=1}^{\infty}\rho_{\max}^{s-1}\rho_{\max}^{t-1}\sqrt{E\|T^{-1/2}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\|^2E\|T^{-1/2}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-t}\|^2}\le\Delta\sum_{s=1}^{\infty}\sum_{t=1}^{\infty}\rho_{\max}^{s-1}\rho_{\max}^{t-1}\le\Delta
\end{aligned}
\tag{B.48}
$$

Similarly, we can prove that $\mathbb{D}_{3.4}=O_p(N^{3/4}T^{-1/2}\delta_{NT}^{-2})$. The term $\mathbb{D}_{3.2}$ is bounded in norm by

$$
\begin{aligned}
&\sup_{1\le i,j\le N}\|T^{-2}\sum_{s=1}^{\infty}\mathbf{X}_i'\mathbf{F}_x^0\mathbf{G}_x(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
\le&\sup_{1\le i\le N}\|T^{-1}\mathbf{X}_i'\mathbf{F}_x^0\|\|\mathbf{G}_x\|\cdot\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}(\widehat{\mathbf{F}}_x-\mathbf{F}_x^0\mathbf{G}_x)'\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
\le&O_p(N^{1/4})\cdot\sup_{1\le i,j\le N}\|N^{-1}T^{-2}\sum_{s=1}^{\infty}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\boldsymbol{\gamma}_{\ell h}^{0\prime}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
&+O_p(N^{1/4})\cdot\sup_{1\le i,j\le N}\|N^{-1}T^{-2}\sum_{s=1}^{\infty}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{F}_x^0\boldsymbol{\gamma}_{\ell h}^0\mathbf{v}_{\ell h}'\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
&+O_p(N^{1/4})\cdot\sup_{1\le i,j\le N}\|N^{-1}T^{-2}\sum_{s=1}^{\infty}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}'\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\|
\end{aligned}
$$

The first term is bounded in norm by

$$
\begin{aligned}
&O_p(N^{1/4})\cdot\sup_{1\le i,j\le N}\|N^{-1}T^{-2}\sum_{s=1}^{\infty}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\boldsymbol{\gamma}_{\ell h}^{0\prime}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
\le&O_p(N^{1/4}T^{-1/2})\cdot\sup_{1\le i\le N}\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|T^{-1/2}\mathbf{F}_x^{0\prime}\boldsymbol{\varepsilon}_{i,-s}\|\cdot\|N^{-1}T^{-1}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\boldsymbol{\gamma}_{\ell h}^{0\prime}\| \\
=&O_p(N^{1/4}T^{-1})+O_p(N^{-1/4}T^{-1/2})
\end{aligned}
$$

with $N^{-1}T^{-1}\sum_{\ell=1}^{k}\sum_{j=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\boldsymbol{\gamma}_{\ell h}^{0\prime} = \mathbf{A}_{kNT} = O_p(N^{-1/2}T^{-1/2}) + O_p(N^{-1})$ and (B.48). Then the second term is bounded in norm by

$$
\begin{aligned}
&O_p(N^{-1/4}T^{-1/2}) \cdot \|T^{-1}\hat{\mathbf{F}}_x'\mathbf{F}_x^0\| \cdot \sup_{1\le i\le N}\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|N^{-1/2}T^{-1/2}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\boldsymbol{\gamma}_{\ell h}^{0}\mathbf{v}_{\ell h}'\boldsymbol{\varepsilon}_{i,-s}\| \\
=&O_p(N^{1/4}T^{-1/2})
\end{aligned}
$$

because $E(\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|N^{-1/2}T^{-1/2}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\boldsymbol{\gamma}_{\ell h}^{0}\mathbf{v}_{\ell h}'\boldsymbol{\varepsilon}_{i,-s}\|)^2 \le \Delta$, which can be proved by following the argument in the proof of (B.48). The third term is bounded in norm by

$$
\begin{aligned}
&O_p(N^{1/4}) \cdot \sup_{1\le i,j\le N}\|N^{-1}T^{-2}\sum_{s=1}^{\infty}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\hat{\mathbf{F}}_x'\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}'\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\| \\
\le&O_p(N^{1/4}T^{-1}) \cdot \|\mathbf{G}_x\| \cdot \sup_{1\le i,j\le N}N^{-1}T^{-1}\sum_{s=1}^{\infty}\sum_{\ell=1}^{k}\sum_{h=1}^{N}\|\mathbf{F}_x^{0\prime}\mathbf{v}_{\ell h}\|\|\mathbf{v}_{\ell h}'\boldsymbol{\varepsilon}_{i,-s}\|\rho_{\max}^{s-1} \\
&+ O_p(N^{1/4}T^{-1}) \cdot \|\widehat{\mathbf{F}}_x - \mathbf{F}_x^0\mathbf{G}_x\| \cdot \sup_{1\le i,j\le N}N^{-1}T^{-1}\sum_{s=1}^{\infty}\|\sum_{\ell=1}^{k}\sum_{h=1}^{N}\mathbf{v}_{\ell h}\mathbf{v}_{\ell h}'\boldsymbol{\varepsilon}_{i,-s}\|\rho_{\max}^{s-1} \\
=&O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-1})
\end{aligned}
$$

Thus, $\mathbb{D}_{3.2} = O_p(N^{1/4}T^{-1/2}) + O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-1})$. Similarly, we can show that $\mathbb{D}_{3.3} = O_p(N^{1/4}T^{-1/2}\delta_{NT}^{-2}) + O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-3}) + O_p(N^{3/4}T^{-1}\delta_{NT}^{-3})$. Combining the above terms, we have $\mathbb{D}_3 = O_p(N^{3/4}T^{-1/2}\delta_{NT}^{-2}) + O_p(NT^{-1}\delta_{NT}^{-2}) + O_p(N^{1/4}T^{-1/2}) + O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-1})$.

Note that $\mathbf{X}_{i,-s} = \mathbf{F}_{x,-s}\boldsymbol{\Gamma}_{xi} + \mathbf{V}_{i,-s}$, we can follow the arguments in the proof of $\mathbb{D}_2$ and $\mathbb{D}_3$, and then have $\mathbb{D}_1 = O_p(N^{1/2}\delta_{NT}^{-2}) + O_p(N^{3/4}T^{-1/2}\delta_{NT}^{-2}) + O_p(NT^{-1}\delta_{NT}^{-2}) + O_p(N^{1/4}T^{-1/2})$. Thus, with the above terms, claims (a) holds.

Analogous to the argument in the proof of (a), we can derive (b). Thus, we complete the proof. ■

Proof of Lemma 17. Consider (a). As $E\|T^{-1/2}[\mathbf{V}_i'\mathbf{V}_i - E(\mathbf{V}_i'\mathbf{V}_i)]\|^4 \le \Delta$ and $E\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|^4 \le \Delta$, $\sup_{1\le i\le N}\|T^{-1/2}[\mathbf{V}_i'\mathbf{V}_i - E(\mathbf{V}_i'\mathbf{V}_i)]\| = O_p(N^{1/4})$ and $\sup_{1\le i\le N}\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|^2 = O_p(N^{1/2})$. Thus

$$
\begin{aligned}
&\sup_{1\le i\le N}\|T^{-1}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{X}_i - T^{-1}E(\mathbf{V}_i'\mathbf{V}_i)\| \\
=&T^{-1/2}\cdot\sup_{1\le i\le N}\|T^{-1/2}[\mathbf{V}_i'\mathbf{V}_i - E(\mathbf{V}_i'\mathbf{V}_i)]\| + T^{-1}\cdot\sup_{1\le i\le N}\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|^2\|(T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0)^{-1}\| \\
=&O_p(N^{1/4}T^{-1/2}) + O_p(N^{1/2}T^{-1}).
\end{aligned}
$$

Similarly, we can prove (b) and (c). Consider (d), which is bounded in norm by

$$
\begin{aligned}
&\sup_{1\le i\le N}\|T^{-1}\mathbf{X}_i'\mathbf{M}_{F_x^0}\mathbf{y}_{i,-1} - T^{-1}\sum_{s=1}^{\infty}E(\mathbf{V}_i'\mathbf{V}_{i,-s})\boldsymbol{\beta}_i\rho_i^{s-1}\| \\
\le&\sup_{1\le i\le N}\|T^{-1}\mathbf{V}_i'\mathbf{M}_{F_x^0}\mathbf{y}_{i,-1} - T^{-1}\sum_{s=1}^{\infty}\mathbf{V}_i'\mathbf{V}_{i,-s}\boldsymbol{\beta}_i\rho_i^{s-1}\| \\
&+\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}[\mathbf{V}_i'\mathbf{V}_{i,-s} - E(\mathbf{V}_i'\mathbf{V}_{i,-s})]\boldsymbol{\beta}_i\rho_i^{s-1}\| \\
\le&\sup_{1\le i\le N}\|T^{-1}\mathbf{V}_i'\mathbf{P}_{F_x^0}\mathbf{y}_{i,-1}\| + \sup_{1\le i\le N}\|T^{-1}\mathbf{V}_i'\mathbf{y}_{i,-1} - T^{-1}\sum_{s=1}^{\infty}\mathbf{V}_i'\mathbf{V}_{i,-s}\boldsymbol{\beta}_i\rho_i^{s-1}\| \\
&+\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}[\mathbf{V}_i'\mathbf{V}_{i,-s} - E(\mathbf{V}_i'\mathbf{V}_{i,-s})]\boldsymbol{\beta}_i\rho_i^{s-1}\| \\
=&\mathbb{F}_1 + \mathbb{F}_2 + \mathbb{F}_3
\end{aligned}
$$

As

$$
\mathbf{y}_{i,-1} = \sum_{s=1}^{\infty}\mathbf{F}_{x,-s}\boldsymbol{\Gamma}_{xi}\boldsymbol{\beta}_i\rho_i^{s-1} + \sum_{s=1}^{\infty}\mathbf{V}_{i,-s}\boldsymbol{\beta}_i\rho_i^{s-1} + \sum_{s=1}^{\infty}\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1} + \sum_{s=1}^{\infty}\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}
$$

The term $\mathbb{F}_1$ is bounded in norm by

$$\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{V}_i'\mathbf{P}_{F_x^0}\mathbf{F}_{x,-s}\mathbf{\Gamma}_{xi}\boldsymbol{\beta}_i\rho_i^{s-1}\|+\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{V}_i'\mathbf{P}_{F_x^0}\mathbf{V}_{i,-s}\boldsymbol{\beta}_i\rho_i^{s-1}\|$$
$$+\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{V}_i'\mathbf{P}_{F_x^0}\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\|+\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{V}_i'\mathbf{P}_{F_x^0}\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\|,$$

The first term is bounded in norm by

$$T^{-1/2}\cdot\sup_{1\le i\le N}(\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|\|\mathbf{\Gamma}_{xi}\|\|\boldsymbol{\beta}_i\|)\cdot\sum_{s=1}^{\infty}\|T^{-1/2}\mathbf{F}_{x,-s}\|\rho_{\max}^{s-1}\cdot\|(T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0)^{-1}\|\|T^{-1/2}\mathbf{F}_x^0\|$$
$$=O_p(T^{-1/2})\cdot\sup_{1\le i\le N}(\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|\|\mathbf{\Gamma}_{xi}\|\|\boldsymbol{\beta}_i\|)\cdot\sum_{s=1}^{\infty}\|T^{-1/2}\mathbf{F}_{x,-s}\|\rho_{\max}^{s-1}=O_p(N^{1/4}T^{-1/2})$$

because

$$E(\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|\|\mathbf{\Gamma}_{xi}\|\|\boldsymbol{\beta}_i\|)^4=E\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|^4E\|\mathbf{\Gamma}_{xi}\|^4E\|\boldsymbol{\beta}_i\|^4\le\Delta$$

and

$$E\sum_{s=1}^{\infty}\|T^{-1/2}\mathbf{F}_{x,-s}\|\rho_{\max}^{s-1}\le\sum_{s=1}^{\infty}\Delta\rho_{\max}^{s-1}\le\Delta$$

Similarly, the third term is $O_p(N^{1/4}T^{-1/2})$. The second term is bounded in norm by

$$T^{-1}\cdot\sup_{1\le i\le N}(\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|\|\boldsymbol{\beta}_i\|)\cdot\sup_{1\le i\le N}\sum_{s=1}^{\infty}\|T^{-1/2}\mathbf{F}_x^{0\prime}\mathbf{V}_{i,-s}\|\rho_{\max}^{s-1}\cdot\|(T^{-1}\mathbf{F}_x^{0\prime}\mathbf{F}_x^0)^{-1}\|$$
$$=O_p(T^{-1})\cdot\sup_{1\le i\le N}(\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_x^0\|\|\boldsymbol{\beta}_i\|)\cdot\sup_{1\le i\le N}\sum_{s=1}^{\infty}\|T^{-1/2}\mathbf{F}_x^{0\prime}\mathbf{V}_{i,-s}\|\rho_{\max}^{s-1}=O_p(N^{1/2}T^{-1})$$

with $E(\sum_{s=1}^{\infty}\|T^{-1/2}\mathbf{F}_x^{0\prime}\mathbf{V}_{i,-s}\|\rho_{\max}^{s-1})^4\le\Delta$. Similarly, the forth term is $O_p(N^{1/2}T^{-1})$. With the above four terms, $\mathbb{F}_1=O_p(N^{1/4}T^{-1/2})+O_p(N^{1/2}T^{-1})$.

The term $\mathbb{F}_2$ is bounded in norm by

$$\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{V}_i'\mathbf{F}_{x,-s}\mathbf{\Gamma}_{xi}\boldsymbol{\beta}_i\rho_i^{s-1}\|$$
$$+\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{V}_i'\mathbf{F}_{y,-s}\boldsymbol{\gamma}_{yi}\rho_i^{s-1}\|+\sup_{1\le i\le N}\|T^{-1}\sum_{s=1}^{\infty}\mathbf{V}_i'\boldsymbol{\varepsilon}_{i,-s}\rho_i^{s-1}\|,$$

The first term is bounded in norm by

$$T^{-1/2}\cdot\sup_{1\le i\le N}\big(\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_{x,-s}\|\cdot\|\mathbf{\Gamma}_{xi}\|\|\boldsymbol{\beta}_i\|\big)=O_p(N^{1/4}T^{-1/2})$$

because

$$E\big(\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_{x,-s}\|\cdot\|\mathbf{\Gamma}_{xi}\|\|\boldsymbol{\beta}_i\|\big)^4$$
$$=E\big(\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|T^{-1/2}\mathbf{V}_i'\mathbf{F}_{x,-s}\|)^4\cdot E\|\mathbf{\Gamma}_{xi}\|^4E\|\boldsymbol{\beta}_i\|^4\le\Delta$$

Similarly, the second and the third terms are $O_p(N^{1/4}T^{-1/2})$. Then $\mathbb{F}_2=O_p(N^{1/4}T^{-1/2})$. Consider the term $\mathbb{F}_3$, which is bounded in norm by

$$T^{-1/2}\cdot\sup_{1\le i\le N}\big(\sum_{s=1}^{\infty}\rho_{\max}^{s-1}\|T^{-1/2}[\mathbf{V}_i'\mathbf{V}_{i,-s}-E(\mathbf{V}_i'\mathbf{V}_{i,-s})]\|\|\boldsymbol{\beta}_i\|\big)=O_p(N^{1/4}T^{-1/2})$$

since the forth moment of the term in the parenthesis is bounded. Consequently, claims (d) holds. Following the argument in the proof of (d), we can show that (e) holds. This completes the proof. ■

Proof of Lemma 18. With the definitions of $\tilde{\mathbf{A}}_{i,T}$, $\tilde{\mathbf{B}}_{i,T}$ and Lemmas 15, 16 and 17, we can derive that

$$\begin{aligned}
&\sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{A}}}_{i,T} - \tilde{\mathbf{A}}_{i,T}\| = O_p([N^{1/2} + N^{3/4}T^{-1/2} + NT^{-1}]\delta_{NT}^{-2}) + O_p(N^{1/4}T^{-1/2}) \\
&\sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{B}}}_{i,T} - \tilde{\mathbf{B}}_{i,T}\| = O_p(N^{1/2}\delta_{NT}^{-2}) \\
&\sup_{1\leq i\leq N} \|\tilde{\mathbf{A}}_{i,T} - \mathbf{A}_{i,T}\| = O_p(N^{1/4}T^{-1/2}) + O_p(N^{1/2}T^{-1}) \\
&\sup_{1\leq i\leq N} \|\tilde{\mathbf{B}}_{i,T} - \mathbf{B}_{i,T}\| = O_p(N^{1/4}T^{-1/2}) + O_p(N^{1/2}T^{-1})
\end{aligned} \tag{B.49}$$

Since $\tilde{\mathbf{B}}_{i,T}^{-1} - \mathbf{B}_{i,T}^{-1} = \mathbf{B}_{i,T}^{-1}(\mathbf{B}_{i,T} - \tilde{\mathbf{B}}_{i,T})\mathbf{B}_{i,T}^{-1} + \mathbf{B}_{i,T}^{-1}(\mathbf{B}_{i,T} - \tilde{\mathbf{B}}_{i,T})(\tilde{\mathbf{B}}_{i,T}^{-1} - \mathbf{B}_{i,T}^{-1})$, we have

$$\begin{aligned}
&\sup_{1\leq i\leq N} \|\tilde{\mathbf{B}}_{i,T}^{-1} - \mathbf{B}_{i,T}^{-1}\| \\
\leq& \sup_{1\leq i\leq N} \|\mathbf{B}_{i,T}^{-1}\| \sup_{1\leq i\leq N} \|\mathbf{B}_{i,T} - \tilde{\mathbf{B}}_{i,T}\| \sup_{1\leq i\leq N} \|\mathbf{B}_{i,T}^{-1}\| \\
&+ \sup_{1\leq i\leq N} \|\mathbf{B}_{i,T}^{-1}\| \sup_{1\leq i\leq N} \|\mathbf{B}_{i,T} - \tilde{\mathbf{B}}_{i,T}\| \sup_{1\leq i\leq N} \|\tilde{\mathbf{B}}_{i,T}^{-1} - \mathbf{B}_{i,T}^{-1}\| \\
=&O(1) \cdot \sup_{1\leq i\leq N} \|\mathbf{B}_{i,T} - \tilde{\mathbf{B}}_{i,T}\| + \left[O_p(N^{1/4}T^{-1/2}) + O_p(N^{1/2}T^{-1})\right] \sup_{1\leq i\leq N} \|\tilde{\mathbf{B}}_{i,T}^{-1} - \mathbf{B}_{i,T}^{-1}\|
\end{aligned} \tag{B.50}$$

Given $N/T^2 \to 0$, the second term is $o_p(1) \cdot \sup_{1\leq i\leq N} \|\tilde{\mathbf{B}}_{i,T}^{-1} - \mathbf{B}_{i,T}^{-1}\|$, it means that the second term is not the leading term while the first term is the leading term, thus

$$\sup_{1\leq i\leq N} \|\tilde{\mathbf{B}}_{i,T}^{-1} - \mathbf{B}_{i,T}^{-1}\| = O_p(N^{1/4}T^{-1/2}) + O_p(N^{1/2}T^{-1}) \tag{B.51}$$

which further means that $\sup_{1\leq i\leq N} \|\tilde{\mathbf{B}}_{i,T}^{-1}\| = O_p(1)$ by triangular inequality with $\sup_{1\leq i\leq N} \|\mathbf{B}_{i,T}^{-1}\| = O(1)$. Then, analogous to the argument in (B.50) and (B.51), we have

$$\sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{B}}}_{i,T}^{-1} - \tilde{\mathbf{B}}_{i,T}^{-1}\| = O_p(N^{1/2}\delta_{NT}^{-2}) \tag{B.52}$$

By Lemma 14 with Assumption 6 (iii), $\sup_{1\leq i\leq N} |\eta_{ir}| = O_p(\ln N)$ for $2 \leq r \leq k+1$, which further implies that $\sup_{1\leq i\leq N} |\beta_{ir}| = O_p(\ln N)$ for $1 \leq r \leq k$. Then, we have $\sup_{1\leq i\leq N} \|\boldsymbol{\beta}_i\| \leq \sup_{1\leq i\leq N} \sum_{r=1}^k |\beta_{ir}| \leq \sum_{r=1}^k (\sup_{1\leq i\leq N} |\beta_{ir}|) = O_p(\ln N)$. Note that

$$\mathbf{A}_{i,T} = \begin{pmatrix} T^{-1}\sum_{s=1}^{\infty} \rho_i^{s-1} E(\mathbf{V}_i'\mathbf{V}_{i,-s}) & T^{-1}E(\mathbf{V}_i'\mathbf{V}_i) \\ T^{-1}\sum_{s=1}^{\infty} \rho_i^{s-1} E(\mathbf{V}_{i,-1}'\mathbf{V}_{i,-s}) & T^{-1}E(\mathbf{V}_{i,-1}'\mathbf{V}_i) \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_k \end{pmatrix}$$

it's easy to show that $\sup_{1\leq i\leq N} \|\mathbf{A}_{i,T}\| = O_p(\ln N)$ with $\rho_{\max} < 1$ and $\sup_{1\leq i\leq N} \|\boldsymbol{\beta}_i\| = O_p(\ln N)$. With (B.49) and the triangular inequality, we have

$$\sup_{1\leq i\leq N} \|\tilde{\mathbf{A}}_{i,T}\| = O_p(\ln N) \tag{B.53}$$

Note that $\sup_{1\leq i\leq N} \|\tilde{\mathbf{B}}_{i,T}^{-1}\| = O_p(1)$, with (B.49), (B.52) and (B.53), we have

$$\begin{aligned}
&\sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{A}}}_{i,T}'\hat{\tilde{\mathbf{B}}}_{i,T}^{-1}\hat{\tilde{\mathbf{A}}}_{i,T} - \tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T}\| \\
\leq& 2\sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{A}}}_{i,T} - \tilde{\mathbf{A}}_{i,T}\| \sup_{1\leq i\leq N} \|\tilde{\mathbf{A}}_{i,T}\| \sup_{1\leq i\leq N} \|\tilde{\mathbf{B}}_{i,T}^{-1}\| + \sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{B}}}_{i,T}^{-1} - \tilde{\mathbf{B}}_{i,T}^{-1}\| \sup_{1\leq i\leq N} \|\tilde{\mathbf{A}}_{i,T}\|^2 \\
&+ \sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{A}}}_{i,T} - \tilde{\mathbf{A}}_{i,T}\|^2 \sup_{1\leq i\leq N} \|\tilde{\mathbf{B}}_{i,T}^{-1}\| + 2\sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{A}}}_{i,T} - \tilde{\mathbf{A}}_{i,T}\| \sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{B}}}_{i,T}^{-1} - \tilde{\mathbf{B}}_{i,T}^{-1}\| \sup_{1\leq i\leq N} \|\tilde{\mathbf{A}}_{i,T}\| \\
&+ \sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{A}}}_{i,T} - \tilde{\mathbf{A}}_{i,T}\|^2 \sup_{1\leq i\leq N} \|\hat{\tilde{\mathbf{B}}}_{i,T}^{-1} - \tilde{\mathbf{B}}_{i,T}^{-1}\| \\
=&O_p(N^{1/2}(\ln N)^2\delta_{NT}^{-2})
\end{aligned} \tag{B.54}$$

Analogously, with (B.49) and (B.51), we have

$$\sup_{1\leq i\leq N} \|\tilde{\mathbf{A}}_{i,T}'\tilde{\mathbf{B}}_{i,T}^{-1}\tilde{\mathbf{A}}_{i,T} - \mathbf{A}_{i,T}'\mathbf{B}_{i,T}^{-1}\mathbf{A}_{i,T}\| = O_p(N^{1/4}(\ln N)^2T^{-1/2}) + O_p(N^{1/2}(\ln N)^2T^{-1})$$

Suppose $\sup_{1\leq i\leq N}\|(\mathbf{A}'_{i,T}\mathbf{B}^{-1}_{i,T}\mathbf{A}_{i,T})^{-1}\| = O_p(\ln N)$, we can follow the argument in (B.50) to show that

$$\sup_{1\leq i\leq N}\|(\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\tilde{\mathbf{A}}_{i,T})^{-1} - (\mathbf{A}'_{i,T}\mathbf{B}^{-1}_{i,T}\mathbf{A}_{i,T})^{-1}\| = O_p(N^{1/4}(\ln N)^4T^{-1/2}) + O_p(N^{1/2}(\ln N)^4T^{-1}) \qquad \text{(B.55)}$$

given $N^{1+\delta}/T^2 \to 0$ for any $\delta > 0$. Then with triangular inequality, $\sup_{1\leq i\leq N}\|(\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\tilde{\mathbf{A}}_{i,T})^{-1}\| = O_p(\ln N)$. Thus, we have

$$\sup_{1\leq i\leq N}\|(\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\| = O_p((\ln N)^2) \qquad \text{(B.56)}$$

Then, we prove (b).

In addition, with (B.54), using the argument in (B.50) again, we can show that

$$\sup_{1\leq i\leq N}\|(\hat{\tilde{\mathbf{A}}}'_{i,T}\hat{\tilde{\mathbf{B}}}^{-1}_{i,T}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1} - (\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\tilde{\mathbf{A}}_{i,T})^{-1}\| = O_p(N^{1/2}(\ln N)^4\delta_{NT}^{-2}) \qquad \text{(B.57)}$$

and $\sup_{1\leq i\leq N}\|(\hat{\tilde{\mathbf{A}}}'_{i,T}\hat{\tilde{\mathbf{B}}}^{-1}_{i,T}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\| = O_p(\ln N)$. Combining (B.52), (B.49) and (B.57), we can derive that

$$\begin{aligned}&\sup_{1\leq i\leq N}\|(\hat{\tilde{\mathbf{A}}}'_{i,T}\hat{\tilde{\mathbf{B}}}^{-1}_{i,T}\hat{\tilde{\mathbf{A}}}_{i,T})^{-1}\hat{\tilde{\mathbf{A}}}'_{i,T}\hat{\tilde{\mathbf{B}}}^{-1}_{i,T} - (\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\tilde{\mathbf{A}}_{i,T})^{-1}\tilde{\mathbf{A}}'_{i,T}\tilde{\mathbf{B}}^{-1}_{i,T}\|\\ =&O_p(N^{1/4}T^{-1/2}\ln N) + O_p(N^{1/2}(\ln N)^5\delta_{NT}^{-2})\end{aligned}$$

Thus, we prove (a). With (B.49), (B.51) and (B.55), we follow the argument in the proof of (a) to prove (c). This complete the proof. ■

Appendix C: Additional Experimental Results

Table C1: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, MGIV^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.8, 3, 1\}$ and $\pi_u = 3/4$, correlated factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.8, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	0.0	-0.1	-0.1	0.0	-0.6	-0.8	-1.0	-1.0	-1.5	-1.5	-1.5	-1.5	-5.8	-6.1	-6.1	-6.2
50	-0.1	0.0	0.0	0.0	-0.2	-0.3	-0.4	-0.5	-0.8	-0.7	-0.7	-0.7	-2.1	-2.1	-2.3	-2.4
100	-0.1	0.0	0.0	0.0	0.0	-0.1	-0.2	-0.2	-0.5	-0.4	-0.4	-0.4	-0.4	-0.6	-0.7	-0.9
200	0.0	0.0	0.0	0.0	0.1	0.0	-0.1	-0.1	-0.3	-0.3	-0.3	-0.3	0.2	0.1	-0.1	-0.2
	RMSE (×100)															
25	5.2	3.6	2.6	1.8	1.5	1.3	1.4	1.2	4.2	3.2	2.6	2.2	6.7	6.7	6.5	6.5
50	3.0	2.1	1.4	1.0	0.8	0.6	0.6	0.5	3.0	2.1	1.7	1.3	2.6	2.4	2.5	2.5
100	2.0	1.3	0.9	0.6	0.5	0.4	0.3	0.3	2.2	1.5	1.1	0.9	1.0	0.9	0.9	1.0
200	1.3	0.8	0.6	0.4	0.3	0.2	0.2	0.1	1.4	1.0	0.8	0.6	0.7	0.5	0.4	0.4
	SIZE: $H_0 : \rho = 0.8$ against $H_1 : \rho \neq 0.8$, at the 5% level															
25	7.5	6.2	4.6	5.3	25.0	34.1	51.8	72.4	8.7	9.7	16.1	25.8	57.9	83.4	96.5	99.4
50	8.3	6.3	6.3	5.7	16.3	20.6	34.3	56.0	5.9	6.2	7.6	10.7	38.3	61.3	86.8	97.4
100	9.1	6.5	5.0	4.8	13.2	14.2	22.7	39.5	6.5	6.4	6.4	8.5	16.1	31.5	54.1	82.5
200	8.4	5.6	5.9	5.1	12.8	11.0	14.4	21.4	5.9	5.2	6.4	8.0	16.6	20.0	25.0	44.4
	POWER (size-adjusted) : $H_0 : \rho = 0.9$ against $H_1 : \rho \neq 0.9$, at the 5% level															
25	66.1	81.4	95.5	99.8	99.6	99.9	99.7	100.0	61.3	79.8	92.8	98.6	13.6	7.9	5.6	2.3
50	89.6	98.8	100.0	100.0	100.0	100.0	100.0	100.0	86.5	96.5	99.7	99.9	93.0	98.8	99.7	99.9
100	99.2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	95.6	99.8	100.0	100.0	100.0	100.0	100.0	100.0
200	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.7	100.0	100.0	100.0	100.0	100.0	100.0	100.0

PANEL B: Results for β_1, homogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.8, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	0.3	0.6	0.1	0.1	-8.2	-7.9	-8.4	-6.4	1.7	2.9	2.0	2.3	-1.0	-0.5	-0.3	0.1
50	0.3	0.2	0.0	-0.2	-9.1	-7.2	-4.6	-2.8	1.8	2.2	2.1	1.7	0.0	0.6	1.1	1.3
100	0.3	-0.1	0.0	0.1	-7.8	-5.9	-2.7	-0.4	1.5	0.7	1.0	1.1	-1.4	-1.3	-0.2	0.5
200	-0.1	0.1	0.0	0.0	-7.7	-4.4	-1.2	0.1	0.4	0.4	0.5	0.5	-3.0	-2.5	-1.6	-0.7
	RMSE (×100)															
25	22.1	15.9	11.0	7.7	28.7	24.0	21.8	17.0	33.3	24.0	16.3	12.0	30.7	22.9	16.4	11.3
50	15.0	10.1	7.2	4.9	26.8	22.2	16.7	11.8	22.0	14.6	10.8	7.5	18.1	12.7	9.1	6.3
100	10.4	7.1	5.0	3.4	26.3	19.1	11.7	6.2	13.9	9.4	6.9	4.8	11.6	8.2	5.8	3.9
200	7.2	5.0	3.4	2.5	25.0	16.1	7.6	2.9	8.9	6.2	4.3	3.1	8.7	6.3	4.0	2.7
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	8.4	6.9	5.1	5.4	58.2	55.9	55.8	49.5	5.7	5.9	5.2	5.8	7.6	7.7	6.9	5.1
50	7.5	5.7	5.9	4.7	66.1	58.7	45.2	29.7	5.5	5.6	5.6	5.6	8.1	6.8	6.0	5.8
100	8.0	5.9	6.9	5.1	69.4	56.1	33.0	15.8	5.4	5.0	5.3	5.2	8.4	7.5	7.5	6.2
200	8.5	5.9	5.2	7.3	74.4	52.7	27.1	11.5	6.2	4.5	4.5	4.9	11.4	11.9	8.2	8.3
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	13.4	12.6	18.0	29.2	3.6	2.9	2.0	2.4	7.4	9.1	12.5	19.3	5.9	6.8	8.9	17.7
50	15.4	21.7	30.6	52.8	2.7	3.1	2.8	3.5	10.5	14.6	22.2	35.3	8.7	14.0	25.8	43.2
100	22.7	33.6	54.0	85.9	3.1	3.0	4.8	63.9	16.2	24.1	40.0	66.2	11.9	18.1	40.9	77.6
200	36.9	57.3	84.0	97.8	2.9	3.5	35.8	96.4	23.1	45.5	71.3	92.0	11.2	20.9	55.2	91.7

Notes: The DGP is the same as the one for Table 3, except $\{\rho, \beta_1, \beta_2\} = \{0.8, 3, 1\}$.

Table C2: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, MGIV^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.8, 3, 1\}$ and $\pi_u = 3/4$, correlated factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.8, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	-0.3	-0.6	-0.5	-0.6	-0.1	-1.0	-2.8	-3.9	-1.5	-1.4	-1.4	-1.5	-5.3	-5.5	-5.8	-6.0
50	-0.3	-0.1	-0.1	-0.1	5.5	5.7	5.5	5.8	-0.8	-0.7	-0.7	-0.7	-1.6	-1.9	-2.1	-2.2
100	0.0	0.1	0.2	0.3	9.6	12.2	13.9	14.7	-0.4	-0.5	-0.5	-0.4	0.0	-0.3	-0.6	-0.7
200	0.3	0.6	0.6	0.7	12.6	15.6	16.9	17.5	-0.3	-0.3	-0.2	-0.3	0.7	0.5	0.2	0.1
	RMSE ($\times$100)															
25	7.7	5.5	3.9	2.7	7.0	6.5	6.8	7.2	4.6	3.5	2.8	2.3	6.5	6.3	6.3	6.3
50	5.4	3.6	2.5	1.8	8.6	8.4	8.1	8.3	3.9	2.7	2.0	1.6	3.2	2.9	2.6	2.6
100	3.7	2.6	1.9	1.3	11.6	13.3	14.5	15.1	3.1	2.3	1.6	1.2	2.5	1.8	1.4	1.3
200	3.1	2.3	1.7	1.3	14.0	16.1	17.1	17.6	2.7	2.0	1.4	1.0	2.4	1.8	1.3	0.9
	SIZE: $H_0 : \rho = 0.8$ against $H_1 : \rho \neq 0.8$, at the 5% level															
25	10.9	8.9	6.8	6.5	73.2	75.1	80.1	85.9	6.6	7.5	11.5	19.3	34.1	57.5	84.1	95.5
50	12.3	7.7	6.6	6.0	81.6	84.6	86.0	89.5	6.9	5.3	6.6	9.7	11.6	21.2	37.2	62.1
100	9.1	7.5	5.8	6.1	92.6	97.3	98.9	99.7	5.6	5.3	5.7	7.0	8.7	8.2	10.0	19.6
200	9.1	8.0	9.1	11.4	97.1	99.8	100.0	100.0	5.5	5.4	5.9	6.0	9.2	7.8	8.2	8.7
	POWER (size-adjusted) : $H_0 : \rho = 0.9$ against $H_1 : \rho \neq 0.9$, at the 5% level															
25	39.0	52.5	72.6	94.2	31.7	38.2	32.7	30.6	50.4	73.4	89.0	97.6	10.1	6.1	4.0	1.9
50	52.2	78.3	95.3	99.9	20.8	30.7	35.0	36.1	66.6	93.0	99.4	100.0	75.2	87.9	96.1	99.2
100	71.3	93.0	99.6	100.0	17.2	25.4	34.7	43.0	84.6	97.5	99.9	100.0	97.2	100.0	100.0	100.0
200	84.2	98.5	100.0	100.0	15.9	19.9	24.4	35.4	92.7	99.7	100.0	100.0	99.4	100.0	100.0	100.0
PANEL B: Results for β_1, heterogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.8, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	-5.0	-5.5	-6.1	-6.1	-16.1	-18.5	-20.4	-23.1	1.8	1.2	1.8	1.9	-1.4	-0.8	-1.2	-0.5
50	-5.6	-6.1	-6.6	-6.8	-22.7	-26.6	-29.5	-31.5	2.0	1.0	1.4	1.5	-0.6	-0.4	0.4	1.2
100	-5.5	-5.4	-6.1	-6.5	-34.2	-48.8	-59.8	-65.2	0.6	0.9	0.8	0.8	-2.1	-0.8	-0.3	0.1
200	-4.4	-4.8	-5.2	-5.4	-48.8	-67.6	-78.9	-84.0	0.3	0.5	0.5	0.2	-3.3	-2.1	-1.3	-0.8
	RMSE ($\times$100)															
25	28.6	20.2	14.7	11.4	34.3	31.5	31.0	31.9	33.4	23.8	16.4	11.8	31.3	23.0	16.5	11.6
50	20.8	15.1	11.6	9.6	40.2	41.0	41.8	42.9	21.5	14.9	10.7	7.7	18.0	13.4	9.4	6.9
100	15.5	11.7	9.6	8.3	51.8	62.4	69.0	72.0	14.0	9.7	6.9	4.8	12.1	8.7	6.2	4.3
200	12.0	9.2	7.7	6.7	65.6	78.0	84.6	87.4	9.3	6.9	4.8	3.2	9.3	6.5	4.4	3.2
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	9.2	9.5	8.1	11.1	63.0	65.8	73.2	81.3	5.7	5.3	4.4	5.4	7.6	6.1	5.5	5.4
50	9.7	9.6	10.6	17.2	73.8	77.8	82.9	86.4	6.5	4.7	5.3	5.7	6.9	6.2	6.4	7.4
100	10.2	9.7	13.9	24.4	82.5	89.8	94.8	97.2	5.6	5.5	5.3	4.5	8.3	8.0	7.4	7.0
200	10.6	11.1	16.1	28.3	90.2	95.5	98.5	99.4	5.8	6.5	6.3	4.6	10.7	10.4	8.5	8.1
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	8.6	7.2	5.2	4.5	2.3	2.1	1.5	1.2	7.3	8.9	13.2	16.3	5.2	6.9	8.5	15.1
50	6.9	7.8	5.8	2.7	2.2	1.9	2.4	1.6	9.3	14.7	20.8	33.7	8.4	13.6	20.9	36.0
100	8.2	8.5	7.8	5.0	2.6	2.4	1.5	1.9	14.9	22.9	34.9	64.7	10.9	19.2	35.5	63.8
200	11.1	11.8	11.9	13.1	2.4	1.7	2.1	1.2	21.0	35.5	61.0	89.2	8.7	22.3	50.4	83.0

Notes: The DGP is the same as the one for Table 4, except $\{\rho, \beta_1, \beta_2\} = \{0.8, 3, 1\}$.

Table C3: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, MGIV^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 0\}$ and $\pi_u = 3/4$, correlated factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 0\}$ and $\pi_u = 3/4$

	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	0.1	0.0	0.0	0.0	-0.4	-0.7	-0.9	-1.1	-0.7	-0.7	-0.6	-0.7	-3.4	-3.6	-3.8	-3.9
50	0.0	0.0	0.0	0.0	0.0	-0.3	-0.5	-0.6	-0.5	-0.4	-0.4	-0.3	-0.9	-1.1	-1.3	-1.5
100	0.0	0.0	0.0	0.0	0.1	-0.1	-0.3	-0.3	-0.3	-0.2	-0.2	-0.2	0.3	0.1	-0.2	-0.4
200	0.0	0.0	0.0	0.0	0.2	0.0	-0.1	-0.2	-0.1	-0.1	-0.1	-0.1	0.8	0.6	0.3	0.1
	RMSE (×100)															
25	3.1	2.2	1.6	1.1	1.7	1.4	1.4	1.3	3.5	2.5	1.8	1.5	4.5	4.3	4.2	4.2
50	2.1	1.4	1.0	0.7	1.1	0.9	0.7	0.7	2.3	1.5	1.2	0.9	1.9	1.7	1.6	1.7
100	1.4	1.0	0.6	0.4	0.8	0.6	0.5	0.4	1.5	1.0	0.7	0.6	1.1	0.9	0.7	0.7
200	1.0	0.6	0.4	0.3	0.6	0.4	0.3	0.2	1.0	0.7	0.5	0.4	1.2	0.9	0.6	0.4
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	8.3	6.0	6.5	5.7	20.2	25.9	41.3	60.8	5.2	7.2	7.9	12.5	31.7	54.5	78.8	94.5
50	9.2	6.2	6.3	4.8	16.4	18.3	27.0	48.4	6.3	5.3	6.1	8.3	17.1	29.5	53.2	78.1
100	9.0	6.9	5.3	4.7	14.0	14.6	20.0	31.4	5.8	6.4	5.7	6.6	13.3	17.8	26.5	44.3
200	9.0	6.0	5.8	5.2	16.6	11.4	12.9	18.2	6.0	5.3	5.5	6.0	31.0	34.4	31.2	28.6
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	88.4	98.6	100.0	100.0	99.8	100.0	100.0	100.0	80.2	93.9	99.4	100.0	39.8	41.2	47.7	52.1
50	99.2	100.0	100.0	100.0	100.0	100.0	100.0	100.0	96.2	99.8	100.0	100.0	99.3	100.0	100.0	100.0
100	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0
200	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

PANEL B: Results for β_2, homogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 0\}$ and $\pi_u = 3/4$

	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	0.2	0.4	0.1	0.0	11.0	9.4	7.6	5.6	0.1	0.1	0.2	0.0	4.0	1.9	1.4	0.4
50	0.1	0.1	-0.1	0.0	9.2	7.1	3.9	1.9	0.3	0.2	-0.1	-0.1	4.2	2.3	1.0	0.5
100	0.0	0.1	0.0	0.0	7.9	4.8	2.0	0.3	0.0	0.1	0.0	0.0	3.7	2.2	1.1	0.6
200	-0.1	0.1	0.0	0.0	6.5	2.9	0.6	0.1	0.0	0.1	0.0	0.0	3.8	2.2	1.1	0.5
	RMSE (×100)															
25	10.7	7.6	5.5	4.0	17.9	14.7	11.9	9.0	15.7	10.9	8.1	5.6	16.1	11.3	8.2	5.6
50	7.4	5.0	3.6	2.5	16.0	12.2	8.2	5.3	9.2	6.5	4.6	3.2	10.1	6.7	4.5	3.2
100	5.0	3.5	2.4	1.6	14.0	9.6	5.9	2.5	6.2	4.2	2.9	2.1	7.1	4.6	2.9	2.0
200	3.5	2.4	1.7	1.2	12.7	7.2	3.2	1.6	4.0	2.8	2.0	1.4	5.7	3.5	2.2	1.4
	SIZE: $H_0 : \beta_2 = 0$ against $H_1 : \beta_2 \neq 0$, at the 5% level															
25	7.6	6.9	6.1	6.0	51.7	51.6	48.8	43.6	5.0	4.7	5.3	5.5	6.9	5.4	5.7	5.3
50	10.3	6.6	5.4	5.7	55.7	47.4	34.0	21.1	5.6	5.3	5.6	4.7	10.3	7.1	5.9	5.8
100	8.5	6.1	6.2	5.5	58.6	39.0	20.6	9.4	6.4	5.4	5.7	4.8	12.6	10.4	7.0	6.2
200	8.7	6.4	6.2	5.9	57.9	32.2	11.4	7.3	6.2	5.2	5.3	4.7	19.9	14.5	9.9	8.1
	POWER (size-adjusted) : $H_0 : \beta_2 = 0.0$ against $H_1 : \beta_2 \neq 0.1$, at the 5% level															
25	23.4	32.2	47.8	74.1	14.5	17.4	21.6	24.6	11.3	17.1	27.3	44.6	16.1	22.4	31.6	48.0
50	36.4	58.1	78.8	97.1	16.5	14.8	17.7	39.4	23.4	38.4	59.9	86.6	26.9	49.2	69.5	90.0
100	58.1	83.6	98.4	100.0	16.5	13.8	15.9	100.0	40.4	68.6	91.3	99.9	50.3	77.7	96.2	100.0
200	82.8	97.9	100.0	100.0	14.2	11.4	99.9	100.0	68.5	93.5	100.0	100.0	75.2	96.2	100.0	100.0

Notes: The DGP is the same as the one for Table 3, except $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 0\}$.

Table C4: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, MGIV^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 0\}$ and $\pi_u = 3/4$, correlated factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 0\}$ and $\pi_u = 3/4$

	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	0.9	1.0	1.0	0.9	1.0	0.9	0.7	0.3	-0.8	-0.7	-0.6	-0.7	-3.4	-3.5	-3.7	-3.9
50	1.0	1.1	1.2	1.2	2.4	2.2	2.1	2.1	-0.5	-0.3	-0.3	-0.3	-0.8	-1.0	-1.2	-1.4
100	1.1	1.2	1.2	1.3	2.9	2.9	2.8	2.8	-0.2	-0.3	-0.3	-0.2	0.4	0.1	-0.2	-0.4
200	1.2	1.3	1.3	1.3	3.2	3.1	3.0	3.1	-0.1	-0.1	-0.1	-0.1	0.9	0.7	0.4	0.1
	RMSE ($\times$100)															
25	4.3	3.2	2.3	1.8	4.4	3.2	2.5	2.0	4.1	3.0	2.2	1.6	5.0	4.5	4.3	4.3
50	3.5	2.6	2.0	1.6	4.3	3.3	2.8	2.5	3.2	2.2	1.6	1.2	3.0	2.4	2.0	1.9
100	3.0	2.3	1.8	1.6	4.4	3.6	3.2	3.0	2.7	2.0	1.4	1.0	2.6	1.9	1.4	1.2
200	2.8	2.2	1.8	1.6	4.6	3.8	3.4	3.3	2.5	1.8	1.3	0.9	2.6	1.9	1.4	1.0
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	11.9	9.5	9.2	11.0	53.1	50.7	51.8	55.8	5.8	5.6	7.3	8.8	20.1	33.7	57.0	82.4
50	13.1	11.4	13.2	20.4	65.2	68.0	75.8	83.6	6.4	5.6	6.3	6.6	10.0	12.5	20.2	37.1
100	11.5	13.0	16.0	27.2	76.6	82.0	90.2	97.3	5.2	5.3	6.0	5.9	8.7	7.8	9.4	14.2
200	12.4	14.9	22.2	33.2	84.7	90.0	95.5	99.5	5.3	5.3	5.1	5.2	9.4	9.8	10.5	10.6
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	74.0	92.6	99.6	100.0	64.5	88.8	96.9	98.9	62.6	88.4	98.4	99.9	25.4	33.3	36.9	40.5
50	87.8	99.0	100.0	100.0	78.3	97.6	99.8	100.0	81.7	98.4	100.0	100.0	85.6	98.2	100.0	100.0
100	96.4	100.0	100.0	100.0	83.7	98.9	100.0	100.0	94.0	99.7	100.0	100.0	97.6	100.0	100.0	100.0
200	98.6	100.0	100.0	100.0	86.1	99.2	100.0	100.0	97.4	99.9	100.0	100.0	99.1	100.0	100.0	100.0

PANEL B: Results for β_2, heterogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 0\}$ and $\pi_u = 3/4$

	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	0.2	0.2	0.0	0.1	10.6	10.6	9.4	7.9	0.5	0.0	-0.2	-0.1	3.9	2.2	0.8	0.4
50	0.1	-0.1	0.0	0.0	9.6	8.5	6.5	4.6	-0.2	-0.2	-0.1	-0.2	3.8	2.0	1.1	0.5
100	0.1	-0.1	0.0	0.1	8.6	7.1	4.7	3.3	0.2	-0.1	0.1	0.0	3.9	2.0	1.1	0.6
200	0.0	0.1	0.0	0.1	7.8	5.3	3.7	3.1	0.0	0.0	0.0	0.1	3.8	2.1	1.1	0.6
	RMSE ($\times$100)															
25	12.6	9.0	6.4	4.7	19.1	15.6	13.3	10.8	15.9	11.2	8.1	5.8	16.2	11.4	8.1	5.8
50	8.4	6.2	4.2	3.0	16.8	13.3	9.9	6.9	9.5	6.8	4.7	3.3	10.0	6.9	4.8	3.2
100	6.2	4.4	3.1	2.1	15.3	10.9	7.3	4.8	6.3	4.4	3.1	2.2	7.3	4.8	3.2	2.2
200	4.6	3.3	2.3	1.6	13.8	8.9	5.5	4.1	4.6	3.3	2.3	1.6	6.0	3.9	2.5	1.6
	SIZE: $H_0 : \beta_2 = 0$ against $H_1 : \beta_2 \neq 0$, at the 5% level															
25	8.6	6.9	6.2	7.1	51.0	54.8	57.2	56.2	5.3	5.2	5.8	6.4	6.5	6.4	5.3	5.4
50	7.7	6.7	5.3	4.5	56.3	55.0	49.8	47.6	5.3	6.0	4.8	5.0	7.6	7.4	6.8	5.0
100	8.4	6.8	5.7	4.9	61.8	55.2	47.2	48.1	6.0	5.1	5.6	5.1	11.3	8.6	7.7	6.2
200	8.1	7.6	5.5	5.7	64.8	55.4	52.2	58.4	5.2	5.8	6.1	4.6	15.2	12.5	8.9	6.7
	POWER (size-adjusted) : $H_0 : \beta_2 = 0.0$ against $H_1 : \beta_2 \neq 0.1$, at the 5% level															
25	18.2	23.5	36.6	58.4	14.6	17.7	22.6	26.4	12.1	16.6	24.3	40.2	16.3	19.3	28.9	46.8
50	28.0	39.9	66.6	91.9	15.5	18.3	25.2	43.0	19.1	30.8	57.8	84.4	28.8	41.7	64.7	90.8
100	44.2	59.2	91.0	99.8	16.3	21.8	36.1	87.4	36.2	60.9	87.7	99.3	46.6	71.2	93.6	99.9
200	60.5	86.7	99.1	100.0	14.6	19.9	71.7	98.3	58.0	85.8	98.8	100.0	65.3	90.2	99.9	100.0

Notes: The DGP is the same as the one for Table 4, except $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 0\}$.

Table C5: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, MGIV^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 1/4$, correlated factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 1/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	0.0	0.0	0.0	0.0	-0.1	-0.2	-0.2	-0.2	0.4	0.4	0.4	0.5	-0.7	-0.9	-1.1	-1.2
50	0.0	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1	0.3	0.2	0.2	0.2	0.4	0.1	-0.2	-0.4
100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	1.0	0.6	0.2	0.0
200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.8	0.4	0.1
	RMSE ($\times$100)															
25	2.5	1.5	0.9	0.6	0.6	0.5	0.3	0.3	2.9	2.1	1.7	1.4	2.4	1.9	1.6	1.5
50	1.3	0.8	0.5	0.4	0.4	0.3	0.2	0.2	1.9	1.4	1.0	0.8	1.5	1.2	0.8	0.6
100	0.8	0.5	0.4	0.3	0.2	0.2	0.1	0.1	1.4	0.9	0.7	0.5	1.5	1.1	0.7	0.3
200	0.5	0.4	0.3	0.2	0.2	0.1	0.1	0.1	0.9	0.6	0.5	0.3	1.6	1.2	0.7	0.3
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	10.9	6.6	5.7	5.8	11.8	14.1	13.9	22.0	7.4	8.1	12.0	19.1	19.5	27.5	43.3	65.3
50	12.3	7.6	5.8	6.2	9.2	8.5	10.3	15.2	6.3	6.2	7.6	8.2	20.3	25.8	34.5	48.3
100	10.9	7.3	5.7	6.4	7.7	8.3	8.7	9.5	5.8	4.9	4.9	5.0	39.3	38.6	33.7	32.7
200	11.4	7.2	6.2	6.0	7.2	7.2	6.5	7.3	4.6	5.0	5.2	5.2	58.1	60.1	49.4	36.0
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	95.2	99.8	100.0	100.0	100.0	100.0	100.0	100.0	92.5	99.4	99.9	100.0	94.5	99.6	100.0	100.0
50	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	98.8	99.9	100.0	100.0	100.0	100.0	100.0	100.0
100	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
200	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
PANEL B: Results for β_1, homogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 1/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	-0.2	-0.1	-0.1	0.0	-0.8	-0.2	0.1	0.2	0.9	0.8	0.7	0.9	-4.9	-3.4	-2.0	-0.9
50	0.1	0.1	0.0	0.0	-0.3	0.0	0.2	0.2	0.3	0.0	-0.1	-0.1	-6.0	-4.5	-2.8	-1.1
100	0.0	-0.1	0.0	0.0	-0.2	0.0	0.1	0.1	0.0	-0.5	-0.2	-0.3	-7.3	-5.5	-3.4	-1.7
200	0.1	0.0	0.0	0.0	-0.2	0.0	0.1	0.0	0.0	-0.1	-0.2	-0.1	-7.9	-6.3	-3.9	-1.9
	RMSE ($\times$100)															
25	8.3	5.4	3.6	2.4	10.5	5.9	3.8	2.6	16.1	11.2	7.9	5.8	15.0	9.7	5.9	3.8
50	5.0	3.4	2.3	1.6	6.6	3.6	2.2	1.5	10.2	7.0	4.9	3.6	11.2	7.7	4.7	2.5
100	3.4	2.3	1.6	1.1	4.5	2.3	1.5	1.0	6.3	4.4	3.2	2.3	10.6	7.1	4.3	2.3
200	2.3	1.6	1.1	0.8	3.2	1.6	1.0	0.7	4.1	2.9	2.1	1.4	10.5	7.4	4.5	2.3
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	8.8	6.9	6.2	5.1	23.1	13.2	9.9	8.0	4.9	5.4	5.3	5.5	12.6	11.4	8.5	6.7
50	8.1	8.2	6.3	5.9	16.2	8.9	7.8	6.9	5.3	5.5	4.2	5.9	19.2	23.5	21.5	12.2
100	9.7	7.2	5.2	6.3	14.0	7.2	6.9	5.7	5.0	5.5	5.6	5.1	34.0	43.0	41.8	29.9
200	10.1	7.1	5.5	6.0	14.4	7.4	6.3	5.0	4.9	4.3	4.6	4.8	50.0	68.9	72.5	55.3
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	36.9	56.8	83.4	98.3	12.1	48.7	83.9	98.4	13.5	17.9	30.2	51.7	6.3	10.9	29.3	70.7
50	64.2	87.6	99.0	100.0	43.3	85.3	99.0	100.0	22.7	33.3	57.5	78.7	6.1	11.1	40.6	95.2
100	86.3	98.5	100.0	100.0	75.8	98.6	100.0	100.0	39.1	57.9	85.7	98.4	4.9	7.8	45.8	97.9
200	98.6	100.0	100.0	100.0	91.8	99.9	100.0	100.0	67.1	91.7	99.5	100.0	6.6	9.7	48.1	97.9

Notes: The DGP is the same as the one for Table 3, except $\pi_u = 1/4$.

Table C6: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, MGIV^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 1/4$, correlated factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 1/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	0.8	0.7	0.8	0.9	1.5	1.7	1.6	1.7	0.5	0.4	0.4	0.4	-0.7	-0.8	-1.0	-1.1
50	1.0	0.9	1.1	1.1	2.4	2.4	2.6	2.5	0.2	0.1	0.3	0.2	0.5	0.2	-0.1	-0.3
100	1.0	1.2	1.2	1.3	2.6	2.7	2.9	3.0	0.0	0.1	0.0	0.1	1.2	0.8	0.3	0.0
200	1.0	1.2	1.3	1.3	2.8	3.0	3.1	3.1	-0.1	0.0	0.0	0.0	1.4	0.9	0.5	0.2
	RMSE (×100)															
25	4.3	3.0	2.2	1.7	4.6	3.6	3.0	2.8	3.6	2.7	2.1	1.6	3.5	2.9	2.3	1.9
50	3.5	2.4	1.9	1.5	4.8	3.8	3.4	3.0	3.0	2.1	1.5	1.2	3.1	2.3	1.8	1.4
100	2.9	2.3	1.8	1.6	4.7	3.8	3.4	3.3	2.7	1.9	1.3	1.0	3.2	2.4	1.8	1.3
200	2.8	2.2	1.8	1.6	4.8	3.9	3.6	3.3	2.5	1.8	1.2	0.9	3.2	2.4	1.8	1.3
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	11.0	9.0	8.8	12.2	65.0	65.5	72.3	78.2	6.3	8.0	9.2	12.2	12.9	18.5	24.5	33.9
50	13.2	11.2	12.6	19.6	76.7	77.4	86.0	91.9	6.6	5.7	5.5	7.3	12.8	14.3	17.1	21.4
100	13.4	14.3	18.7	29.8	84.0	85.4	92.8	98.1	6.1	5.7	5.3	5.7	15.4	16.6	19.7	21.7
200	13.5	15.4	21.6	36.9	87.9	90.9	95.9	98.7	5.9	6.3	5.6	4.5	16.6	18.9	20.0	19.6
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	76.1	94.4	99.4	100.0	62.9	86.9	96.6	99.0	82.5	96.0	99.8	100.0	74.0	87.9	98.3	99.8
50	90.3	99.1	100.0	100.0	64.7	93.9	99.6	100.0	89.4	99.6	100.0	100.0	91.5	99.4	100.0	100.0
100	96.5	100.0	100.0	100.0	72.5	96.4	99.9	100.0	94.3	99.8	100.0	100.0	94.3	99.6	100.0	100.0
200	98.7	100.0	100.0	100.0	73.7	96.9	100.0	100.0	97.4	100.0	100.0	100.0	96.0	99.9	100.0	100.0
PANEL B: Results for β_1, heterogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 1/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	-1.4	-1.6	-1.2	-1.5	-6.6	-7.3	-6.7	-6.4	1.1	0.7	1.0	0.7	-4.1	-3.2	-2.0	-0.9
50	-1.3	-1.3	-1.2	-1.3	-7.8	-6.6	-5.8	-5.5	-0.2	-0.1	0.0	-0.1	-6.1	-4.6	-2.7	-1.1
100	-1.0	-1.1	-1.2	-1.1	-6.8	-5.7	-5.3	-4.9	-0.2	-0.2	-0.3	-0.1	-7.2	-5.2	-3.5	-1.6
200	-1.0	-1.1	-1.1	-1.1	-7.2	-5.6	-5.2	-4.7	-0.2	-0.1	-0.2	-0.1	-7.9	-6.1	-3.9	-2.0
	RMSE (×100)															
25	12.8	9.0	6.2	4.7	20.8	17.3	14.4	12.5	16.3	11.5	8.0	5.7	15.3	10.2	6.4	3.9
50	9.2	6.4	4.5	3.4	19.5	15.2	11.6	9.9	10.3	7.2	5.0	3.7	11.8	8.0	4.9	3.0
100	6.6	4.6	3.3	2.4	17.4	12.7	9.8	8.1	6.8	4.6	3.4	2.4	11.0	7.3	4.9	2.8
200	4.9	3.5	2.5	2.0	17.8	12.4	9.1	7.0	4.8	3.3	2.4	1.7	10.9	7.7	5.1	2.9
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	8.2	6.9	5.9	7.1	51.2	50.0	52.8	57.2	5.7	5.8	5.2	5.3	11.4	10.3	8.4	6.4
50	10.0	8.1	6.5	7.8	57.7	55.3	55.8	61.6	6.3	5.0	4.7	5.9	19.1	19.7	17.4	15.3
100	9.4	9.1	8.3	9.5	63.6	60.0	63.7	67.9	5.8	4.7	4.9	4.9	28.9	33.6	36.9	27.1
200	11.5	9.0	10.0	12.1	73.3	68.5	72.5	76.9	5.5	5.0	5.2	4.9	39.5	51.0	51.9	41.4
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	14.5	18.7	31.9	47.4	3.6	2.8	3.0	2.6	10.8	18.3	29.9	49.8	6.2	10.0	26.6	65.4
50	21.4	33.1	53.8	79.1	2.4	2.9	3.2	3.0	19.3	31.1	55.3	76.1	5.2	8.8	34.2	86.3
100	35.2	54.8	80.9	96.6	3.2	3.5	2.2	2.8	33.4	56.9	82.3	97.8	4.0	8.5	29.4	88.4
200	53.5	79.2	95.7	99.8	2.3	3.0	1.7	2.6	54.2	82.7	97.8	100.0	4.0	5.7	28.9	88.5

Notes: The DGP is the same as the one for Table 4, except $\pi_u = 1/4$.

Table C7: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, MGIV^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$, zero mean factor loadings

PANEL A: Results for ρ, homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	-0.1	-0.1	-0.1	-0.1	-0.6	-0.8	-1.0	-1.3	-0.7	-0.8	-0.7	-0.8	-2.6	-2.7	-2.7	-2.8
50	0.0	-0.1	0.0	0.0	0.0	-0.3	-0.6	-0.7	-0.4	-0.4	-0.4	-0.4	-0.3	-0.3	-0.3	-0.3
100	0.0	0.0	0.0	0.0	0.2	-0.1	-0.3	-0.4	-0.2	-0.3	-0.2	-0.2	0.9	0.8	0.8	0.9
200	0.0	0.0	0.0	0.0	0.3	0.0	-0.2	-0.2	-0.1	-0.2	-0.1	-0.1	1.3	1.3	1.3	1.3
	RMSE ($\times$100)															
25	3.1	2.3	1.6	1.1	1.8	1.6	1.5	1.6	3.5	2.6	1.9	1.5	3.9	3.6	3.3	3.3
50	2.0	1.4	1.0	0.7	1.1	0.9	0.9	0.9	2.2	1.5	1.1	0.9	1.7	1.4	1.2	1.1
100	1.4	0.9	0.6	0.4	0.8	0.6	0.5	0.5	1.5	1.0	0.7	0.6	1.4	1.3	1.1	1.1
200	1.0	0.6	0.4	0.3	0.7	0.4	0.3	0.3	1.0	0.7	0.5	0.4	1.6	1.5	1.5	1.4
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	9.4	8.3	6.0	5.2	21.2	27.8	43.5	65.4	6.7	8.1	9.1	12.0	24.1	39.9	58.2	74.8
50	8.3	6.2	6.5	5.8	15.9	18.9	32.2	55.9	5.7	5.7	6.7	8.0	12.1	18.9	25.7	38.1
100	8.9	6.8	5.6	6.5	16.2	16.2	21.7	35.8	5.1	5.3	6.4	7.1	23.6	32.7	45.7	62.9
200	9.4	7.0	5.6	5.5	20.2	13.5	16.5	21.3	5.2	5.5	4.7	6.4	53.9	74.2	90.1	96.1
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	89.7	98.2	100.0	100.0	99.1	99.9	100.0	100.0	76.5	92.1	99.3	100.0	54.5	66.8	78.2	83.9
50	99.5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	96.9	99.7	100.0	100.0	99.9	100.0	100.0	100.0
100	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0
200	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
PANEL B: Results for β_1, homogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	0.3	0.3	0.0	0.1	0.8	0.6	1.2	1.4	1.2	2.0	1.6	1.4	2.7	2.5	2.8	2.5
50	-0.3	0.1	0.1	-0.1	-0.1	0.2	1.0	1.1	0.5	0.8	0.7	0.6	0.6	0.7	0.5	0.5
100	-0.1	-0.1	0.1	0.0	-0.7	0.3	0.5	0.7	0.3	0.3	0.5	0.3	-1.2	-1.3	-1.1	-1.3
200	0.1	0.0	0.0	0.0	-0.7	0.3	0.3	0.3	0.3	0.2	0.2	0.2	-2.3	-2.5	-2.4	-2.4
	RMSE ($\times$100)															
25	11.6	8.4	6.0	4.2	14.6	11.3	9.2	7.1	16.8	11.8	8.5	5.9	16.4	12.1	9.4	7.6
50	7.9	5.7	3.9	2.8	13.0	9.8	6.8	4.4	9.8	6.8	4.9	3.5	10.3	7.9	6.1	4.9
100	5.5	4.0	2.8	1.9	11.7	7.7	4.6	2.4	6.3	4.4	3.1	2.2	8.5	6.1	4.8	4.0
200	4.0	2.6	2.0	1.3	10.3	6.1	2.9	1.5	4.3	2.9	2.2	1.4	7.4	5.8	4.7	3.9
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	9.0	7.3	6.5	5.4	36.1	35.7	35.1	33.8	5.2	5.6	5.9	6.3	8.1	9.2	11.5	16.0
50	8.6	6.5	5.4	5.2	43.4	39.1	28.3	19.5	5.1	4.3	4.7	5.4	7.9	8.8	11.2	15.9
100	8.2	7.4	5.7	5.7	48.6	34.5	18.5	10.4	6.2	4.8	3.9	5.5	9.0	8.6	12.0	18.2
200	9.1	5.4	7.5	5.4	51.1	29.9	12.3	8.1	5.0	4.7	5.6	3.8	9.4	12.1	17.0	26.0
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	20.9	28.0	45.2	70.7	10.9	15.7	19.3	29.3	13.4	19.8	31.6	50.6	12.2	18.3	28.2	39.0
50	32.0	52.3	75.3	94.0	10.9	13.4	22.5	68.3	22.9	40.5	63.8	86.5	18.3	29.3	39.9	60.2
100	51.7	75.3	96.1	99.9	9.4	15.3	68.5	99.6	38.6	67.4	92.1	99.4	18.9	32.0	46.9	56.0
200	77.9	96.2	99.9	100.0	10.6	19.6	98.5	100.0	68.1	91.9	99.7	100.0	16.6	26.7	35.3	50.5

Notes: The DGP is the same as the one for Table 3, except all factgor loadings have mean zero hence the rank condition for CCEMG is not satisfied.

Table C8: Bias, root mean squared error (RMSE) of IV2^b, bias-corrected QMLE, MGIV^b and CCEMG estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$, zero mean factor loadings

PANEL A: Results for ρ, heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	0.5	0.6	0.7	0.8	0.7	0.3	-0.1	-0.4	-0.8	-0.7	-0.8	-0.8	-2.5	-2.7	-2.7	-2.7
50	0.8	0.9	1.0	1.1	2.0	1.6	1.4	1.2	-0.4	-0.4	-0.4	-0.4	-0.3	-0.4	-0.2	-0.3
100	1.0	1.1	1.2	1.2	2.5	2.2	1.9	1.8	-0.3	-0.2	-0.2	-0.2	0.9	0.9	0.8	0.8
200	1.1	1.2	1.3	1.3	2.6	2.4	2.2	2.2	-0.1	-0.1	-0.1	-0.2	1.5	1.4	1.4	1.3
	RMSE ($\times$100)															
25	4.2	3.1	2.2	1.7	4.2	3.3	2.5	2.1	4.2	3.0	2.2	1.7	4.4	3.8	3.5	3.3
50	3.4	2.5	1.9	1.6	4.2	3.0	2.3	1.8	3.3	2.3	1.6	1.2	2.9	2.1	1.6	1.4
100	3.0	2.3	1.9	1.6	4.1	3.3	2.6	2.2	2.7	1.9	1.4	1.0	2.7	2.0	1.6	1.4
200	2.8	2.2	1.8	1.6	4.1	3.2	2.6	2.4	2.4	1.8	1.3	0.9	2.8	2.2	1.9	1.7
	SIZE: $H_0: \rho = 0.5$ against $H_1: \rho \neq 0.5$, at the 5% level															
25	10.8	8.2	8.5	10.6	51.7	49.8	50.2	51.1	5.9	6.3	8.0	10.3	15.8	24.4	40.9	59.3
50	11.9	11.2	11.9	17.8	63.8	60.2	62.7	64.9	6.4	6.6	5.6	6.1	9.5	9.6	10.2	17.7
100	13.6	12.6	17.4	26.3	74.7	75.8	79.7	86.7	6.1	5.8	5.8	6.2	10.2	11.0	15.2	22.8
200	12.3	13.9	20.9	32.1	81.1	84.6	89.0	96.1	5.1	5.4	5.0	5.4	12.9	14.9	26.0	39.1
	POWER (size-adjusted) : $H_0: \rho = 0.6$ against $H_1: \rho \neq 0.6$, at the 5% level															
25	75.9	93.6	99.5	100.0	66.9	85.7	96.2	98.4	64.5	86.5	97.7	100.0	42.5	55.2	65.2	76.9
50	88.8	99.4	100.0	100.0	76.3	97.4	100.0	100.0	82.6	97.5	100.0	100.0	89.6	99.6	100.0	100.0
100	95.4	100.0	100.0	100.0	85.2	98.7	100.0	100.0	92.2	99.8	100.0	100.0	97.7	100.0	100.0	100.0
200	98.4	100.0	100.0	100.0	86.5	99.5	100.0	100.0	98.0	100.0	100.0	100.0	99.4	100.0	100.0	100.0
PANEL B: Results for β_1, heterogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	IV2^b				QMLE				IVMG^b				CCEMG			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	-1.1	-0.9	-1.2	-1.2	-2.0	-2.0	-1.8	-1.9	1.6	1.6	1.5	1.3	2.2	2.3	2.3	2.5
50	-1.0	-1.1	-0.9	-1.1	-3.2	-2.5	-2.2	-1.9	0.8	0.8	0.8	0.7	0.7	0.7	0.6	0.7
100	-1.0	-1.0	-1.0	-1.0	-3.5	-2.7	-2.1	-2.1	0.4	0.4	0.5	0.4	-1.4	-1.3	-1.3	-1.2
200	-1.0	-1.1	-1.1	-1.1	-3.8	-2.8	-2.4	-2.3	0.1	0.1	0.1	0.2	-2.6	-2.3	-2.4	-2.3
	RMSE ($\times$100)															
25	13.4	9.3	6.9	4.9	15.8	12.5	9.8	7.6	17.3	11.6	8.4	6.3	16.3	12.0	9.2	7.5
50	9.3	6.7	4.6	3.4	14.4	10.5	7.9	5.5	10.2	7.2	4.9	3.6	10.9	8.0	6.2	5.1
100	6.8	4.8	3.4	2.5	13.0	9.3	6.0	4.3	6.8	4.5	3.3	2.4	8.9	6.3	4.8	4.1
200	5.0	3.6	2.6	2.1	12.4	7.8	4.9	3.6	4.9	3.3	2.4	1.7	7.8	5.8	4.8	4.0
	SIZE: $H_0: \beta_1 = 3$ against $H_1: \beta_1 \neq 3$, at the 5% level															
25	9.7	7.5	8.0	6.2	38.2	38.3	36.4	37.4	5.4	5.2	5.8	7.1	7.6	8.0	10.2	16.1
50	9.4	7.4	5.7	6.2	45.1	42.9	36.8	33.8	5.9	5.9	4.8	5.5	8.3	9.2	11.0	16.3
100	9.5	8.1	7.0	7.2	52.5	46.4	37.1	37.6	5.8	3.7	4.4	4.5	8.5	8.8	10.4	17.1
200	9.4	7.8	8.7	12.1	61.3	50.8	44.4	50.1	6.0	4.9	4.8	4.8	8.7	9.1	15.1	23.6
	POWER (size-adjusted) : $H_0: \beta_1 = 3.1$ against $H_1: \beta_1 \neq 3.1$, at the 5% level															
25	12.5	19.1	24.8	47.0	7.7	9.5	10.9	14.3	13.4	21.0	29.3	47.1	12.8	19.9	27.3	38.4
50	21.5	30.4	53.1	78.8	7.9	11.3	10.1	21.4	19.2	34.3	61.5	85.0	16.6	27.1	40.4	56.2
100	31.9	49.8	79.7	97.5	5.7	8.5	20.3	49.9	32.5	64.8	89.7	99.3	16.9	26.6	44.2	60.7
200	46.4	71.4	94.2	99.7	6.1	9.5	34.2	65.0	52.5	84.5	98.8	100.0	15.8	24.2	34.8	50.5

Notes: The DGP is the same as the one for Table 4, except all factgor loadings have mean zero hence the rank condition for CCEMG is not satisfied

Table C9: Bias, root mean squared error (RMSE) of IV2^a, IV2^b IVMG^a, IVMG^b estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$, correlated factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$

	IV2^a				IV2^b				IVMG^a				IVMG^b			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	-0.2	0.0	0.0	-0.7	-0.7	-0.6	-0.7
50	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	-0.5	-0.4	-0.4	-0.3
100	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.3	-0.2	-0.2	-0.2
200	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1
	RMSE (×100)															
25	3.4	2.4	1.8	1.2	3.1	2.2	1.6	1.1	6.4	5.4	3.4	3.6	3.4	2.6	1.9	1.5
50	2.3	1.6	1.1	0.8	2.1	1.4	1.0	0.7	3.0	2.1	1.5	1.1	2.3	1.5	1.1	0.9
100	1.5	1.0	0.7	0.5	1.4	1.0	0.6	0.4	1.7	1.2	0.8	0.6	1.5	1.0	0.7	0.6
200	1.0	0.7	0.5	0.3	1.0	0.6	0.4	0.3	1.1	0.7	0.5	0.4	1.0	0.7	0.5	0.4
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	8.0	6.6	6.1	5.5	7.5	6.4	6.1	5.4	4.7	4.4	4.8	4.7	5.8	7.0	7.9	11.5
50	8.4	6.8	5.4	5.2	8.7	7.0	6.7	4.9	5.5	4.7	5.1	4.2	5.9	5.4	6.5	8.7
100	9.3	6.9	5.4	5.1	8.9	6.5	5.2	4.8	5.5	5.2	3.8	5.4	5.9	6.4	5.2	6.6
200	8.8	6.0	5.5	5.3	8.6	5.4	6.1	5.3	5.1	4.7	5.1	5.3	6.2	5.4	5.5	6.6
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	85.0	97.1	100.0	100.0	89.0	98.0	100.0	100.0	60.6	79.2	90.3	96.2	78.9	93.7	99.0	99.9
50	98.3	100.0	100.0	100.0	99.1	100.0	100.0	100.0	91.9	98.3	99.2	99.7	96.5	99.6	100.0	100.0
100	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.5	100.0	100.0	100.0	99.8	100.0	100.0	100.0
200	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

PANEL B: Results for β_1, homogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$

	IV2^a				IV2^b				IVMG^a				IVMG^b			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	-0.2	0.2	0.0	-0.1	-0.2	0.2	0.0	-0.1	1.4	1.3	1.0	1.4	1.3	1.7	1.4	1.5
50	0.0	0.0	0.0	-0.1	0.0	0.0	0.0	-0.1	0.2	0.1	0.3	0.0	0.9	0.7	0.8	0.5
100	0.2	-0.1	0.0	0.0	0.2	-0.1	0.0	0.0	0.2	-0.3	0.0	0.0	0.6	0.2	0.3	0.4
200	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.2	0.2	0.2
	RMSE (×100)															
25	11.9	8.7	6.2	4.4	11.8	8.7	6.1	4.3	23.7	20.9	13.2	13.0	16.9	11.9	8.5	6.2
50	8.2	5.6	4.1	2.8	8.1	5.6	4.0	2.7	11.5	8.0	5.5	4.5	10.0	6.8	5.0	3.5
100	5.9	4.0	2.8	1.9	5.8	3.9	2.8	1.9	6.7	4.6	3.3	2.2	6.4	4.4	3.2	2.2
200	4.0	2.8	1.9	1.4	3.9	2.8	1.9	1.4	4.4	3.1	2.2	1.6	4.3	3.1	2.1	1.5
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	8.6	7.6	6.2	6.2	8.5	7.5	6.2	6.1	5.2	4.8	5.2	5.0	5.8	5.9	5.9	6.3
50	8.2	5.6	6.1	4.5	8.8	5.4	6.3	4.4	5.0	4.3	5.1	4.5	5.1	4.9	6.0	5.0
100	8.7	6.5	5.9	5.1	8.0	6.3	6.4	5.0	5.3	4.5	5.3	4.0	5.1	4.4	5.4	4.3
200	8.2	6.1	4.7	6.9	8.2	5.7	5.0	7.0	5.4	4.6	4.5	5.3	5.7	5.0	4.6	5.5
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	19.4	25.6	41.2	66.1	19.6	26.2	41.4	66.3	10.7	17.6	23.1	38.5	13.0	20.3	29.9	50.2
50	31.4	45.4	72.0	94.6	31.8	47.7	73.1	95.4	19.3	34.1	53.5	77.0	22.3	37.6	60.8	86.3
100	49.4	74.3	94.0	99.9	50.4	75.8	94.4	99.9	38.3	60.5	84.5	99.2	43.2	66.0	89.7	99.6
200	73.8	94.4	99.6	100.0	73.5	94.8	99.7	100.0	61.3	89.8	99.0	100.0	65.9	91.2	99.2	100.0

Notes: The DGP is identical to that of Table 3. The IV2 and IVMG estimators are defined by (23) and (31), respectively. IV2^a and IVMG^a use the instruments set $(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_{i,-1})$ and IV2^b and IVMG^b $(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_{i,-1}, \hat{\mathbf{X}}_{i,-2})$, where $\hat{\mathbf{X}}_i = \mathbf{M}_{\hat{F}_x}\mathbf{X}_i$ and $\hat{\mathbf{X}}_{i,-j} = \mathbf{M}_{\hat{F}_{x,-j}}\mathbf{X}_{i,-j}$ for $j = 1, 2$.

Table C10: Bias, root mean squared error (RMSE) of IV2^a, IV2^b IVMG^a, IVMG^b estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$, correlated factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$

	IV2^a				IV2^b				IVMG^a				IVMG^b			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	0.3	0.4	0.4	0.4	0.8	0.9	0.9	0.9	0.1	0.0	-0.1	-0.1	-0.7	-0.7	-0.7	-0.7
50	0.4	0.5	0.6	0.6	0.9	1.1	1.2	1.2	-0.1	0.0	0.0	0.0	-0.4	-0.3	-0.3	-0.3
100	0.5	0.7	0.7	0.7	1.1	1.2	1.2	1.2	0.0	0.0	0.0	0.0	-0.2	-0.3	-0.2	-0.2
200	0.6	0.7	0.8	0.8	1.1	1.3	1.3	1.3	0.0	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.1
	RMSE ($\times$100)															
25	4.5	3.2	2.3	1.6	4.4	3.1	2.3	1.7	7.4	5.4	6.0	3.5	4.1	3.0	2.2	1.7
50	3.5	2.5	1.8	1.3	3.5	2.6	2.0	1.6	4.0	2.7	2.0	1.4	3.2	2.2	1.6	1.2
100	2.9	2.1	1.6	1.2	3.0	2.3	1.8	1.6	2.8	2.1	1.4	1.0	2.7	2.0	1.4	1.0
200	2.6	1.9	1.5	1.2	2.8	2.2	1.8	1.6	2.5	1.8	1.3	0.9	2.5	1.8	1.3	0.9
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	10.6	7.3	6.6	6.9	11.7	9.2	9.3	10.5	5.3	4.2	4.3	4.8	6.3	6.6	7.4	8.9
50	11.2	8.9	7.9	9.1	12.7	11.5	12.9	19.8	5.8	5.2	5.2	6.0	6.2	5.1	6.8	6.8
100	9.2	9.3	8.9	12.5	11.6	13.1	15.9	26.3	5.8	5.5	5.5	5.4	5.9	5.6	5.5	6.2
200	9.9	8.7	12.1	15.7	12.3	14.1	21.9	33.3	5.9	5.3	4.9	5.5	5.5	5.3	4.9	5.4
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	67.2	89.8	99.1	100.0	71.9	92.8	99.6	100.0	52.8	74.2	89.0	95.5	61.2	87.9	98.3	100.0
50	85.1	98.6	100.0	100.0	87.2	98.9	100.0	100.0	77.6	95.3	98.7	99.8	80.7	98.2	100.0	100.0
100	95.4	99.9	100.0	100.0	96.3	100.0	100.0	100.0	92.8	99.6	99.9	100.0	93.6	99.7	100.0	100.0
200	98.4	100.0	100.0	100.0	98.6	100.0	100.0	100.0	97.0	99.9	100.0	100.0	97.3	100.0	100.0	100.0

PANEL B: Results for β_1, heterogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$

	IV2^a				IV2^b				IVMG^a				IVMG^b			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS ($\times$100)															
25	-0.8	-0.8	-0.8	-1.0	-1.2	-1.3	-1.4	-1.5	1.8	0.9	1.4	1.3	1.7	1.4	1.5	1.4
50	-0.2	-0.5	-0.4	-0.4	-1.0	-1.3	-1.2	-1.2	0.5	-0.1	0.2	0.3	0.9	0.4	0.7	0.7
100	-0.3	-0.1	-0.2	-0.2	-1.2	-1.0	-1.1	-1.2	-0.2	0.1	0.0	-0.1	0.2	0.5	0.4	0.3
200	-0.3	0.0	-0.1	-0.2	-1.3	-1.0	-1.1	-1.2	-0.1	0.1	0.0	-0.1	0.1	0.3	0.2	0.1
	RMSE ($\times$100)															
25	13.4	9.9	6.8	5.0	13.4	9.8	6.8	5.1	23.9	19.2	16.6	12.9	16.9	11.9	8.5	6.1
50	9.7	6.9	4.7	3.4	9.6	6.9	4.8	3.5	12.2	8.0	6.2	4.3	10.2	7.1	5.2	3.7
100	6.9	5.1	3.4	2.4	6.9	5.1	3.5	2.6	7.0	5.1	3.5	2.4	6.7	4.8	3.4	2.3
200	5.2	3.6	2.5	1.8	5.3	3.8	2.7	2.1	5.1	3.6	2.5	1.7	5.0	3.6	2.5	1.7
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	9.2	7.4	6.3	5.9	9.6	7.5	6.5	6.6	5.5	5.3	5.7	5.4	6.3	5.8	6.2	6.7
50	8.6	6.9	5.7	6.9	8.8	7.5	6.3	7.7	5.9	4.7	5.6	6.0	6.2	4.8	5.7	5.7
100	8.8	8.0	5.8	6.0	9.0	9.0	7.3	8.2	5.1	5.6	4.9	4.8	5.5	5.5	5.3	4.8
200	9.6	7.4	6.4	5.6	10.8	8.6	8.6	12.0	6.1	6.8	5.8	4.2	6.2	6.5	6.0	4.6
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	14.0	17.7	28.4	45.3	14.2	16.6	25.0	39.6	10.6	15.2	21.8	34.4	11.6	16.5	29.8	47.4
50	23.2	32.2	53.3	80.8	20.5	28.3	46.9	72.3	18.0	30.0	47.4	70.7	19.2	35.0	55.0	82.9
100	32.8	50.6	82.5	98.0	28.5	45.7	74.3	95.4	31.1	53.1	82.4	98.1	33.0	60.2	86.1	99.1
200	47.5	79.0	97.1	100.0	40.8	66.5	92.6	99.8	47.9	78.9	97.2	100.0	50.3	80.5	98.0	100.0

Notes: The DGP is identical to that of Table 4. The IV2 and IVMG estimators are defined by (23) and (31), respectively. IV2^a and IVMG^a use the instruments set $(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_{i,-1})$ and IV2^b and IVMG^b $(\hat{\mathbf{X}}_i, \hat{\mathbf{X}}_{i,-1}, \hat{\mathbf{X}}_{i,-2})$, where $\hat{\mathbf{X}}_i = \mathbf{M}_{\hat{F}_x} \mathbf{X}_i$ and $\hat{\mathbf{X}}_{i,-j} = \mathbf{M}_{\hat{F}_{x,-j}} \mathbf{X}_{i,-j}$ for $j = 1, 2$.

Table C11: Bias and root mean squared error (RMSE) of CCEMG, BC-CCEMG, CCEMG*, BC-CCEMG* estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$, independent factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, homogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	MGCCE				BC-MGCCE				MGCCE*				BC-MGCCE*			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	-3.2	-3.4	-3.7	-3.9	3.4	2.9	2.8	2.5	-5.9	-6.5	-6.9	-7.4	0.5	-0.2	-0.3	-0.8
50	-0.8	-1.0	-1.2	-1.5	1.9	1.8	1.5	1.3	-1.9	-2.2	-2.7	-3.1	4.8	4.6	4.2	3.7
100	0.4	0.1	-0.1	-0.4	1.5	1.3	1.0	0.7	0.1	-0.3	-0.7	-1.1	2.4	2.0	1.6	1.1
200	0.9	0.7	0.4	0.1	1.5	1.2	0.9	0.6	1.1	0.6	0.2	-0.3	2.0	1.5	1.1	0.6
	RMSE (×100)															
25	4.1	4.0	4.0	4.1	16.8	8.7	7.0	5.1	7.2	7.2	7.3	7.6	13.3	10.3	8.1	7.1
50	1.8	1.5	1.5	1.6	3.0	2.5	2.0	1.6	2.9	2.8	3.0	3.3	6.4	5.6	4.9	4.2
100	1.1	0.9	0.6	0.6	2.0	1.7	1.3	0.9	1.4	1.2	1.1	1.3	3.1	2.6	1.9	1.4
200	1.2	0.9	0.7	0.4	1.7	1.4	1.1	0.8	1.6	1.1	0.7	0.5	2.4	1.9	1.3	0.8
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	28.2	51.1	81.3	97.7	56.7	60.8	64.3	69.2	41.8	71.5	92.7	99.8	52.5	56.6	63.6	65.8
50	13.8	22.7	49.8	81.8	35.1	43.9	51.9	61.1	22.9	46.2	80.0	97.2	54.1	67.8	80.1	87.9
100	12.4	14.4	18.4	39.0	42.2	52.1	56.2	58.1	12.5	19.4	40.9	80.4	49.6	58.1	62.3	65.9
200	32.8	36.1	34.5	24.7	58.7	68.1	73.0	73.1	31.5	29.4	25.5	37.5	62.4	65.1	62.4	54.0
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	48.2	54.0	57.4	63.2	54.2	61.8	78.5	89.5	5.1	2.8	0.8	0.0	16.5	17.9	26.7	31.5
50	99.7	100.0	100.0	100.0	99.4	100.0	100.0	100.0	86.2	93.0	96.5	98.6	93.0	98.4	99.9	100.0
100	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
200	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

PANEL B: Results for β_1, homogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	MGCCE				BC-MGCCE				MGCCE*				BC-MGCCE*			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	2.9	2.9	3.3	3.2	7.8	6.9	7.8	7.8	3.9	4.4	4.4	4.3	22.3	19.3	17.5	15.8
50	1.1	1.4	1.9	2.0	-0.4	-0.2	0.8	1.3	2.8	2.9	3.7	3.8	2.0	2.0	4.0	4.9
100	-0.6	-0.2	0.3	0.7	-2.2	-1.9	-1.2	-0.7	-0.1	0.5	1.3	1.9	-3.1	-2.2	-1.1	-0.3
200	-1.8	-1.2	-0.8	-0.2	-2.9	-2.2	-1.8	-1.1	-2.1	-1.1	-0.3	0.5	-3.8	-2.7	-1.8	-0.9
	RMSE (×100)															
25	17.1	11.9	9.0	6.7	82.1	46.1	35.5	24.3	20.6	14.8	11.3	8.3	70.6	55.5	44.4	35.4
50	9.7	6.7	5.1	3.9	12.9	9.1	6.7	5.0	11.1	7.8	6.3	5.3	19.8	14.1	10.8	8.8
100	6.4	4.2	3.2	2.2	7.8	5.3	3.9	2.6	6.8	4.6	3.7	3.0	9.1	6.4	4.5	3.0
200	4.7	3.3	2.3	1.5	5.6	4.0	2.9	2.0	5.1	3.5	2.3	1.7	6.5	4.7	3.2	2.0
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	7.2	6.4	8.1	8.8	51.9	54.4	56.0	61.8	7.3	6.7	8.5	8.6	54.4	58.2	61.5	65.6
50	6.8	6.0	6.8	10.1	16.6	17.1	17.5	19.3	7.8	6.9	11.4	19.4	31.1	32.4	35.4	42.6
100	7.0	6.5	7.2	7.3	13.4	12.0	12.9	12.5	6.8	6.5	9.3	16.0	17.4	17.7	16.3	14.6
200	7.8	9.2	9.0	8.7	14.6	15.8	18.5	18.6	9.1	10.3	8.8	9.6	19.7	21.2	20.8	16.9
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	12.5	21.9	32.4	54.0	7.6	7.8	8.7	10.5	11.9	17.7	24.6	42.1	8.6	8.9	9.6	11.1
50	21.9	41.4	70.1	89.8	13.2	19.1	40.5	61.2	20.3	40.1	63.9	84.4	11.9	15.7	24.3	39.2
100	34.7	62.9	89.8	99.7	17.0	33.8	64.5	93.1	33.8	61.8	87.9	99.2	12.7	25.9	52.4	90.0
200	45.3	76.4	97.5	100.0	24.0	49.8	85.4	99.9	35.3	70.9	97.3	100.0	13.7	35.9	77.1	99.4

Notes: The DGP is identical to that of Table 1. The CCEMG and BC-CCEMG estimators are defined by (38) and (40), respectively. CCEMG uses cross-section average of $(y_{it}, ..., y_{it-p}; \mathbf{x}'_{it}; 1)$, whilst CCEMG* uses $(y_{it}, ..., y_{it-p}; \mathbf{x}'_{it}, ..., \mathbf{x}'_{it-p}; 1)$ with p being the interger part of $T^{1/3}$. BC-CCEMG* is the bias-corrected version of CCEMG*.

Table C12: Bias, root mean squared error (RMSE) of CCEMG, BC-CCEMG, CCEMG*, BC-CCEMG* estimates and size and power of the associated t-tests, for the panel ARDL(1,0) model with heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$, independent factor loadings in x_{1it} & u_{it}

PANEL A: Results for ρ, heterogeneous slopes with $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	MGCCE				BC-MGCCE				MGCCE*				BC-MGCCE*			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	-3.1	-3.3	-3.6	-3.8	3.2	2.9	3.1	2.6	-5.9	-6.2	-6.7	-7.1	0.2	-0.1	-0.3	-0.8
50	-0.6	-0.9	-1.2	-1.5	2.2	2.0	1.7	1.3	-1.5	-2.0	-2.5	-3.0	5.2	4.9	4.2	3.8
100	0.4	0.3	-0.1	-0.3	1.6	1.4	1.1	0.8	0.3	0.0	-0.6	-1.0	2.6	2.3	1.7	1.3
200	1.0	0.8	0.5	0.2	1.5	1.3	1.0	0.7	1.1	0.8	0.3	-0.1	2.1	1.7	1.2	0.7
	RMSE (×100)															
25	4.8	4.3	4.1	4.1	13.2	13.7	7.7	5.6	7.6	7.2	7.2	7.4	13.4	10.2	8.2	6.6
50	2.8	2.2	1.9	1.8	4.0	3.1	2.4	1.8	3.4	3.1	3.0	3.2	7.0	6.0	5.1	4.4
100	2.5	1.9	1.3	1.0	3.0	2.4	1.8	1.3	2.7	2.0	1.6	1.4	3.9	3.2	2.3	1.7
200	2.6	1.8	1.3	0.9	2.9	2.1	1.6	1.2	2.8	1.9	1.4	1.0	3.3	2.5	1.8	1.2
	SIZE: $H_0 : \rho = 0.5$ against $H_1 : \rho \neq 0.5$, at the 5% level															
25	18.6	30.7	56.7	84.2	43.6	50.3	55.8	59.9	32.5	54.2	83.9	97.6	48.4	52.1	56.5	62.9
50	8.3	10.2	16.2	36.4	19.4	22.5	28.2	32.0	10.7	19.7	42.8	79.1	40.6	53.8	64.3	77.0
100	7.3	8.0	8.0	9.6	14.9	18.4	18.0	19.7	8.6	8.3	11.8	23.2	22.5	29.5	30.7	31.4
200	9.3	8.0	9.0	7.1	12.3	14.2	16.2	16.5	11.8	9.5	9.0	8.1	17.8	20.0	20.1	18.0
	POWER (size-adjusted) : $H_0 : \rho = 0.6$ against $H_1 : \rho \neq 0.6$, at the 5% level															
25	31.0	40.6	46.9	51.8	37.8	50.5	69.9	85.5	5.5	3.0	0.9	0.1	14.1	17.1	24.2	30.5
50	89.2	98.6	100.0	100.0	91.3	99.3	100.0	100.0	70.6	80.4	90.2	94.0	76.5	93.2	99.3	100.0
100	98.1	100.0	100.0	100.0	98.0	100.0	100.0	100.0	95.7	99.9	100.0	100.0	94.1	99.9	100.0	100.0
200	99.2	100.0	100.0	100.0	99.1	100.0	100.0	100.0	98.4	100.0	100.0	100.0	97.9	100.0	100.0	100.0
PANEL B: Results for β_1, heterogeneous slopes $\{\rho, \beta_1, \beta_2\} = \{0.5, 3, 1\}$ and $\pi_u = 3/4$																
	MGCCE				BC-MGCCE				MGCCE*				BC-MGCCE*			
T,N	25	50	100	200	25	50	100	200	25	50	100	200	25	50	100	200
	BIAS (×100)															
25	2.2	3.3	2.9	3.2	7.3	9.1	6.8	9.2	3.3	4.4	3.9	4.3	22.9	21.4	17.8	17.3
50	0.8	1.3	1.8	1.9	-0.9	0.0	0.6	1.1	2.2	2.6	3.4	3.6	1.4	2.3	3.4	4.7
100	-0.4	-0.4	0.1	0.6	-2.1	-2.1	-1.4	-0.8	-0.1	0.0	1.0	1.6	-2.9	-2.7	-1.5	-0.5
200	-1.6	-1.4	-0.9	-0.3	-2.7	-2.5	-1.8	-1.2	-1.9	-1.4	-0.6	0.3	-3.7	-3.1	-2.1	-1.1
	RMSE (×100)															
25	16.8	12.2	9.0	6.7	63.7	46.2	34.4	28.7	20.5	15.2	11.0	8.5	70.8	58.3	43.9	34.4
50	9.8	7.2	5.3	3.8	13.5	9.7	6.7	4.8	10.9	8.2	6.4	5.2	20.1	14.6	10.8	8.7
100	6.7	4.7	3.2	2.3	7.9	5.8	4.0	2.7	7.1	5.0	3.5	2.9	9.2	6.8	4.5	3.0
200	5.3	3.8	2.6	1.7	6.1	4.5	3.2	2.2	5.6	4.0	2.7	1.8	6.9	5.1	3.6	2.3
	SIZE: $H_0 : \beta_1 = 3$ against $H_1 : \beta_1 \neq 3$, at the 5% level															
25	6.7	6.5	7.6	7.8	50.0	55.1	57.3	61.0	6.7	6.7	7.0	9.7	56.5	59.4	62.8	67.3
50	6.1	6.3	7.5	9.1	16.6	16.1	16.6	18.1	6.5	7.6	11.0	17.8	29.2	31.0	34.9	42.7
100	5.7	6.4	5.3	6.3	11.8	12.6	12.4	11.1	6.4	6.7	6.9	12.2	15.3	16.6	14.6	12.8
200	8.5	9.4	8.6	6.8	12.8	15.9	16.8	14.7	9.4	9.3	8.8	7.1	16.7	20.2	19.7	15.0
	POWER (size-adjusted) : $H_0 : \beta_1 = 3.1$ against $H_1 : \beta_1 \neq 3.1$, at the 5% level															
25	12.8	19.7	30.9	53.4	8.4	8.8	8.9	11.2	10.9	16.7	25.5	41.4	9.7	9.0	9.2	11.9
50	19.8	38.1	62.6	89.9	12.6	18.0	37.5	63.8	19.5	35.7	59.2	83.6	11.2	15.0	25.0	42.0
100	33.4	51.3	87.7	99.5	17.0	29.4	56.5	92.7	30.7	52.2	89.0	99.2	10.2	19.1	46.9	86.5
200	34.6	64.0	94.4	99.9	19.8	38.4	73.2	99.1	29.2	60.4	93.3	100.0	10.9	24.4	63.5	97.7

Notes: The DGP is identical to that of Table 2. The CCEMG and BC-CCEMG estimators are defined by (38) and (40), respectively. CCEMG uses cross-section average of $(y_{it}, ..., y_{it-p}; \mathbf{x}'_{it}; 1)$, whilst CCEMG* uses $(y_{it}, ..., y_{it-p}; \mathbf{x}'_{it}, ..., \mathbf{x}'_{it-p}; 1)$ with p being the interger part of $T^{1/3}$. BC-CCEMG* is the bias-corrected version of CCEMG*.