Two-Stage Instrumental Variable Estimation of Linear Panel Data Models with Interactive Effects

Working Paper Series

No 90 / 2021
Two-Stage Instrumental Variable Estimation of Linear Panel Data Models with Interactive Effects

Guowei Cui

Milda Norkutė

Vasilis Sarafidis

Takashi Yamagata

---

1 School of Economics, Huazhong University of Science and Technology (HUST), Wuhan 430074, China
2 Bank of Lithuania, Totoriu g. 4, LT-01121, Vilnius, Lithuania; E-mail address: MNorkute@lb.lt. This author notes that the views expressed in the present paper do not necessarily represent those of the Bank of Lithuania and the Eurosystem.
3 Department of Econometrics and Business Statistics, Monash University, VIC 3145, Australia. E-mail address: vasilis.sarafidis@monash.edu. This author gratefully acknowledges financial support from the Australian Research Council, under research Grant Number DP-170103135.
4 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 JP18K01545.
ABSTRACT

This paper analyses the instrumental variables (IV) approach put forward by Norkutė et al. (2021), in the context of static linear panel data models with interactive effects present in the error term and the regressors. Instruments are obtained from transformed regressors, thereby it is not necessary to search for external instruments. We consider a two-stage IV (2SIV) and a mean-group IV (MGIV) estimator for homogeneous and heterogeneous slope models, respectively. The asymptotic analysis reveals that: (i) the $\sqrt{NT}$-consistent 2SIV estimator is free from asymptotic bias that may arise due to the estimation error of the interactive effects, whilst (ii) existing estimators can suffer from asymptotic bias; (iii) the proposed 2SIV estimator is asymptotically as efficient as existing estimators that eliminate interactive effects jointly in the regressors and the error, whilst; (iv) the relative efficiency of the estimators that eliminate interactive effects only in the error term is indeterminate. A Monte Carlo study confirms good approximation quality of our asymptotic results.

Keywords: Large panel data, interactive effects, common factors, principal components analysis, instrumental variables.

JEL Classification: C13, C15, C23, C26.
1 Introduction

Panel data sets with large cross-section and time-series dimensions ($N$ and $T$, respectively) have become increasingly available in the social sciences. As a result, regression analysis of large panels has gained an ever-growing popularity. A central issue in these models is how to properly control for rich sources of unobserved heterogeneity, including common shocks and interactive effects (see e.g. Sarafidis and Wansbeek (2020) for a recent overview).

Broadly speaking, there are two popular estimation approaches currently advanced in the field. The first one involves eliminating the interactive effects from the error term and the regressors jointly, in a single stage. Representative methods include the Common Correlated Effects approach of Pesaran (2006), which involves least-squares on a regression model augmented by cross-sectional averages (CA) of observables; and the Principal Components (PC) estimator considered first by Kapetanios and Pesaran (2005) and analysed subsequently by Westerlund and Urbain (2015). The second approach asymptotically eliminates the interactive effects from the error term only. The representative method is the Iterative Principal Components (IPC) estimator of Bai (2009), further developed by Moon and Weidner (2015, 2017), among many others. An attractive feature of CA (as well as PC) is that it permits estimation of models with heterogeneous slopes. On the other hand, an advantage of IPC is that it does not assume that regressors are subject to a factor structure.

In models with homogeneous slopes, Westerlund and Urbain (2015) showed that both CA and PC estimators suffer from asymptotic bias due to the incidental parameter problem (see Juodis et al. (2021) for additional results on the asymptotic properties of CA). A similar outcome was shown by Bai (2009) for the IPC estimator. Thus in all three cases, bias correction is necessary for asymptotically valid inferences. In addition, the CA estimator requires the so-called rank condition, which assumes that the number of factors does not exceed the rank of the (unknown) matrix of cross-sectional averages of the factor loadings. On the other hand, IPC involves non-linear optimisation, and so convergence to the global optimum might not be guaranteed (see e.g. Jiang et al. (2017)).

This paper analyses the instrumental variables (IV) approach put forward by Norkutė et al. (2021) in the context of a static linear panel data model. Their approach differs from CA, PC and IPC because it asymptotically eliminates the interactive effects in the error term and in the regressors separately, in two stages. In particular, for models with homogeneous slopes, in the first stage the interactive effects are projected out from the regressors. Subsequently, the transformed regressors are used as instruments to obtain consistent estimates of the model parameters. This way, it is not necessary to search for external instruments. In the second stage, the interactive effects in the error term are eliminated using the first-stage residuals, and a second IV regression is run. That is, IV regression is performed in both of two stages. The resulting two-stage IV (2SIV) estimator is shown to be $\sqrt{NT}$-consistent and asymptotically normal. For models with heterogeneous slopes, we analyse a mean-group IV (MGIV) estimator and establish $\sqrt{N}$-consistency and asymptotic normality. The asymptotic results established in this paper are completely new, as we permit weak cross-section and time-series dependence in the idiosyncratic errors. The weak dependence assumption is typically employed by the static panel data literature, such as Bai (2009). In contrast, Norkutė et al. (2021) focus on dynamic panels with interactive effects, assuming cross-sectional and serial independence of the idiosyncratic disturbances.

In addition, the present paper offers new insights into the literature by comparing and contrasting the asymptotic properties of 2SIV, IPC, PC and CA. Such a task was not considered by Norkutė
et al. (2021). To be more specific, we analytically show why the proposed two-stage approach makes
the 2SIV estimator free from asymptotic bias, whilst under the same conditions IPC, PC and CA
are subject to biases. In brief, the reason for the lack of asymptotic bias of 2SIV is that the factors
in the regressors and the errors are estimated separately in two stages. This makes the endogeneity
caused by the estimation errors of the interactive effects asymptotically negligible. Moreover, our
analysis reveals that 2SIV is asymptotically as efficient as the bias-corrected versions of PC and
CA, whereas the relative efficiency of the bias-corrected IPC estimator is indeterminate, in general.
This is because the IPC estimator (i) does not necessarily eliminate the factors contained in the
regressors; (ii) requires a transformation of the regressors, which is due to the estimation error of
the interactive effects.

A Monte Carlo study confirms good approximation quality of our asymptotic results and com-
petent performance of 2SIV and MGIV relative to existing estimators. Furthermore, the results
demonstrate that the bias-corrections of IPC and PC can noticeably inflate the dispersion of the
estimators in finite samples. We apply our methodology to study the effect of climate shocks
on economic growth using an unbalanced panel of 125 countries over the period 1961-2003. The
implications of our results are different from those obtained in existing literature.

A Stata algorithm that implements our approach, has been recently developed by Kripfganz
and Sarafidis (2021) and is available to all Stata users.1

The remainder of this paper is organised as follows. Section 2 introduces a panel data model with
homogeneous slopes and interactive effects, and describes the set of assumptions employed. Section
3 studies the asymptotic properties of the proposed 2SIV estimator. Section 4 analyses a mean-
group IV estimator for models with heterogeneous slopes and establishes its properties in large
samples. Section 5 provides an asymptotic comparison among 2SIV, IPC, CA and PC. Section
6 studies the finite sample performance of these estimators and Section 7 provides an empirical
illustration. Section 8 concludes. Proofs of main theoretical results with necessary lemmas and
auxiliary lemmas are relegated to Online Supplement.

**Notation:** Throughout, we denote the largest eigenvalues of the $N \times N$ matrix $A = (a_{ij})$ by
$\mu_{\text{max}}(A)$, its trace by $\text{tr}(A) = \sum_{i=1}^{N} a_{ii}$, its Frobenius norm by $\|A\| = \sqrt{\text{tr}(A' A)}$. The projection
matrix on $A'$ is $P_A = A (A' A)^{-1} A'$ and $M_A = I - P_A$. $C$ is a generic positive constant large
enough, $C_{\text{min}}$ is a small positive constant sufficiently away from zero, $\delta_{N,T}^2 = \min\{N, T\}$. We use
$N, T \to \infty$ to denote that $N$ and $T$ pass to infinity jointly.

## 2 Model and assumptions

We consider the following panel data model:

$$
\begin{align*}
    y_{it} &= x_{it}' \beta + u_i; & u_i &= \varphi_i^0 h_i^0 + \varepsilon_{it}, \\
    x_{it} &= \Gamma_i^0 r_i^0 + v_{it}; & i &= 1, \ldots, N; \quad t = 1, \ldots, T,
\end{align*}
$$

(2.1)

where $y_{it}$ denotes the value of the dependent variable for individual $i$ at time $t$, $x_{it}$ is a $k \times 1$ vector
of regressors and $\beta$ is the corresponding vector of slope coefficients. $u_i$ follows a factor structure,
where $h_i^0$ is an $r_2 \times 1$ vector of latent factors, $\varphi_i^0$ is the associated factor loading vector, and $\varepsilon_{it}$
denotes an idiosyncratic error. The regressors are assumed to be strictly exogenous with respect to

1See http://www.kripfganz.de/stata/xtivdfreg.html.
\( \varepsilon_{it} \), however they are subject to a factor model, where \( \mathbf{f}_i^0 \) denotes an \( r_1 \times 1 \) vector of latent factors, \( \mathbf{\Gamma}_i^0 \) is a \( r_1 \times k \) matrix of factor loadings, and \( \varepsilon_{it} \) is an idiosyncratic error of dimension \( k \times 1 \). We treat \( r_1 \) and \( r_2 \) as given.\(^2\)

Estimation of the model above has been studied by Pesaran (2006), Bai and Li (2014), Westerlund and Urbain (2015), Juodis and Sarafidis (2020, 2021), Cui et al. (2019) to mention a few. Such model has been employed in a wide variety of fields, including economics and finance.

**Remark 2.1** Permitting different sets of interactive effects in \( \mathbf{x}_{it} \) and \( u_{it} \) is important not only from the empirical perspective but also from the theoretical perspective. It plays a crucial role when we analytically compare the estimators that eliminate the factors in the error term and the regressors separately (as in our approach), and those estimators that eliminate the factors in the error term only (as in the IPC approach of Bai (2009)). This remark does not apply to estimators that extract factors in \( \mathbf{x}_{it} \) and \( u_{it} \) jointly, as in the approaches considered by Pesaran (2006) and Westerlund and Urbain (2015)); see Section 5 for more details.

Stacking Eq. (2.1) over \( t \), we have

\[
\begin{align*}
\mathbf{y}_i &= \mathbf{X}_i \beta + \mathbf{u}_i; \\
\mathbf{u}_i &= \mathbf{H}^0 \varphi_i^0 + \mathbf{\varepsilon}_i,
\end{align*}
\]

where \( \mathbf{y}_i = (y_{i1}, \ldots, y_{iT})' \), \( \mathbf{X}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iT})' \), \( \mathbf{F}_0 = (\mathbf{f}_1^0, \ldots, \mathbf{f}_T^0)' \), \( \mathbf{H}^0 = (\mathbf{h}_1^0, \ldots, \mathbf{h}_k^0)' \), \( \mathbf{\varepsilon}_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})' \) and \( \mathbf{V}_i = (\mathbf{v}_{i1}, \ldots, \mathbf{v}_{iT})' \).

Following Norkutė et al. (2021), we consider an IV estimation approach that involves two stages. In the first stage, the common factors in \( \mathbf{X}_i \) are asymptotically eliminated using principal components analysis. Next, the transformed regressors are used to construct instruments and estimate the model parameters. To illustrate the first-stage IV estimator, suppose that \( \mathbf{F}_0 \) is observed. Pre-multiplying \( \mathbf{X}_i \) by \( \mathbf{M}_{\mathbf{F}_0} \) yields

\[
\mathbf{M}_{\mathbf{F}_0} \mathbf{X}_i = \mathbf{M}_{\mathbf{F}_0} \mathbf{V}_i.
\]

Assuming \( \mathbf{V}_i \) is independent of \( \varepsilon_i \), \( \mathbf{H}^0 \) and \( \varphi_i^0 \), it is easily seen that \( E[\mathbf{X}'_i \mathbf{M}_{\mathbf{F}_0} \mathbf{u}_i] = E[\mathbf{V}'_i \mathbf{M}_{\mathbf{F}_0} (\mathbf{H}^0 \varphi_i^0 + \varepsilon_i)] = \mathbf{0} \). Together with the fact that \( \mathbf{M}_{\mathbf{F}_0} \mathbf{X}_i \) is correlated with \( \mathbf{X}_i \) through \( \mathbf{V}_i \), \( \mathbf{M}_{\mathbf{F}_0} \mathbf{X}_i \) can be regarded as an instrument for \( \mathbf{X}_i \).

The first-stage (infeasible) estimator is defined as

\[
\hat{\beta}_{1SIV}^{inf} = \left( \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\mathbf{F}_0} \mathbf{X}_i \right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\mathbf{F}_0} \mathbf{y}_i.
\]

In the second stage, the space spanned by \( \mathbf{H}^0 \) is estimated from the residual \( \tilde{\mathbf{u}}_i^{inf} = \mathbf{y}_i - \mathbf{X}_i \hat{\beta}_{1SIV}^{inf} \) and then it is projected out. To illustrate, suppose that \( \mathbf{H}^0 \) is also observed; one can instrument \( \mathbf{X}_i \) using \( \mathbf{M}_{\mathbf{H}^0} \mathbf{M}_{\mathbf{F}_0} \mathbf{X}_i \). Note that \( E[\mathbf{X}'_i \mathbf{M}_{\mathbf{F}_0} \mathbf{M}_{\mathbf{H}^0} \mathbf{u}_i] = E[\mathbf{V}'_i \mathbf{M}_{\mathbf{F}_0} \mathbf{M}_{\mathbf{H}^0} \varepsilon_i] = \mathbf{0} \). The (infeasible) second-stage IV (2SIV) estimator of \( \beta \) is given by

\[
\hat{\beta}_{2SIV}^{inf} = \left( \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\mathbf{F}_0} \mathbf{M}_{\mathbf{H}^0} \mathbf{X}_i \right)^{-1} \sum_{i=1}^{N} \mathbf{X}_i' \mathbf{M}_{\mathbf{F}_0} \mathbf{M}_{\mathbf{H}^0} \mathbf{y}_i.
\]

\(^2\)In practice, \( r_1 \) can be estimated from the raw data \( \{\mathbf{X}_i\}_{i=1}^{N} \) using methods already available in the literature, such as the information criteria of Bai and Ng (2002) or the eigenvalue-based tests of Kapetanios (2010) and Ahn and Horenstein (2013). \( r_2 \) can be estimated in the same way from the residual covariance matrix. An asymptotic justification of such practice is discussed in Bai (2009b, Section C.3). In the Monte Carlo section of the paper we show that these methods provide quite accurate determination of the number of factors.
In practice, \( F^0 \) and \( H^0 \) are typically unobserved. As it will be discussed in detail below, we replace these quantities with estimates obtained using principal components analysis, as advanced in Bai (2003) and Bai (2009).

To obtain our theoretical results it is sufficient to make the following assumptions.

**Assumption A (idiiosyncratic error in y) :** We assume that

1. \( E(\varepsilon_{it}) = 0 \) and \( \mathbb{E} \| \varepsilon_{it} \|^{8+\delta} \leq C \) for some \( \delta > 0 \);

2. Let \( \sigma_{ij,st} \equiv \mathbb{E}(\varepsilon_{is} \varepsilon_{jt}) \). We assume that there exist \( \tilde{\sigma}_{ij} \) and \( \tilde{\sigma}_{st} \), \( |\sigma_{ij,st}| \leq \tilde{\sigma}_{ij} \) for all \((s,t)\), and \( |\sigma_{ij,st}| \leq \tilde{\sigma}_{st} \) for all \((i,j)\), such that \( N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \tilde{\sigma}_{ij} \leq C \); \( T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{\sigma}_{st} \leq C \).

3. For every \((s,t)\), \( \mathbb{E} \| N^{-1/2} \sum_{i=1}^{N} [\varepsilon_{is} \varepsilon_{it} - \sigma_{is,it}] \|^{4} \leq C \).

4. For each \( j \), \( \mathbb{E} \| N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} [\varepsilon_{it} \varepsilon_{jt} - \mathbb{E}(\varepsilon_{it} \varepsilon_{jt})] \varphi_{jt}^{0} \|^{2} \leq C \). Also, for each \( s \), \( \mathbb{E} \| N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} [\varepsilon_{is} \varepsilon_{it} - \mathbb{E}(\varepsilon_{is} \varepsilon_{it})] \vartheta_{i}^{0} \|^{2} \leq C \), where \( \vartheta_{i}^{0} = (\varphi_{it}^{0}, \vartheta_{it}^{0})' \).

5. \( N^{-1} T^{-2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{s=1}^{T} \sum_{t_{1}=1}^{T} \sum_{t_{2}=1}^{T} \sum_{s_{1}=1}^{T} \sum_{t_{3}=1}^{T} \| \text{cov}(\varepsilon_{is}, \varepsilon_{is}, \varepsilon_{jt}, \varepsilon_{jt}) \| \leq C \).

**Assumption B (idiiosyncratic error in x)** Let \( \Sigma_{ij,st} \equiv \mathbb{E}(v_{is} v_{jt}') \). We assume that

1. \( v_{it} \) is group-wise independent from \( \varepsilon_{it} \), \( \mathbb{E}(v_{it}) = 0 \) and \( \mathbb{E} \| v_{it} \|^{8+\delta} \leq C \);

2. There exist \( \tilde{\tau}_{ij} \) and \( \tilde{\tau}_{st} \), \( \| \Sigma_{ij,st} \| \leq \tilde{\tau}_{ij} \) for all \((s,t)\), and \( \| \Sigma_{ij,st} \| \leq \tilde{\tau}_{st} \) for all \((i,j)\), such that \( N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \tilde{\tau}_{ij} \leq C \); \( T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{\tau}_{st} \leq C \);

3. For every \((s,t)\), \( \mathbb{E} \| N^{-1/2} \sum_{i=1}^{N} [v_{is} v_{it}' - \Sigma_{is,it}] \|^{4} \leq C \).

4. For each \( j \), \( \mathbb{E} \| N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \varphi_{jt}^{0} \otimes [v_{it} v_{jt}' - \mathbb{E}(v_{it} v_{jt}')] \varphi_{jt}^{0} \|^{2} \leq C \). Additionally, for each \( s \), \( \mathbb{E} \| N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} [v_{is} v_{it}' - \mathbb{E}(v_{is} v_{it}')] \vartheta_{i}^{0} \|^{2} \leq C \).

5. \( N^{-1} T^{-2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{s=1}^{T} \sum_{t_{1}=1}^{T} \sum_{t_{2}=1}^{T} \sum_{s_{1}=1}^{T} \sum_{t_{3}=1}^{T} \| \text{cov}(v_{is}, v_{is}, v_{jt}, v_{jt}) \| \leq C \).

**Assumption C (factors) \( E(\| F^{0} \|^{4}) \leq C, T^{-1} F^{0} F^{0} \xrightarrow{p} \Sigma_{F} \) as \( T \to \infty \) for some non-random positive definite matrix \( \Sigma_{F} \). \( E(\| h^{0} \|^{4}) \leq C, T^{-1} H^{0} H^{0} \xrightarrow{p} \Sigma_{H} \) as \( T \to \infty \) for some non-random positive definite matrix \( \Sigma_{H} \).**

**Assumption D (loadings) \( E(\| \Gamma^{0} \|^{4}) \leq C, \Gamma^{0} = N^{-1} \sum_{i=1}^{N} \Gamma^{0}_{i} \Gamma^{0}_{i} \xrightarrow{p} \Gamma^{0} \) as \( N \to \infty \), and \( E(\| \varphi^{0} \|^{4}) \leq C, \varphi^{0} = N^{-1} \sum_{i=1}^{N} \varphi^{0}_{i} \varphi^{0}_{i} \xrightarrow{p} \varphi^{0} \) as \( N \to \infty \) for some non-random positive definite matrices \( \Gamma^{0} \) and \( \varphi^{0} \). In addition, \( \Gamma^{0}_{i} \) and \( \varphi^{0}_{i} \) are independent groups from \( \varepsilon_{it}, v_{it}, f_{it}^{0} \) and \( h_{it}^{0} \).**

**Assumption E (identification) The matrix \( T^{-1} X_{i}' M_{\varphi^{0}} X_{i} \) has full column rank and \( \mathbb{E} \| T^{-1} X_{i}' M_{\varphi^{0}} X_{i} \|^{2+2\delta} \leq C \) for all \( i \).**

Unlike Norkuté et al. (2021), Assumptions A and B permit weak cross-sectional and serial dependence in \( \varepsilon_{it} \) and \( v_{it} \), in a similar manner to Bai (2009). Assumptions C and D on the moments and the limit variance of factors and factor loadings are standard and in line with Bai (2009). Note that these assumptions permit that \( T^{-1} G^{0} G^{0} \xrightarrow{p} \Sigma_{G}^{0} \), a positive semi-definite matrix, where \( G^{0} = (F^{0}, H^{0}) \). Assumption E is sufficient for identification of heterogeneous slope coefficients.
3 Estimation of models with homogeneous slopes

We propose the following two-stage IV procedure:

1. Estimate the span of $\mathbf{F}$ by $\hat{\mathbf{F}}$, defined as $\sqrt{T}$ times the eigenvectors corresponding to the $r_1$ largest eigenvalues of the $T \times T$ matrix $N^{-1}T^{-1} \sum_{i=1}^{N} X_i X_i'$. Then estimate $\beta$ as

$$
\hat{\beta}_{1SIV} = \left( \sum_{i=1}^{N} X_i M_{\hat{\mathbf{F}}} X_i \right)^{-1} \sum_{i=1}^{N} X_i M_{\hat{\mathbf{F}}} y_i. \tag{3.1}
$$

2. Let $\hat{u}_i = y_i - X_i \hat{\beta}_{1SIV}$. Define $\hat{\mathbf{H}}$ to be $\sqrt{T}$ times the eigenvectors corresponding to the $r$ largest eigenvalues of the $T \times T$ matrix $(NT)^{-1} \sum_{i=1}^{N} \hat{u}_i \hat{u}_i'$. The second-stage estimator of $\beta$ is defined as follows:

$$
\hat{\beta}_{2SIV} = \left( \sum_{i=1}^{N} X_i M_{\hat{\mathbf{F}}} M_{\hat{\mathbf{H}}} X_i \right)^{-1} \sum_{i=1}^{N} X_i M_{\hat{\mathbf{F}}} M_{\hat{\mathbf{H}}} y_i. \tag{3.2}
$$

In order to establish the asymptotic properties of these estimators, we first expand (3.1) as follows:

$$
\sqrt{NT}(\hat{\beta}_{1SIV} - \beta) = \left( \frac{1}{NT} \sum_{i=1}^{N} X_i M_{\hat{\mathbf{F}}} X_i \right)^{-1} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i M_{\hat{\mathbf{F}}} u_i. \tag{3.3}
$$

The following Proposition demonstrates $\sqrt{NT}$-consistency of the first-stage estimator, $\hat{\beta}_{1SIV}$:

**Proposition 3.1** Under Assumptions A-E, we have

$$
N^{-1/2}T^{-1/2} \sum_{i=1}^{N} X_i' M_{\hat{\mathbf{F}}} u_i = N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i' M_{\hat{\mathbf{F}}} u_i + b_{0F} + b_{1F} + b_{2F} + O_p(\sqrt{NT}\delta^{-3}_{NT})
$$

with

$$
b_{0F} = - N^{-1/2}T^{-1/2} \sum_{i=1}^{N} N^{-1} \sum_{\ell=1}^{N} \Gamma_0^0(\mathbf{Y}^0)^{-1} \Gamma_0^0 V_i' M_{\hat{\mathbf{F}}} u_i;$$

$$
b_{1F} = - \sqrt{\frac{T}{N}} \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{h=1}^{N} \sum_{\ell=1}^{N} \mathbb{E}(V_i V_h) \Gamma_0^0(\mathbf{Y}^0)^{-1} (T^{-1} \mathbf{F}^0 \mathbf{F}^0)^{-1} \mathbf{F}^0 \mathbf{H}^0 \varphi_i^0 + \Gamma_0^0(\mathbf{Y}^0)^{-1} \Gamma_0^0 \mathbb{E}(V_i V_h) \Gamma_0^0(\mathbf{Y}^0)^{-1} (T^{-1} \mathbf{F}^0 \mathbf{F}^0)^{-1} \mathbf{F}^0 \mathbf{H}^0 \varphi_i^0;$$

$$
b_{2F} = - \sqrt{\frac{N}{NT}} \frac{1}{NT} \sum_{i=1}^{N} \Gamma_0^0(\mathbf{Y}^0)^{-1} (T^{-1} \mathbf{F}^0 \mathbf{F}^0)^{-1} \mathbf{F}^0 \Sigma \mathbf{M}_{\hat{\mathbf{F}}} \mathbf{H}^0 \varphi_i^0,
$$

where $\mathbf{Y}^0 = \sum_{i=1}^{N} \mathbf{F}^0 \mathbf{F}^0 / N$, $\Sigma = N^{-1} \sum_{i=1}^{N} \mathbb{E}(V_i V_i)$, and $N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i' M_{\hat{\mathbf{F}}} u_i$, $b_{0F}$, $b_{1F}$ and $b_{2F}$ are $O_p(1)$ when $N/T \rightarrow C$. Consequently,

$$
\sqrt{NT}(\hat{\beta}_{1SIV} - \beta) = O_p(1).
$$

---

3. Letting $\hat{\varphi}_i = (\hat{\mathbf{H}}' \hat{\mathbf{H}})^{-1} \hat{\mathbf{H}}' \hat{\mathbf{u}}_i$, an alternative second-stage estimator can be defined by $(\sum_{i=1}^{N} X_i' M_{\hat{\mathbf{F}}} X_i)^{-1} \sum_{i=1}^{N} X_i' M_{\hat{\mathbf{F}}}(y_i - P_{\hat{\mathbf{H}}} \hat{\mathbf{u}}_i)$. We do not discuss this estimator since the finite sample performance was slightly worse than that of $\hat{\beta}_{2SIV}$.  

Proposition 3.1 implies that \( \hat{\beta}_{1SIV} \) is consistent but asymptotically biased. Rather than bias-correcting this estimator, we show that the second-stage IV estimator is free from asymptotic bias. To begin with, we make use of the following expansion:

\[
\sqrt{NT}(\hat{\beta}_{2SIV} - \beta) = \left( \frac{1}{ NT } \sum_{i=1}^{N} X_i' M_{\hat{F}} M_{\hat{H}} X_i \right)^{-1} \frac{1}{ \sqrt{NT} } \sum_{i=1}^{N} X_i' M_{\hat{F}} M_{\hat{H}} u_i. \tag{3.4}
\]

The next proposition provides an asymptotic representation of \( \hat{\beta}_{2SIV} \).

**Proposition 3.2** Under Assumptions A-E, as \( N, T \to \infty, N/T \to C \), we have

\[
\sqrt{NT}(\hat{\beta}_{2SIV} - \beta) = \left( \frac{1}{ NT } \sum_{i=1}^{N} X_i' M_{F^0} H_{H^0} X_i \right)^{-1} \frac{1}{ \sqrt{NT} } \sum_{i=1}^{N} X_i' M_{F^0} H_{H^0} \varepsilon_i + O_p(\sqrt{NT} \delta_{NT}^{-3})
\]

\[
= \left( \frac{1}{ NT } \sum_{i=1}^{N} V_i' V_i \right)^{-1} \frac{1}{ \sqrt{NT} } \sum_{i=1}^{N} V_i' \varepsilon_i + O_p(\sqrt{NT} \delta_{NT}^{-3}).
\]

Proposition 3.2 shows that the effects of estimating \( F^0 \) from \( X_i \) and \( H^0 \) from \( \hat{u}_i = y_i - X_i \hat{\beta}_{1SIV} \) are asymptotically negligible. Moreover, \( \hat{\beta}_{2SIV} \) is asymptotically equivalent to a least-squares estimator obtained by regressing \( (y_i - H^0 \varphi_0^0) \) on \( (X_i - F^0 \varphi^0) \).

To establish asymptotic normality under weak cross-sectional and serial error dependence, we place the following additional assumption, which is in line with Assumption E in Bai (2009).

**Assumption F** \( \text{plim} N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} V_i' \varepsilon_i \varepsilon_j' V_j / T = B \), and \( \frac{1}{ \sqrt{NT} } \sum_{i=1}^{N} V_i' \varepsilon_i \xrightarrow{d} N(0, B) \), for some non-random positive definite matrix \( B \).

Using Proposition 3.2 and Assumption F, it is straightforward to establish the asymptotic distribution of \( \hat{\beta}_{2SIV} \):

**Theorem 3.1** Under Assumptions A-F, as \( N, T \to \infty, N/T \to C \), we have

\[
\sqrt{NT}(\hat{\beta}_{2SIV} - \beta) \xrightarrow{d} N(0, \Psi)
\]

where \( \Psi = A^{-1} B A^{-1} \).

Note that despite the fact that our assumptions permit serial correlation and heteroskedasticity in \( v_{it} \) and \( \varepsilon_{it} \), \( \hat{\beta}_{2SIV} \) is not subject to any asymptotic bias. We discuss this property in more detail in Section 5.

As discussed in Bai (2009) and Norkutė et al. (2021), in general consistent estimation of \( \Psi \) is not feasible when the idiosyncratic errors are both cross-sectional and time-series dependent. Following Norkutė et al. (2021) and Cui et al. (2019), we propose using the following estimator:

\[
\hat{\Psi} = \hat{A}^{-1} \hat{B} \hat{A}^{-1}
\]

with

\[
\hat{A} = \frac{1}{ NT } \sum_{i=1}^{N} X_i' M_{\hat{F}} M_{\hat{H}} X_i; \quad \hat{B} = \frac{1}{ NT } \sum_{i=1}^{N} X_i' M_{\hat{F}} M_{\hat{H}} \hat{u}_i \hat{u}_i' M_{\hat{H}} M_{\hat{F}} X_i,
\]

where \( \hat{u}_i = y_i - X_i \hat{\beta}_{2SIV} \). In line with the discussion in Hansen (2007), it can be shown that when \( \{v_{it}, \varepsilon_{it}\} \) follows a certain strong mixing process over \( t \) and is independent over \( i \), \( \hat{\Psi} - \Psi \xrightarrow{p} 0 \) as \( N, T \to \infty, N/T \to C \).
4 Models with heterogeneous slopes

We now turn our focus on models with heterogeneous coefficients:

\[ y_i = X_i \beta_i + H^0 \varphi^0_i + \varepsilon_i, \]
\[ X_i = F^0 \Gamma^0_i + V_i. \]  

(4.1)

We first consider the following individual-specific estimator

\[ \hat{\beta}_i = (X_i' \mathbf{M}_F X_i)^{-1} X_i' \mathbf{M}_F y_i. \]

Proposition 4.1 Under Assumptions A-E, for each \( i \) we have

\[ \sqrt{T}(\hat{\beta}_i - \beta_i) = (T^{-1} X_i' \mathbf{F}_0 X_i)^{-1} \times T^{-1/2} X_i' \mathbf{F}_0 u_i + O_p(\delta^{-1}) + O_p\left(T^{1/2} \delta^{-2}\right) \]

and

\[ T^{-1/2} X_i' \mathbf{F}_0 u_i \xrightarrow{d} N(0, \Omega_i) \]

where \( \Omega_i = T^{-1} \lim_{T \to \infty} \sum_{t=1}^{T} \sum_{t=1}^{T} \bar{u}_{it} \bar{u}_{it} E(v_{it} v_{it}') \) and \( \bar{u}_i = \mathbf{M}_F u_i \equiv (\bar{u}_{i1}, \ldots, \bar{u}_{iT})' \).

We also consider inference on the mean of \( \beta_i \). We make the following assumptions.

Assumption G (random coefficients) \( \beta_i = \beta + e_i \), where \( e_i \) is independently and identically distributed over \( i \) with mean zero and variance \( \Sigma_\beta \). Furthermore, \( e_i \) is independent with \( \mathbf{F}_j, \varphi_j^0, \varepsilon_{jt}, v_{jt}, f_t^0 \) and \( h_t^0 \) for all \( i, j, t \).

Assumption H (moments) For each \( i \),

\[ E\|T^{-1/2} \varepsilon_i' \mathbf{F}_0\|^4 \leq C, \quad E\|T^{-1/2} \varepsilon_i' \Sigma \mathbf{F}_0\|^4 \leq C, \]

\[ E\|T^{-1/2} \sum_{t=1}^{T} \varepsilon_i' \varepsilon_t' \mathbf{F}_0\|^4 \leq C, \quad E\|T^{-1/2} \sum_{t=1}^{T} [\mathbf{V}_i' \mathbf{V}_i - \Sigma]\|^4 \leq C, \]

\[ E\|N^{-1/2} T^{-1/2} \sum_{t=1}^{T} (\mathbf{V}_i' \mathbf{V}_t - E(\mathbf{V}_i' \mathbf{V}_t)) \mathbf{F}_t\|^4 \leq C, \quad \text{and} \quad 0 < C_{\min} \leq \|\Sigma\| \leq C. \]

We propose the following mean-group IV (MGIV) estimator:

\[ \hat{\beta}_{MGIV} = N^{-1} \sum_{i=1}^{N} \hat{\beta}_i. \]  

(4.2)

Theorem 4.1 Under Assumptions A-E and G-H, we have

\[ \sqrt{N}(\hat{\beta}_{MGIV} - \beta) = N^{-1/2} \sum_{i=1}^{N} e_i + O_p(N^{3/4} T^{-1}) + O_p(N T^{-3/2}) + O_p(N^{1/2} \delta^{-2}), \]

such that for \( N^3 / T^4 \to 0 \) as \( N, T \to \infty \), we obtain

\[ \sqrt{N}(\hat{\beta}_{MGIV} - \beta) \xrightarrow{d} N(0, \Sigma_\beta). \]

Furthermore, \( \hat{\Sigma}_\beta - \Sigma_\beta \xrightarrow{p} 0 \), where

\[ \hat{\Sigma}_\beta = \frac{1}{N-1} \sum_{i=1}^{N} (\hat{\beta}_i - \hat{\beta}_{MGIV})(\hat{\beta}_i - \hat{\beta}_{MGIV})'. \]  

(4.3)
5 Asymptotic comparison of $\hat{\beta}_{2SIV}$ with existing estimators

This section investigates asymptotic bias properties and relative efficiency of the 2SIV, IPC, PC and CA estimators for the models with homogeneous slopes. For this purpose, let $G^0 = (F^0, H^0)$ denote a $T \times r$ matrix. We shall assume that $G^0G^0/T \cdot \sum_G > 0$, a positive definite matrix. Note that, together with Assumption C, this implies that $F^0$ and $H^0$ are linearly independent of each other (and can be correlated), which is slightly stronger than Assumption C.

5.1 2SIV estimator

Recall that $X_i = F^0\gamma_i + V_i$ and $u_i = H^0\phi_i + \varepsilon_i$. Proposition 3.2 in Appendix B demonstrates that under Assumptions A-E $\left(N^{-1}T^{-1} \sum_{i=1}^{N} X_i'M_F'M_H u_i\right) \sqrt{NT} \left(\hat{\beta}_{2SIV} - \beta\right)$ can be expanded as follows:

$$
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i'M_F'M_H u_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i\varepsilon_i + b_{0FH} + b_{1FH} + b_{2FH} + O_p \left(\sqrt{NT}\delta_{NT}^{-3}\right),
$$

where

$$
b_{0FH} = \frac{1}{N^{1/2}} \frac{1}{NT^{1/2}} \sum_{i=1}^{N} \sum_{j=1}^{N} \left(F_i^0\gamma_i^0\right)^{-1} T_j^0 + \phi_j^0 \left(F_i^0\gamma_i^0\right)^{-1} \phi_j^0 V_j\varepsilon_i;
$$

$$
b_{1FH} = \frac{1}{N^{1/2}} \frac{1}{NT^{1/2}} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{l=1}^{N} \Gamma_i^0 \left(F_i^0\gamma_i^0\right)^{-1} \Gamma_j^0 \left(V_j \varepsilon_j\right) \phi_j^0 \left(F_i^0\gamma_i^0\right)^{-1} \phi_j^0;
$$

$$
b_{2FH} = \frac{1}{T^{1/2}} \frac{1}{N^{3/2}T} \sum_{i=1}^{N} \sum_{j=1}^{N} \Gamma_i^0 \left(F_i^0\gamma_i^0\right)^{-1} \Gamma_j^0 V_j^0 \Sigma_\varepsilon \left(\frac{H_i^0H_j^0}{T}\right)^{-1} \left(F_i^0\gamma_i^0\right)^{-1} \phi_j^0;
$$

with $\Sigma_\varepsilon = \frac{1}{T} \sum_{i=1}^{N} \left(\varepsilon_i \varepsilon_i^T\right)$. It is easily seen that (see proof of Proposition 3.2) $b_{0FH} = O_p \left(N^{-1/2}\right)$, $b_{1FH} = O_p \left(N^{-1/2}\right)$ and $b_{2FH} = O_p \left(T^{-1/2}\right)$.

Hence, we have

$$
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i'M_F'M_H u_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i\varepsilon_i + o_p (1).
$$

5.2 Asymptotic bias of Bai’s (2009a) IPC-type estimator

It is instructive to consider a PC estimator that is asymptotically equivalent to Bai (2009) but avoids iterations:

$$
\tilde{\beta}_{2SIV} = \left(\sum_{i=1}^{N} X_i'M_H X_i\right)^{-1} \sum_{i=1}^{N} X_i'M_H Y_i.
$$

Observe that this estimator projects out $\hat{H}$ from $(X_i, y_i)$, but it does not eliminate $\hat{F}$ from $X_i$. $\tilde{H}$ is estimated using the residuals of the first-stage IV estimator, $\hat{u}_i = y_i - X_i\hat{\beta}_{1SIV}$.

Using similar derivations as in Section 5.1, Proposition 5.1 below shows that $\left(N^{-1}T^{-1} \sum_{i=1}^{N} X_i'M_H X_i\right) \times \sqrt{NT} \left(\hat{\beta}_{2SIV} - \beta\right)$ has the following asymptotic expansion:

**Proposition 5.1** Under Assumptions A-E, we have

$$
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i'M_H u_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i'M_H \varepsilon_i + b_{0H} + b_{1H} + b_{2H} + O_p \left(\sqrt{NT}\delta_{NT}^{-3}\right)
$$

(5.2)
where

\[ b_{0H} = -\frac{1}{N^{3/2}T^{1/2}} \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} X'_j M_{H^0} \varepsilon_i; \]
\[ b_{1H} = -\frac{T}{N N T} \sum_{i=1}^{N} \sum_{j=1}^{N} X'_i H^0 \left( \frac{H^0 H^0}{T} \right)^{-1} (Y^0_\varphi)^{-1} \varphi_i^0 \mathbb{E} (\varepsilon'_i \varepsilon_i / T); \]
\[ b_{2H} = -\frac{T}{N T N T} \sum_{i=1}^{N} X'_i M_{H^0} \Sigma_\varepsilon H^0 \left( \frac{H^0 H^0}{T} \right)^{-1} (Y^0_\varphi)^{-1} \varphi_i^0, \]

with

The above asymptotic bias terms are identical to those of the IPC estimator of Bai (2009). As a result, it suffices to compare \( \tilde{\beta}_{2SIV} \) with \( \beta_{2SIV} \). Incidentally, as shown in Bai (2009), the term \( b_{0H} \) tends to a normal random vector, which necessitates the transformation of the regressor matrix to \( X'_i \); see equation (5.3) below.

The terms \( b_{0H}, b_{1H} \) and \( b_{2H} \) in (5.2) are comparable to the terms \( b_{0FH}, b_{1FH} \) and \( b_{2FH} \), respectively, in (5.1). One striking result is that \( b_{0H}, b_{1H} \) and \( b_{2H} \) are not asymptotically ignorable, whereas \( b_{0FH}, b_{1FH} \) and \( b_{2FH} \) are. This difference stems solely from the fact that \( \tilde{\beta}_{2SIV} \) asymptotically projects out \( F^0 T'_i \) from \( X_i \) and \( H^0 \varphi_i^0 \) from \( u_i \) separately, whereas \( \beta_{2SIV} \) projects out \( H^0 \varphi_i^0 \) from \( u_i \) only. Therefore, the asymptotic bias terms of \( \tilde{\beta}_{2SIV}, b_{0H}, b_{1H} \) and \( b_{2H} \), contain correlations between the regressors \( X_i \) and the disturbance \( u_i(= H^0 \varphi_i^0 + \varepsilon_i) \) since the estimation error of \( \tilde{H} \) contains \( u_i \). Recalling that \( X_i = F^0 T'_i + V_i \), such correlations are asymptotically non-negligible because \( H^0 F^0 / T = O_p(1) \) and \( \sum_{i=1}^{N} \varphi_i^0 \text{vec}(\Gamma_i^0) / N = O_p(1) \).

On the other hand, \( \tilde{\beta}_{2SIV} \) asymptotically projects out \( F^0 T'_i \) from \( X_i \) as well as \( H^0 \varphi_i^0 \) from \( u_i \). Therefore, \( b_{0FH}, b_{1FH} \) and \( b_{2FH} \) contain correlations between \( M_{F^0} X_i = M_{F^0} V_i \) and \( u_i \). Since \( V_i, H^0 \varphi_i^0 \) and \( \varepsilon_i \) are independent of each other, such correlations are asymptotically negligible. As a result, our estimator \( \tilde{\beta}_{2SIV} \) does not suffer from asymptotic bias.

Using similar reasoning, it turns out that in some special cases, some of the bias terms of \( \tilde{\beta}_{2SIV} \) may disappear as well. For instance, when \( F^0 \subseteq H^0 \), we have \( M_{H^0} X_j = M_{H^0} V_j \) because \( M_{H^0} F^0 = 0 \). Thus, \( b_{0H} = O_p \left( N^{-1/2} \right) \) and \( b_{2H} = O_p \left( T^{-1/2} \right) \), although \( b_{1H} \) remains \( O_p(1) \). Note that under our assumptions all three bias terms, \( b_{0H}, b_{1H} \) and \( b_{2H} \), are asymptotically negligible only if \( H^0 = F^0 \), which can be a highly restrictive condition in practice.\(^4\)

### 5.3 Asymptotic bias of PC and CA estimators

Pesaran (2006) and Westerlund and Urbain (2015) put forward pooled estimators in which the whole set of factors in \( X_i \) and \( u_i \) are estimated jointly, rather than separately. This difference makes these estimators asymptotically biased. To show this, we rewrite the model as

\[ Z_i = (y_i, X_i) = G^0 \Lambda^0_i + U_i, \]

where

\[ \Lambda^0_i = \begin{pmatrix} \Gamma^0_i \beta \\ \varphi^0_i \\ 0 \end{pmatrix}, \quad U_i = (V_i \beta + \varepsilon_i, V_i). \]

\(^4\)When \( \varepsilon_i \sim i.i.d.(0, \sigma^2) \), \( b_{0H} \) remains \( O_p(1) \) whilst \( b_{1H} \) and \( b_{2H} \) become asymptotically negligible. See Corollary 1 in Bai (2009).
Define

$$\mathbf{Y}_A^0 = N^{-1} \sum_{i=1}^N \mathbf{A}_i^0 \mathbf{A}_i^0 = \left( N^{-1} \sum_{i=1}^N \mathbf{I}_i^0 \left( \beta \beta' + \mathbf{I}_i \right) \mathbf{I}_i^0 + \mathbf{Y}_A^0 \right) N^{-1} \sum_{i=1}^N \phi_i^0 \beta \mathbf{T}_i^0 \).$$

In the PC approach of Westerlund and Urbain (2015), a span of $\mathbf{G}^0$ is estimated as $\sqrt{T}$ times the eigenvectors corresponding to the first $r$ largest eigenvalues of $\sum_{i=1}^N \mathbf{Z}_i \mathbf{Z}_i'/N$, which is denoted by $\hat{\mathbf{G}}_z$. The resulting PC estimator is defined as

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

In line with Pesaran (2006), the CA estimator of Westerlund and Urbain (2015) approximates a span of $\mathbf{G}^0$ by a linear combination of $\bar{\mathbf{Z}} = N^{-1} \sum_{i=1}^N \mathbf{Z}_i$. The associated CA estimator is given by

$$\hat{\beta}_{CA} = \left( \sum_{i=1}^N \mathbf{X}_i' \mathbf{M}_{\bar{\mathbf{Z}}} \mathbf{X}_i \right)^{-1} \sum_{i=1}^N \mathbf{X}_i' \mathbf{M}_{\bar{\mathbf{Z}}} \mathbf{y}_i.$$

As discussed in Westerlund and Urbain (2015), both PC and CA are asymptotically biased due to the correlation between the estimation error of $\hat{\mathbf{G}}_z$ and $\{ \mathbf{X}_i, \mathbf{u}_i \}$. The estimation error of $\hat{\mathbf{G}}_z$ contains the error term of the system equation $\mathbf{U}_i$, which is a function of both $\mathbf{V}_i$ and $\mathbf{e}_i$. Therefore, the estimation error of $\hat{\mathbf{G}}_z$ is correlated with $\mathbf{M}_\mathbf{G} \mathbf{X}_i$ and $\mathbf{M}_\mathbf{G} \mathbf{u}_i$, which causes the asymptotic bias. In what follows, we shall focus on the PC estimator as the bias analysis for the CA estimator is very similar.

Following Westerlund and Urbain (2015), we expand $(N^{-1}T^{-1} \sum_{i=1}^N \mathbf{X}_i' \mathbf{M}_{\hat{\mathbf{G}}_z} \mathbf{X}_i) \sqrt{NT} \left( \hat{\beta}_{PC} - \beta \right)$ as follows:

**Proposition 5.2** Under Assumptions A-E

$$\frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{X}_i' \mathbf{M}_{\hat{\mathbf{G}}_z} \mathbf{u}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathbf{V}_i' \mathbf{e}_i + b_{1G} + b_{2G} + b_{3G} + O_p \left( \sqrt{NT} \delta_{NT}^{-3} \right),$$

$$b_{1G} = -\sqrt{T} \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \left( \mathbf{r}_i^0, 0' \right) \left( \mathbf{Y}_A^0 \right)^{-1} \mathbf{A}_i^0 \mathbf{E} \left( \mathbf{U}_i' \mathbf{e}_i / T \right);$$

$$b_{2G} = -\sqrt{T} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \left( \mathbf{r}_i^0, 0' \right) \left( \mathbf{r}_j^0 \right)^{-1} \mathbf{A}_i^0 \mathbf{E} \left( \mathbf{U}_i' \mathbf{U}_j / T \right) \mathbf{A}_j^0 \left( \mathbf{Y}_A^0 \right)^{-1} \left( \frac{\mathbf{G}^0 \mathbf{G}^0}{T} \right)^{-1} \frac{\mathbf{G}^0 \mathbf{H}^0}{T} \phi_i^0;$$

$$b_{3G} = -\sqrt{T} \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N \sum_{k=1}^N \mathbf{V}_i' \mathbf{U}_j \left( \mathbf{Y}_A^0 \right)^{-1} \left( \frac{\mathbf{G}^0 \mathbf{G}^0}{T} \right)^{-1} \frac{\mathbf{G}^0 \mathbf{H}^0}{T} \phi_i^0.$$

It is easily seen that $b_{1G}, b_{2G}$ and $b_{3G}$ are all $O_p(1)$. Note that the asymptotic bias terms are functions of $\mathbf{A}_i^0$ and $\mathbf{Y}_A^0$, which depend on the slope coefficient vector $\beta$.

### 5.4 Relative asymptotic efficiency of 2SIV, IPC, PC and CA estimators

Finally, we compare the asymptotic efficiency of the estimators. To make the problem tractable and as succinct as possible, we shall assume that $\mathbf{e}_{it}$ is i.i.d. over $i$ and $t$ with $\mathbf{E}(\mathbf{e}_{it}) = 0$ and
\[ \mathbb{E}(\varepsilon_i^2) = \sigma^2 \varepsilon. \] In this case, it is easily seen that the asymptotic variance of \( \hat{\beta}_{2SIV} \) is
\[ \Psi = \sigma^2 \left( \text{plim} N^{-1} T^{-1} \sum_{i=1}^{N} V_i' V_i \right)^{-1}. \]

Next, using Proposition 5.2, consider the bias-corrected PC estimator
\[ \hat{\beta}_{PC} = \hat{\beta}_{PC} - N^{1/2} T^{1/2} \left( \sum_{i=1}^{N} V_i' V_i \right)^{-1} (b_{1G} + b_{2G} + b_{3G}). \]

We can see that the asymptotic variance of the bias-corrected PC estimator is identical to \( \Psi \). Therefore, the 2SIV and the bias-corrected PC estimators are asymptotically equivalent.

Consider now \( \tilde{\beta}_{2SIV} \). Noting that \( b_{0H} \) tends to a normal distribution, and following Bai (2009), the bias-corrected estimator with transformed regressors can be expressed as:
\[ \tilde{\beta}_{2SIV}^* = \tilde{\beta}_{2SIV} - N^{1/2} T^{1/2} \left( \sum_{i=1}^{N} X_i' M_{H^0} X_i \right)^{-1} (b_{1H} + b_{2H}), \]

where
\[ \tilde{\beta}_{2SIV} = \left( \sum_{i=1}^{N} X_i' M_{H^0} X_i \right)^{-1} \sum_{i=1}^{N} X_i' M_{H^0} y_i. \] (5.3)

The asymptotic variance of this bias-corrected estimator is given by
\[ \tilde{\Psi} = \sigma^2 \left( \text{plim} N^{-1} T^{-1} \sum_{i=1}^{N} X_i' M_{H^0} X_i \right)^{-1}. \]

There exist two differences compared to \( \Psi \). First, in general \( M_{H^0} X_i \neq M_{H^0} V_i \) as the factors in \( X_i \) may not be identical to the factors in \( u_i \). Second, regressors are to be transformed as \( X_i = X_i - N^{-1} \sum_{\ell=1}^{N} a_{it} X_\ell \). Therefore, \( \Psi - \tilde{\Psi} \) can be positive semi-definite or negative semi-definite. Thus, the asymptotic efficiency of the bias-corrected IPC estimator of Bai (2009) relative to 2SIV and the bias-corrected PC/CA estimators, is indeterminate. However, in the special case where \( F^0 \subseteq H^0 \), we have \( M_{H^0} X_i = M_{H^0} V_i \), with \( V_i = V_i - N^{-1} \sum_{\ell=1}^{N} a_{it} V_\ell \). The second term of \( V_i \) is \( O_p(N^{-1/2}) \) because \( V_\ell \) and \( a_{it} \) are independent. Hence, in this case \( \tilde{\Psi} = \Psi \), and the bias-corrected IPC estimator is asymptotically as efficient as the bias-corrected PC/CA estimator and 2SIV.

### 6 Monte Carlo Simulations

We conduct a small-scale Monte Carlo simulation exercise in order to assess the finite sample behaviour of the proposed approach in terms of bias, standard deviation (s.d.), root mean squared error (RMSE), empirical size and power of the t-test. More specifically, we investigate the performance of 2SIV, defined in (3.2), and MGIV defined in (4.2). For the purposes of comparison, we also consider the (bias-corrected) IPC of Bai (2009) and the PC estimator, labeled as (BC-)IPC and (BC-)PC respectively, the CA estimator, as well as the mean-group versions of PC and CA (denoted as MGPC and MGCA), which were put forward by Pesaran (2006), Westerlund and Urbain (2015) and Reese and Westerlund (2018). The t-statistics for 2SIV and MGIV are computed using the variance estimators defined by (3.5) and (4.3), respectively. The t-statistics for IPC, PC and CA estimators and their MG versions (if any) employ analogous variance estimators.
6.1 Design

We consider the following panel data model:

\[ y_{it} = \alpha_i + \sum_{\ell=1}^{k} \beta_{\ell i} x_{\ell it} + u_{it}; \quad u_{it} = \sum_{s=1}^{m_y} \gamma_{si}^0 f_{s,t}^0 + \varepsilon_{it}, \quad (6.1) \]

\( i = 1, ..., N, \quad t = -49, ..., T, \) where the process for the covariates is given by

\[ x_{\ell it} = \mu_{\ell i} + \sum_{s=1}^{m_x} \gamma_{\ell st}^0 f_{s,t}^0 + v_{\ell it}; \quad i = 1, 2, ..., N; \quad t = -49, -48, ..., T. \quad (6.2) \]

We set \( k = 2, \) \( m_y = 2 \) and \( m_x = 3. \) This implies that the first two factors in \( u_{it}, f_{1t}^0 \) and \( f_{2t}^0, \) are also in the DGP of \( x_{\ell it} \) for \( \ell = 1, 2, \) while \( f_{1t}^0 \) is included in \( x_{\ell it} \) only. Observe that, using notation of earlier sections, \( h_i^0 = (f_{1t}^0, f_{2t}^0)' \) and \( f_i^0 = (f_{1t}^0, f_{2t}^0, f_{3t}^0)' \).

The factors \( f_{s,t}^0 \) are generated using the following AR(1) process:

\[ f_{s,t}^0 = \rho_{fs} f_{s,t-1}^0 + (1 - \rho_{fs}^2)^{1/2} \zeta_{s,t}, \quad (6.3) \]

where \( \rho_{fs} = 0.5 \) and \( \zeta_{s,t} \sim i.i.d. N(0, 1) \) for \( s = 1, ..., 3. \)

The idiosyncratic error of \( y_{it}, \varepsilon_{it}, \) is non-normal and heteroskedastic across both \( i \) and \( t, \) such that \( \varepsilon_{it} = \zeta \sigma_i (\varepsilon_{it} - 1) / \sqrt{2}, \) \( \varepsilon_{it} \sim i.i.d. \chi_1^2, \) with \( \sigma_i^2 = \eta_i \varphi_t, \) \( \eta_i \sim i.i.d. \chi_2^2 / 2, \) and \( \varphi_t = t / T \) for \( t = 0, 1, ..., T \) and unity otherwise. We define \( \pi_u := \varepsilon / (m_y + \varepsilon^2) \) which is the proportion of the average variance of \( u_{it} \) due to \( \varepsilon_{it}. \) This implies \( \varepsilon^2 = \pi_u m_y (1 - \pi_u)^{-1}. \) We set \( \varepsilon^2 \) such that \( \pi_u \in \{1/4, 3/4\}. \)

The idiosyncratic errors of the covariates follow an AR(1) process

\[ v_{\ell it} = \rho_{v,\ell i} v_{\ell it-1} + (1 - \rho_{v,\ell i}^2)^{1/2} w_{\ell it}; \quad w_{\ell it} \sim i.i.d. N(0, \varepsilon_{\ell}^2), \quad (6.4) \]

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

We define the signal-to-noise ratio (SNR) as \( SNR := (\beta_1^2 + \beta_2^2) \varepsilon^2 \) where \( \rho_{\ell} := \rho_{v,\ell} \) for \( \ell = 1, 2. \) Solving for \( \varepsilon^2 \) gives \( \varepsilon^2 = \varepsilon\varepsilon^2 \) \( SNR (\beta_1^2 + \beta_2^2)^{-1}. \) We set \( SNR = 4, \) which lies within the values considered by Bun and Kiviet (2006) and Juodis and Sarafidis (2018).

The individual-specific effects are generated by drawing initially mean-zero random variables as

\[ \mu_{\ell i}^* = \rho_{\mu,\ell i} \alpha_i^* + (1 - \rho_{\mu,\ell i}^2)^{1/2} \omega_{\ell i}, \quad (6.5) \]

where \( \alpha_i^* \sim i.i.d. N(0,1), \) \( \omega_{\ell i} \sim i.i.d. N(0,1), \) for \( \ell = 1, 2. \) We set \( \rho_{\mu,\ell} = 0.5 \) for \( \ell = 1, 2. \) Subsequently, we set

\[ \alpha_i = \alpha + \alpha_i^*, \quad \mu_{\ell i} = \mu + \mu_{\ell i}^*, \quad (6.6) \]

where \( \alpha = 1/2, \mu_1 = 1, \mu_2 = -1/2, \) for \( \ell = 1, 2. \)

Similarly, the factor loadings in \( u_{it} \) are generated at first instance as mean-zero random variables such that \( \gamma_{si}^0 \sim i.i.d. N(0,1) \) for \( s = 1, ..., m_y = 2, \ell = 1, 2. \) the factor loadings in \( x_{1it} \) and \( x_{2it} \) are generated as

\[ \gamma_{\ell s i} = \rho_{\gamma,\ell i s} \gamma_{si}^0 + (1 - \rho_{\gamma,\ell s i}^2)^{1/2} \xi_{\ell s i}; \quad \xi_{\ell s i} \sim i.i.d. N(0,1); \quad (6.7) \]

\[ 5\text{Tables E1-E3 in Appendix E present results for a different specification, where } m_y = 3 \text{ and } m_x = 2. \text{ To save space, we do not discuss these results here but it suffices to say that the conclusions are similar to those in Section 6.2.} \]
\[ \gamma_{13t}^0 = \rho_{\gamma,13} \gamma_{11t}^0 + (1 - \rho_{\gamma,13}^2)^{1/2} \xi_{13t}; \quad \xi_{13t} \sim i.i.d. N(0,1); \]  
\[ \gamma_{23t}^0 = \rho_{\gamma,23} \gamma_{22t}^0 + (1 - \rho_{\gamma,23}^2)^{1/2} \xi_{23t}; \quad \xi_{23t} \sim i.i.d. N(0,1). \]  
(6.8)  
(6.9)

The process (6.7) allows the factor loadings to fit the within transformation in order to eliminate individual-specific effects. For the CA and MGCA alternatives are considered.

The results are obtained based on 2,000 replications, and all tests are conducted at the 5% significance level. On the other hand, (6.8) and (6.9) ensure that the factor loadings to fit the within transformation in order to eliminate individual-specific effects.

\[ \rho_i \sim i.i.d. U[-c, +c], \]  
\[ (6.8) \]  
\[ (6.9) \]

Prior to computing the estimators except for CA and MGCA, the data are demeaned using the within transformation in order to eliminate individual-specific effects. For the CA and MGCA estimators, the untransformed data are used, but a \( T \times 1 \) vector of ones is included along with the cross-sectional averages. The number of factors \( m_x \) and \( m_y \) are estimated in each replication using the eigenvalue ratio (ER) statistic proposed by Ahn and Horenstein (2013).
6.2 Results

Tables 1–3 report results for $\beta_1$ in terms of bias, standard deviation, RMSE, empirical size and power for the model in (6.1).\footnote{The results for $\beta_2$ are qualitatively similar and so we do not report them to save space. These results are available upon request.}

Table 1 focuses on the case where $N = T = 200$ and $\pi_u$ alternates between $\{1/4, 3/4\}$. Consider first the homogeneous model with $\pi_u = 3/4$. As we can see, the bias ($\times 100$) for 2SIV and MGIV is very close to zero and takes the smallest value compared to the remaining estimators. The bias of BC-IPC is larger in absolute value than that of IPC but of opposite sign. This may suggest that bias-correction over-corrects in this case. MGPC and PC perform similarly and exhibit larger bias than IPC. Last, both CA and MGCA are subject to substantial bias, which is not surprising as these estimators may require bias-correction in the present DGP.

In regards to the dispersion of the estimators, the standard deviation of 2SIV and PC is very similar, which is in line with our theoretical results. For this specific design, IPC takes the smallest s.d. value among the estimators under consideration. On the other hand, when it comes to the bias-corrected estimators, bias-correction appears to inflate dispersion and thus the standard deviation of BC-IPC and BC-PC is relatively large (equal to 0.805 and 0.885, respectively). As a result, 2SIV outperforms BC-IPC and BC-PC, with a s.d. value equal to 0.586.

In terms of RMSE, IPC appears to perform best, although this estimator is not recommended in practice due to its asymptotic bias. 2SIV takes the second smallest RMSE value, followed by MGIV. CA and MGCA exhibit the largest RMSE values, an outcome that reflects the large bias of these estimators.

Next, we turn our attention to the model with heterogeneous slopes and $\pi_u = 3/4$. In comparison to the homogeneous model, all estimators suffer a substantial increase in bias; the only exception is MGIV, which has the smallest bias. MGPC and MGCA are severely biased, both in absolute magnitude as well as relative to the remaining inconsistent estimators. The s.d. values of MGIV and MGPC are very similar and relatively small compared to the other estimators. The smallest RMSE value is that of MGIV.

We now discuss the results in the lower panel of Table 1, which correspond to $\pi_u = 1/4$. The relative performance of the estimators is similar to the case where $\pi_u = 3/4$, except for a noticeable improvement in the performance of BC-IPC. Thus, the results for BC-IPC and IPC are quite comparable, suggesting that the bias-correction term is close to zero and so over-correction is avoided. The results for 2SIV are very similar to those for $\pi_u = 3/4$, which indicates that the estimator is robust to different values of the variance ratio. The conclusions with heterogeneous slopes for $\pi_u = 1/4$ are similar to those for $\pi_u = 3/4$.

In regards to inference, the size of the t-test associated with 2SIV and MGIV is close to the nominal value of 5% under the setting of homogeneous slopes. The same appears to hold true for BC-IPC when $\pi_u = 1/4$, although there are substantial distortions when $\pi_u = 3/4$. The t-test associated with BC-PC is oversized when $\pi_u = 3/4$ and the distortion becomes more severe with $\pi_u = 1/4$. CA and MGCA have the largest size distortions. In the case of heterogeneous slopes, MGIV performs well and size is close to 5%. MGPC and MGCA have substantial size distortions regardless of the value of $\pi_u$.

Table 2 presents results for the case where $(N, T) = (200, 25)$ (i.e. $N$ is large relative to $T$)
and \((N, T) = (25, 200)\) \((N\) is small relative to \(T)\) for \(\pi_u = 3/4\). In the former case, 2SIV performs best in terms of bias. IPC has the smallest RMSE, followed by 2SIV. CA has the largest bias and RMSE. In the case of heterogeneous slopes, MGIV has smaller absolute bias than MGPC and MGCA. Therefore, MGIV is superior among mean-group type estimators, which are the only consistent estimators in this design. In the case where \(T\) is large relative to \(N\), 2SIV and MGIV again outperform BC-IPC, BC-PC and CA in terms of bias, standard deviation and RMSE.

In regards to the properties of the t-test, 2SIV and MGIV have the smallest size distortions relative to the other estimators, and inference based on 2SIV and MGIV remains credible even for small values of \(N\) or \(T\). Moreover, 2SIV and MGIV exhibit good power properties, whereas MGPC has the lowest power when \(N\) is small relative to \(T\).

Table 3 shows the bias of the estimators, scaled by \(\sqrt{NT} / (\sqrt{N})\) for different values of \(N = T\) with \(\pi_u = \{1/4, 3/4\}\) when the slopes are homogeneous (heterogeneous). The performance of 2SIV and MGIV is in agreement with our theoretical results. More specifically, the bias monotonically decreases as the sample size goes up. In contrast, for \(\pi_u = 3/4\) it appears that a relatively large sample size is necessary so that bias-correction works for BC-IPC. BC-PC appears to require even larger sample sizes.

In a nutshell, the results presented in Tables 1-3 and the associated discussion above suggest that 2SIV and MGIV have good small sample properties and outperform existing popular estimators for the experimental designs considered here.

## 7 Illustration

In this section we apply our methodology to study the effect of climate shocks on economic growth using an unbalanced panel of 125 countries over the period 1961-2003. The data set is taken from Dell et al. (2012).

In line with existing literature (e.g. Dell et al. (2014)), we consider the following benchmark static panel data model:

\[
g_{it} = \beta_1 \text{temp}_{it} + \gamma_1 \text{prec}_{it} + \beta_2 D_i \text{temp}_{it} + \gamma_2 D_i \text{prec}_{it} + \eta_i + \tau_t + u_{it}, \tag{7.1}
\]

where \(g_{it}\) denotes the growth rate of per-capita output for country \(i\) at year \(t\), while \(\text{temp}_{it}\) and \(\text{prec}_{it}\) denote the level of temperature (in degrees Celcius) and precipitation (in units of 100 mm) for country \(i\) at year \(t\), \(i = 1, \ldots, 126\), \(t = 1, \ldots, T_i\), where \(\min\{T_i\} = 21\), \(\max\{T_i\} = 43\) and \(\bar{T} = N^{-1} \sum_{i=1}^{N} T_i \approx 40\). \(D_i\) denotes a binary variable that equals one if the country \(i\) is characterised as “developing” and zero if it is characterised as “developed”. Thus, \(\beta_1\) and \(\gamma_1\) reflect the effect of temperature and precipitation, respectively, on economic growth rate for developed economies, whereas \(\beta_2\) and \(\gamma_2\) capture the corresponding differential effects between developing and developed economies. The main reason for such a distinction is that developing economies are often reliant on agriculture or outdoor activities, and therefore they are vulnerable to climate shocks. Following Dell et al. (2012), in the present application a country is defined as developing if it has below-median PPP-adjusted per capita GDP in the first year the country enters the dataset, otherwise it is defined as developed.

As it is common practice in the literature (e.g. Colacito et al. (2018)), we include country effects \(\eta_i\) and year effects, \(\tau_t\). In addition to these additive effects, we also allow for unobserved
interactive effects. This offers wider scope for controlling for omitted variables, including situations where there is cross-sectional dependence. In particular, \( u_{it} \) is given by

\[
u_{it} = \phi_0^i h_0^t + \varepsilon_{it},\]

where \( h_0^t \) is a \( r_2 \times 1 \) vector of year-specific unobserved common shocks with corresponding country-specific loadings given by \( \phi_0^i \), whereas \( \varepsilon_{it} \) is a purely idiosyncratic error.

We employ four estimators; namely, the two-stage IV (2SIV) estimator analysed in this paper, a fixed effects (FE) estimator that allows for two-way clustering per country and region-year\(^7\), as in Cameron et al. (2011), the pooled common correlated effects (CA) estimator of Pesaran (2006), and the iterative principal components (IPC) estimator of Bai (2009).

The FE estimator imposes \( u_{it} = \varepsilon_{it} \) by construction, i.e. it assumes there exists no factor structure. However, in order to try and neutralise the effect of common shocks, we follow Dell et al. (2012) and include year fixed effects interacted with region dummies, as well as year fixed effects interacted with the developing country dummy.\(^8\) For 2SIV, the number of factors in regressors and \( u_{it}, r_1 \) and \( r_2 \) respectively, is estimated using the eigenvalue ratio test of Ahn and Horenstein (2013). In order to carry out a specification test of our model (the well-known overidentifying restrictions J-test), we make use of present and lagged values of all defactored regressors as instruments. Thus, the total number of instruments equals 8. Since temperature and precipitation are measured in rather different units, we defactor these variables separately. The CA estimator is implemented using year-specific cross-sectional averages of all regressors. For the IPC estimator of Bai (2009), the number of factors in \( u_{it} \) is estimated using the Bai and Ng (2002) model information criteria.

The results are presented in Table 4. For all estimators, we run two different models. Column (A) corresponds to a specification that imposes the restriction \( \beta_2 = \gamma_2 = 0 \); that is, developing and developed economies are pooled together. The estimates of the coefficients of temperature and precipitation are negative across all four estimators. However, the temperature effect is statistically significant only for 2SIV and CA, both at the 10% level. Moreover, the precipitation coefficient is statistically significant only for 2SIV (at the 5% level). The J-test statistic of 2SIV rejects the specification of the model, which implies that developing and developed economies may be affected in a different manner.

Column (B) corresponds to the specification in (7.1), which relaxes the pooling restriction. In this case, FE replicates the main panel results reported in Table 4, Column (3), by Dell et al. (2012). As we can see, the effect of temperature on growth appears to be positive for developed economies and highly negative for developing ones, indicating substantial heterogeneity between the two groups. However, the estimate of \( \beta_1 \) is statistically significant only for 2SIV and IPC but not FE or CA. Thus, the implications are substantially different. In particular, the results obtained by 2SIV indicate that a 1°C rise in temperature appears to increase (decrease) growth rates for developed (developing) economies by .530 (1.42) percentage points (hereafter, p.p.), all other things being equal.\(^9\) The specification of the model is not rejected by the J-test statistic. The estimated effect of temperature obtained by IPC is somewhat smaller in absolute magnitude than that of 2SIV, both in terms of developing and developed economies. Last, for FE (CA) the estimated coefficients

---

\(^7\)See footnote 12 in Dell et al. (2012) p. 74 for the definition of geographical regions.  
\(^8\)Thus, \( \tau_t \) cannot be separately identified per se.  
\(^9\)The latter estimate is obtained by adding \( \hat{\beta}_1 \) and \( \hat{\beta}_2 \).
indicate that a $1^\circ$C rise in temperature decreases the growth rate for developing economies by 1.61 (1.76) p.p., whereas it does not exert a statistically significant impact on developed economies.

In regards to precipitation, the results obtained by 2SIV indicate that an extra 100 mm of annual rainfall is expected to decrease growth rates for both developed and developing economies by approximately .08 p.p. all other things being equal. On the other hand, the estimated effect of precipitation obtained from IPC and CA is not statistically significant for either group of economies. Finally, for FE the estimated precipitation effect for developed economies is very close to that obtained by 2SIV. However, the estimated precipitation effect is significantly positive for developing economies.

8 Conclusions

We put forward IV estimators for linear panel data models with interactive effects in the error term and regressors. The instruments are transformed regressors, and so it is not necessary to search for external instruments. Models with homogeneous and heterogeneous slope coefficients have been considered. In the former model, we propose a two-stage IV estimator. In the first stage, we asymptotically projects out the interactive effects from the regressors and use the defactored regressors as instruments. In the second stage, we asymptotically eliminate the interactive effects in the error term based on their estimates using the first-stage residuals. We established the $\sqrt{NT}$-consistency and the asymptotic normality of the 2SIV estimator. For the heterogeneous slopes, we put forward a mean-group IV estimator (MGIV) and established $\sqrt{N}$-consistency and asymptotic normality.

Having derived the theoretical properties of our IV estimators, we compared the asymptotic expressions of our 2SIV estimator, IPC of Bai (2009), PC and CA of Westerlund and Urbain (2015) and Pesaran (2006), for the models with homogeneous slopes. Under the conditions similar to those in Bai (2009), it has emerged that 2SIV is free from asymptotic bias, whereas the remaining estimators suffer from asymptotic bias. In addition, it is revealed that 2SIV is asymptotically as efficient as the bias-corrected versions of PC and CA, while the relative efficiency of the bias-corrected IPC estimator is generally indeterminate. The theoretical results are corroborated in a Monte Carlo simulation exercise, which shows that 2SIV and MGIV perform competently and can outperform existing estimators.
Table 1: Bias, root mean squared error (RMSE) of the estimators of $\beta_1$, and size and power of the associated t-tests when $\pi_u = \{1/4, 3/4\}$ and $N = T = 200$.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Homogeneous Slopes</th>
<th>Heterogeneous Slopes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>S.D.</td>
</tr>
<tr>
<td>2SIV</td>
<td>0.003</td>
<td>0.586</td>
</tr>
<tr>
<td>BC-IPC</td>
<td>-0.149</td>
<td>0.805</td>
</tr>
<tr>
<td>IPC</td>
<td>0.020</td>
<td>0.528</td>
</tr>
<tr>
<td>BC-PC</td>
<td>0.306</td>
<td>0.885</td>
</tr>
<tr>
<td>PC</td>
<td>-0.638</td>
<td>0.589</td>
</tr>
<tr>
<td>CA</td>
<td>1.859</td>
<td>0.806</td>
</tr>
<tr>
<td>MGIV</td>
<td>0.000</td>
<td>0.593</td>
</tr>
<tr>
<td>MGPC</td>
<td>-0.650</td>
<td>0.595</td>
</tr>
<tr>
<td>MGCA</td>
<td>1.623</td>
<td>0.722</td>
</tr>
</tbody>
</table>

Table 2: Bias, root mean squared error (RMSE) of the estimators of $\beta_1$, and size and power of the associated t-tests when $\pi_u = 3/4$, $N = 200$, $T = 25$ and $N = 25$, $T = 200$.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>Homogeneous Slopes</th>
<th>Heterogeneous Slopes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>S.D.</td>
</tr>
<tr>
<td>2SIV</td>
<td>0.126</td>
<td>1.941</td>
</tr>
<tr>
<td>BC-IPC</td>
<td>-1.180</td>
<td>2.610</td>
</tr>
<tr>
<td>IPC</td>
<td>0.374</td>
<td>1.870</td>
</tr>
<tr>
<td>BC-PC</td>
<td>0.825</td>
<td>2.746</td>
</tr>
<tr>
<td>PC</td>
<td>-0.211</td>
<td>2.756</td>
</tr>
<tr>
<td>CA</td>
<td>2.084</td>
<td>2.000</td>
</tr>
<tr>
<td>MGIV</td>
<td>0.482</td>
<td>2.534</td>
</tr>
<tr>
<td>MGPC</td>
<td>-0.414</td>
<td>2.554</td>
</tr>
<tr>
<td>MGCA</td>
<td>1.850</td>
<td>2.127</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Bias</th>
<th>S.D.</th>
<th>RMSE</th>
<th>Size</th>
<th>Power</th>
<th>Bias</th>
<th>S.D.</th>
<th>RMSE</th>
<th>Size</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SIV</td>
<td>0.016</td>
<td>1.715</td>
<td>1.715</td>
<td>9.2</td>
<td>99.9</td>
<td>0.480</td>
<td>2.736</td>
<td>2.777</td>
<td>8.7</td>
<td>97.7</td>
</tr>
<tr>
<td>IPC</td>
<td>0.639</td>
<td>2.883</td>
<td>2.953</td>
<td>14.8</td>
<td>98.2</td>
<td>0.939</td>
<td>3.885</td>
<td>3.996</td>
<td>13.2</td>
<td>91.1</td>
</tr>
<tr>
<td>BC-PC</td>
<td>2.547</td>
<td>5.525</td>
<td>6.083</td>
<td>29.5</td>
<td>95.7</td>
<td>2.910</td>
<td>6.102</td>
<td>6.759</td>
<td>24.5</td>
<td>87.7</td>
</tr>
<tr>
<td>PC</td>
<td>-5.703</td>
<td>2.103</td>
<td>6.078</td>
<td>82.5</td>
<td>57.8</td>
<td>-5.413</td>
<td>3.011</td>
<td>6.194</td>
<td>42.6</td>
<td>33.2</td>
</tr>
<tr>
<td>CA</td>
<td>5.971</td>
<td>3.267</td>
<td>6.805</td>
<td>64.3</td>
<td>100.0</td>
<td>6.277</td>
<td>4.086</td>
<td>7.489</td>
<td>39.9</td>
<td>99.7</td>
</tr>
<tr>
<td>MGIV</td>
<td>0.038</td>
<td>1.742</td>
<td>1.742</td>
<td>6.6</td>
<td>99.9</td>
<td>0.036</td>
<td>2.725</td>
<td>2.725</td>
<td>5.6</td>
<td>94.7</td>
</tr>
<tr>
<td>MGPC</td>
<td>-6.047</td>
<td>2.179</td>
<td>6.427</td>
<td>83.6</td>
<td>48.3</td>
<td>-5.997</td>
<td>3.018</td>
<td>6.713</td>
<td>48.3</td>
<td>26.5</td>
</tr>
<tr>
<td>MGCA</td>
<td>4.705</td>
<td>2.610</td>
<td>5.380</td>
<td>54.6</td>
<td>100.0</td>
<td>4.689</td>
<td>3.416</td>
<td>5.801</td>
<td>32.0</td>
<td>99.5</td>
</tr>
</tbody>
</table>
Table 3: Scaled bias of the estimators of $\beta_1$.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>$N = T$</th>
<th>Homogeneous Slopes ($\sqrt{NT \times Bias}$)</th>
<th>Heterogeneous Slopes ($\sqrt{NT \times Bias}$)</th>
<th>$\pi_u = 3/4$</th>
<th>$\pi_u = 1/4$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>25</td>
<td>50</td>
<td>100</td>
<td>200</td>
<td>25</td>
</tr>
<tr>
<td>2SIV</td>
<td>0.162</td>
<td>0.044</td>
<td>0.015</td>
<td>0.005</td>
<td>0.094</td>
</tr>
<tr>
<td>BC-IPC</td>
<td>-0.142</td>
<td>-1.228</td>
<td>-0.771</td>
<td>-0.288</td>
<td>0.003</td>
</tr>
<tr>
<td>IPC</td>
<td>0.551</td>
<td>0.384</td>
<td>0.116</td>
<td>0.040</td>
<td>0.145</td>
</tr>
<tr>
<td>IPC-PC</td>
<td>0.753</td>
<td>0.771</td>
<td>0.604</td>
<td>0.612</td>
<td>0.195</td>
</tr>
<tr>
<td>PC</td>
<td>-1.061</td>
<td>-1.390</td>
<td>-1.317</td>
<td>-1.277</td>
<td>-0.174</td>
</tr>
<tr>
<td>CA</td>
<td>1.509</td>
<td>2.353</td>
<td>3.157</td>
<td>3.718</td>
<td>0.356</td>
</tr>
<tr>
<td>MGIIV</td>
<td>0.258</td>
<td>0.072</td>
<td>0.025</td>
<td>-0.001</td>
<td>0.058</td>
</tr>
<tr>
<td>MGPC</td>
<td>-1.229</td>
<td>-1.463</td>
<td>-1.351</td>
<td>-1.301</td>
<td>-0.240</td>
</tr>
<tr>
<td>MGCA</td>
<td>1.228</td>
<td>1.891</td>
<td>2.604</td>
<td>3.245</td>
<td>0.256</td>
</tr>
</tbody>
</table>

Table 4: Climate shocks and economic growth

<table>
<thead>
<tr>
<th></th>
<th>2SIV (A)</th>
<th>FE (B)</th>
<th>CA (A)</th>
<th>IPC (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}_1$</td>
<td>-0.427*</td>
<td>.530*</td>
<td>-3.28</td>
<td>.412*</td>
</tr>
<tr>
<td></td>
<td>(.241)</td>
<td>(.315)</td>
<td>(.285)</td>
<td>(.230)</td>
</tr>
<tr>
<td>$\hat{\beta}_2$</td>
<td>-1.946***</td>
<td>-1.610***</td>
<td>-1.764***</td>
<td>-1.368***</td>
</tr>
<tr>
<td></td>
<td>(.534)</td>
<td>(.485)</td>
<td>(.528)</td>
<td>(.306)</td>
</tr>
<tr>
<td>$\hat{\gamma}_1$</td>
<td>-0.089***</td>
<td>-0.079*</td>
<td>-0.008</td>
<td>-0.083</td>
</tr>
<tr>
<td></td>
<td>(.041)</td>
<td>(.046)</td>
<td>(.044)</td>
<td>(.034)</td>
</tr>
<tr>
<td>$\hat{\gamma}_2$</td>
<td>-0.016</td>
<td>.153*</td>
<td>.068</td>
<td>.064</td>
</tr>
<tr>
<td></td>
<td>(.088)</td>
<td>(.078)</td>
<td>(.102)</td>
<td>(.069)</td>
</tr>
<tr>
<td>$\hat{\beta}_1 + \hat{\gamma}_1$</td>
<td>-1.417***</td>
<td>-1.348***</td>
<td>-1.570***</td>
<td>-0.976***</td>
</tr>
<tr>
<td></td>
<td>(.429)</td>
<td>(.408)</td>
<td>(.474)</td>
<td>(.272)</td>
</tr>
<tr>
<td>$\hat{\beta}_2 + \hat{\gamma}_2$</td>
<td>-0.033</td>
<td>.070*</td>
<td>.015</td>
<td>.012</td>
</tr>
<tr>
<td></td>
<td>(.056)</td>
<td>(.042)</td>
<td>(.060)</td>
<td>(.046)</td>
</tr>
</tbody>
</table>

Notes: Standard errors in parentheses and p-values in square brackets. * p<0.10, ** p<0.05, *** p<0.01.

In column (B), the effect on developing economies is obtained as $\hat{\beta}_1 + \hat{\gamma}_1$ (or $\hat{\beta}_2 + \hat{\gamma}_2$).
Appendices: Proofs of the main theoretical results

In Appendices A-D, proofs of main theoretical results with necessary Lemmas are provided. Proofs of used lemmas are available in Online Supplement.

Appendix A  Lemmas and proof of Proposition 3.1

Throughout the appendix, we use C to denote a generic finite constant large enough, which need not to be the same at each appearance. Denote the projection matrix \( P_A = A(A'A)^{-1}A' \) and the residual maker \( M_A = I - P_A \) for a matrix \( A \). Let \( \Xi \) be \( r_1 \times r_1 \) diagonal matrix that consist of the first \( r_1 \) largest eigenvalues of the \( T \times T \) matrix \( (NT)^{-1} \sum_{i=1}^N X_iX_i' \). Then by the definition of eigenvalues and \( \hat{F} \), \( \Xi = (NT)^{-1} \sum_{i=1}^N X_iX_i' \). It’s easy to show that \( \Xi \) is invertible following the proof of Lemma A.3 in Bai (2003). Then

\[
\hat{F} - \tilde{F} = \frac{1}{NT} \sum_{\ell=1}^N F\Gamma_\ell^0 \tilde{V}_\ell \Xi^{-1} + \frac{1}{NT} \sum_{\ell=1}^N \tilde{V}_\ell \Gamma_\ell^0 F0 \Xi^{-1} + \frac{1}{NT} \sum_{\ell=1}^N \tilde{V}_\ell \tilde{V}_\ell' \Xi^{-1} \tag{A.1}
\]

where \( \mathbf{R} = (NT)^{-1} \sum_{\ell=1}^N \Gamma_\ell^0 F0 \Xi^{-1} \). Following the proof of Lemma A.3 in Bai (2003) again, we can show that \( \mathbf{R} \) is invertible.

**Lemma A.1** Under Assumptions B to D, we have

(a) \( T^{-1}||\hat{F} - \tilde{F}|| = O_p(\delta_{NT}^{-1}) \),

(b) \( T^{-1}(\hat{F} - \tilde{F}^0 R)' \tilde{F}^0 = O_p(\delta_{NT}^{-2}) \),

(c) \( \Xi = O_p(1), \mathbf{R} = O_p(1), \Xi^{-1} = O_p(1), \mathbf{R}^{-1} = O_p(1) \),

(d) \( \mathbf{M}_F - \tilde{M}_F = O_p(\delta_{NT}^{-1}) \),

(e) \( N^{-1}T^{-1} \sum_{\ell=1}^N \Gamma_\ell^0 \tilde{V}_\ell (\hat{F} - \tilde{F}^0 R) = O_p(N^{-1}) + O_p(N^{-1/2} \delta_{NT}^{-2}) \),

**Lemma A.2** Under Assumptions A to D, we have

(a) \( N^{-1} \sum_{i=1}^N \| \Phi_i^0 \| \| \tilde{\epsilon}_i (\tilde{F}^0 - \tilde{F} \tilde{R}^{-1}) \| = O_p(\delta_{NT}^2) \)

(b) \( N^{-1} \sum_{i=1}^N \| \tilde{\epsilon}_i (\tilde{F} - \tilde{F}^0 R) \| = O_p(\delta_{NT}^{-2}) \)

(c) \( N^{-1} \sum_{i=1}^N \| \tilde{\epsilon}_i (\tilde{F} - \tilde{F}^0 R) \| \| \tilde{V}_i (\tilde{F} - \tilde{F}^0 R) \| = O_p(\delta_{NT}^{-2}) \)

**Lemma A.3** Under Assumptions A to D, we have

\[
N^{-1/2}T^{-1/2} \sum_{\ell=1}^N \Gamma_\ell^0 \tilde{F}^0 \tilde{M}_F \tilde{u}_i = \]

\[
- N^{-3/2}T^{-1/2} \sum_{i=1}^N \sum_{\ell=1}^N \Gamma_\ell^0 (\tilde{Y}_i^0)' \Gamma_\ell^0 \tilde{V}_\ell \tilde{M}_F \tilde{u}_i =
\]

\[
- N^{-3/2}T^{-1/2} \sum_{i=1}^N \sum_{\ell=1}^N \Gamma_\ell^0 (\tilde{Y}_i^0)' (T^{-1} \tilde{F} \tilde{F}^0)'^{-1} \tilde{F}' \tilde{V}_\ell \tilde{M}_F \tilde{u}_i =
\]

\[
+ O_p(N^{-1/2}T^{1/2} \delta_{NT}^{-2}) + O_p(\delta_{NT}^{-2})
\]

24
Lemma A.4  Under Assumptions A to D, we have

\[-N^{-3/2}T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma^{0}_{i}(\mathbf{Y}^{0})^{-1} \Gamma_{\ell}^{0} \mathbf{V}_{i} \mathbf{M}_{p} \mathbf{u}_{i} = -N^{-3/2}T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma^{0}_{i}(\mathbf{Y}^{0})^{-1} \Gamma_{\ell}^{0} \mathbf{V}_{i} \mathbf{M}_{p} \mathbf{u}_{i} + N^{-5/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma^{0}_{i}(\mathbf{Y}^{0})^{-1} \Gamma_{\ell}^{0} \mathbf{V}_{i} \mathbf{V}_{h} \Gamma^{0}_{h}(\mathbf{Y}^{0})^{-1} (T^{-2} \mathbf{F}^{0} \mathbf{F}^{0})^{-1} \mathbf{F}^{0} \mathbf{u}_{i} + O_{p}(T^{1/2} \delta^{-2}_{NT})\]

Lemma A.5  Under Assumptions A to D, we have

\[
N^{-3/2}T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma^{0}_{i}(\mathbf{Y}^{0})^{-1} (T^{-2} \mathbf{F}^{0} \mathbf{F}^{0})^{-1} \mathbf{F}^{0} \mathbf{V}_{i} \mathbf{V}_{\ell} \mathbf{M}_{p} \mathbf{u}_{i} = N^{-3/2}T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma^{0}_{i}(\mathbf{Y}^{0})^{-1} (T^{-2} \mathbf{F}^{0} \mathbf{F}^{0})^{-1} \mathbf{F}^{0} \mathbf{E}(\mathbf{V}_{i} \mathbf{V}_{\ell}) \mathbf{M}_{p} \mathbf{u}_{i} + O_{p}(T^{1/2} \delta^{-2}_{NT}) + O_{p}(N^{1/2}T^{-1/2} \delta^{-1}_{NT})
\]

Lemma A.6  Under Assumptions A to D, we have

\[
N^{-1/2}T^{-1/2} \sum_{i=1}^{N} \mathbf{V}_{i} \mathbf{M}_{p} \mathbf{u}_{i} = N^{-1/2}T^{-1/2} \sum_{i=1}^{N} \mathbf{V}_{i} \mathbf{M}_{p} \mathbf{u}_{i} - N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} \mathbf{V}_{i} \mathbf{V}_{h} \Gamma^{0}_{h}(\mathbf{Y}^{0})^{-1} (T^{-1} \mathbf{F}^{0} \mathbf{F}^{0})^{-1} \mathbf{F}^{0} \mathbf{u}_{i} + O_{p}(N^{1/2}T^{1/2} \delta^{-3}_{NT})
\]

Lemma A.7  Under Assumptions A to D, we have

\[
\frac{1}{NT} \sum_{i=1}^{N} \sum_{h=1}^{N} \mathbf{V}_{i} \mathbf{V}_{h} \Gamma^{0}_{h}(\mathbf{Y}^{0})^{-1} (T^{-2} \mathbf{F}^{0} \mathbf{F}^{0})^{-1} (T^{-1} \mathbf{F}^{0} \mathbf{u}_{i}) = \frac{1}{NT^{2}} \sum_{i=1}^{N} \sum_{h=1}^{N} \mathbf{E}(\mathbf{V}_{i} \mathbf{V}_{h}) \Gamma^{0}_{h}(\mathbf{Y}^{0})^{-1} (T^{-2} \mathbf{F}^{0} \mathbf{F}^{0})^{-1} \mathbf{F}^{0} \mathbf{H} \varphi_{i}^{0} + O_{p}(T^{-1/2})
\]

\[
- \frac{1}{NT^{2}} \sum_{i=1}^{N} \sum_{h=1}^{N} \sum_{\ell=1}^{N} \Gamma^{0}_{h}(\mathbf{Y}^{0})^{-1} \Gamma_{\ell}^{0} \mathbf{E}(\mathbf{V}_{h} \mathbf{V}_{\ell}) \Gamma^{0}_{h}(\mathbf{Y}^{0})^{-1} (T^{-1} \mathbf{F}^{0} \mathbf{F}^{0})^{-1} \mathbf{F}^{0} \mathbf{H} \varphi_{i}^{0} + O_{p}(T^{-1/2})
\]

\[
+ 1 - \frac{1}{NT^{2}} \sum_{i=1}^{N} \sum_{h=1}^{N} \Gamma^{0}_{h}(\mathbf{Y}^{0})^{-1} (T^{-1} \mathbf{F}^{0} \mathbf{F}^{0})^{-1} \mathbf{F}^{0} \mathbf{M}_{p} \mathbf{u}_{i} \Sigma = N^{-1} \sum_{i=1}^{N} \mathbf{E}(\mathbf{V}_{i} \mathbf{V}_{i}) = \frac{1}{NT^{2}} \sum_{i=1}^{N} \sum_{h=1}^{N} \mathbf{E}(\mathbf{V}_{h} \mathbf{V}_{\ell}) \Gamma^{0}_{h}(\mathbf{Y}^{0})^{-1} \Gamma^{0}_{\ell} \mathbf{F}^{0} \mathbf{F}^{0} \mathbf{H} \varphi_{i}^{0} + O_{p}(T^{-1/2})
\]

which are \( O_{p}(1) \), where \( \mathbf{x}_{i} = \mathbf{x}_{i} - N^{-1} \sum_{\ell=1}^{N} \mathbf{x}_{i} \mathbf{F}_{\ell}^{0}(\mathbf{Y}_{\ell}^{0})^{-1} \mathbf{F}_{i}^{0} \), \( \mathbf{v}_{i} = \mathbf{v}_{i} - N^{-1} \sum_{\ell=1}^{N} \mathbf{v}_{i} \mathbf{F}_{\ell}^{0}(\mathbf{Y}_{\ell}^{0})^{-1} \mathbf{F}_{i}^{0} \), \( \mathbf{Y}^{0} = N^{-1} \sum_{i=1}^{N} \mathbf{F}_{i}^{0} \mathbf{F}_{i}^{0} \) and \( \Sigma = N^{-1} \sum_{i=1}^{N} \mathbf{E}(\mathbf{v}_{i} \mathbf{v}_{i}) \).

Proof of Proposition 4.1. With Lemmas A.3, A.4, A.5, we have

\[
N^{-1/2}T^{-1/2} \sum_{i=1}^{N} \Gamma^{0}_{i}(\mathbf{Y}_{i}^{0})^{-1} \Gamma^{0}_{i} \mathbf{M}_{p} \mathbf{u}_{i}
\]

\[
= -N^{-3/2}T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma^{0}_{i}(\mathbf{Y}_{i}^{0})^{-1} \Gamma^{0}_{\ell} \mathbf{V}_{i} \mathbf{M}_{p} \mathbf{u}_{i}
\]

\[
+ N^{-5/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma^{0}_{i}(\mathbf{Y}_{i}^{0})^{-1} \Gamma^{0}_{\ell} \mathbf{V}_{i} \mathbf{V}_{h} \Gamma^{0}_{h}(\mathbf{Y}_{h}^{0})^{-1} (T^{-1} \mathbf{F}^{0} \mathbf{F}^{0})^{-1} \mathbf{F}^{0} \mathbf{u}_{i}
\]

\[
- N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma^{0}_{i}(\mathbf{Y}_{i}^{0})^{-1} (T^{-1} \mathbf{F}^{0} \mathbf{F}^{0})^{-1} \mathbf{F}^{0} \mathbf{E}(\mathbf{V}_{i} \mathbf{V}_{i}) \mathbf{M}_{p} \mathbf{u}_{i} + O_{p}(T^{1/2} \delta^{-2}_{NT}) + O_{p}(N^{1/2}T^{1/2} \delta^{-1}_{NT})\]
and with Lemma A.6,

\[
N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i^T M_{\hat{p}} u_i
\]

\[
= N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i^T M_{\hat{p}} u_i - N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i^T V_h \Gamma_h^0 (Y^0)^{-1} (T^{-1} F^0 F^0)^{-1} F^0 u_i + O_p(N^{1/2}T^{1/2} \delta^{-3}_{NT}).
\]

Then, we have

\[
N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i^T M_{\hat{p}} u_i
\]

\[
= N^{-1/2}T^{-1/2} \sum_{i=1}^{N} X_i^T M_{p0} u_i - N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i^T V_h \Gamma_h^0 (Y^0)^{-1} (T^{-1} F^0 F^0)^{-1} F^0 u_i
\]

\[
- N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} \Gamma_i^0 (Y^0)^{-1} (T^{-1} F^0 F^0)^{-1} F^0 E (V_i V_i') M_{p0} u_i + O_p(N^{1/2}T^{1/2} \delta^{-3}_{NT})
\]

\[
= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T M_{p0} u_i + \sqrt{\frac{T}{N}} a_1 + \sqrt{\frac{N}{T}} a_2 + O_p(N^{1/2}T^{1/2} \delta^{-3}_{NT})
\]

where

\[
a_1 = - \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i^T V_h \Gamma_h^0 (Y^0)^{-1} (T^{-1} F^0 F^0)^{-1} F^0 u_i
\]

\[
a_2 = - \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \Gamma_i^0 (Y^0)^{-1} (T^{-1} F^0 F^0)^{-1} F^0 \Sigma M_{p0} u_i
\]

with \( V_i = V_i - N^{-1} \sum_{\ell=1}^{N} V_i \Gamma_i^\ell (Y^0)^{-1} F_i^\ell \), \( Y^0 = N^{-1} \sum_{\ell=1}^{N} \Gamma_i^\ell \) and \( \Sigma = N^{-1} \sum_{\ell=1}^{N} E(V_i V_i') \). Applying Lemma A.7 to \( a_1 \) and \( a_2 \), we can derive that \( N^{-1/2}T^{-1/2} \sum_{i=1}^{N} X_i^T M_{\hat{p}} u_i = O_p(N^{-1/2}T^{1/2}) + O_p(N^{1/2}T^{-1/2} \delta^{-3}_{NT}) \) and

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i^T M_{\hat{p}} u_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T M_{p0} u_i + b_1 + b_2 + O_p(N^{1/2}T^{1/2} \delta^{-3}_{NT})
\]

where

\[
b_1 = - \sqrt{\frac{T}{N}} \frac{1}{\sqrt{NT^2}} \sum_{i=1}^{N} \sum_{h=1}^{N} E(V_i V_h) \Gamma_h^0 (Y^0)^{-1} (T^{-1} F^0 F^0)^{-1} F^0 \varphi_i^0
\]

\[
+ \sqrt{\frac{T}{N}} \frac{1}{\sqrt{NT^2}} \sum_{i=1}^{N} \sum_{h=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^\ell E(V_i V_h) \Gamma_h^0 (Y^0)^{-1} (T^{-1} F^0 F^0)^{-1} F^0 \varphi_i^0
\]

\[
b_2 = - \sqrt{\frac{N}{T}} \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \Gamma_i^0 (Y^0)^{-1} (T^{-1} F^0 F^0)^{-1} F^0 \Sigma M_{p0} H^0 \varphi_i^0
\]

as required. With the facts that \( N^{-1} \sum_{i=1}^{N} \|T^{-1/2} X_i\| \geq O_p(1) \), \( \|M_{\hat{p}} - M_{p0}\| = O_p(\delta^{-1}_{NT}) \) and \( N^{-1} \sum_{i=1}^{N} V_i^T M_{p0} V_i = N^{-1} \sum_{i=1}^{N} V_i V_i = O_p(T^{-1}) \) we have

\[
N^{-1}T^{-1} \sum_{i=1}^{N} X_i^T M_{\hat{p}} X_i = N^{-1}T^{-1} \sum_{i=1}^{N} V_i^T V_i = O_p(\delta^{-1}_{NT}) + O_p(T^{-1})
\]

so that, with continuous mapping theorem, \( \sqrt{NT}(\hat{\beta}_{1STV} - \beta) = O_p(1) \). This completes the proof. \( \square \)
Appendix B   Lemmas and proofs of Proposition 3.2 and Theorem 3.1

Let $\Xi$ be $r_2 \times r_2$ diagonal matrix that consist of the first $r_2$ largest eigenvalues of the $T \times T$ matrix $N^{-1}T^{-1} \sum_{i=1}^{N} \tilde{u}_i \tilde{u}_i'$. Then by the definition of eigenvalues and $\tilde{H}$, $\tilde{H} \Xi = N^{-1}T^{-1} \sum_{i=1}^{N} \tilde{u}_i \tilde{u}_i' \tilde{H}$. It’s easy to show that $\Xi$ is invertible following the proof of Proposition A.1 (i) in Bai (2009) given $\tilde{\beta}_{ISIV} - \beta = O_p(N^{-1/2}T^{-1/2})$.

$$\tilde{H} - H^0 \mathcal{R}$$

$$= N^{-1}T^{-1} \sum_{i=1}^{N} X_i (\beta - \tilde{\beta}_{ISIV})(\beta - \tilde{\beta}_{ISIV})' X_i \tilde{H} \Xi^{-1}$$

$$+ N^{-1}T^{-1} \sum_{i=1}^{N} X_i (\beta - \tilde{\beta}_{ISIV}) u_i' \tilde{H} \Xi^{-1} + N^{-1}T^{-1} \sum_{i=1}^{N} u_i (\beta - \tilde{\beta}_{ISIV})' X_i \tilde{H} \Xi^{-1}$$

$$+ N^{-1}T^{-1} \sum_{i=1}^{N} H^0 \varphi_i^0 \varepsilon_i' \tilde{H} + N^{-1}T^{-1} \sum_{i=1}^{N} \varepsilon_i \varphi_i^0 H^0 \tilde{H} \Xi^{-1} + N^{-1}T^{-1} \sum_{i=1}^{N} \varepsilon_i \varepsilon_i' \tilde{H} \Xi^{-1}$$

(B.1)

where $\mathcal{R} = T^{-1} \sum_{i=1}^{N} X_i \tilde{H} \Xi^{-1} \tilde{H}$. Following the proof of Proposition A.1 (ii) in Bai (2009), we can show that $\mathcal{R}$ is invertible.

**Lemma B.1** Under Assumptions A to D, we have

(a) $T^{-1} \| \tilde{H} - H^0 \mathcal{R} \|^2 = O_p(\delta_{NT}^2)$,

(b) $T^{-1} (\tilde{H} - H^0 \mathcal{R})' H^0 = O_p(\delta_{NT}^2), T^{-1} (\tilde{H} - H^0 \mathcal{R})' \tilde{H} = O_p(\delta_{NT}^2)$,

(c) $\Xi = O_p(1), \mathcal{R} = O_p(1), \Xi^{-1} = O_p(1), \mathcal{R}^{-1} = O_p(1)$.

(d) $\mathcal{R} \mathcal{R}' - (T^{-1} H^0 H^0)^{-1} = O_p(\delta_{NT}^2)$,

(e) $M_{\tilde{H}} - M_{H^0} = O_p(\delta_{NT}^2)$.

(f) $N^{-1}T^{-1} \sum_{i=1}^{N} \varphi_i \varphi_i' (\tilde{H} - H^0 \mathcal{R}) = O_p(N^{-1}) + O_p(N^{-1/2} \delta_{NT}^{-2})$.

**Lemma B.2** Under Assumptions A to D, we have

(a) $N^{-1} \sum_{i=1}^{N} \| T^{-1/2} V_i \| || T^{-1} (\tilde{H} - H^0 \mathcal{R})' \varepsilon_i || = O_p(\delta_{NT}^2)$,

(b) $N^{-1} \sum_{i=1}^{N} \| \varphi_i^0 \| || T^{-1} V_i (\tilde{H} - H^0 \mathcal{R}) || = O_p(\delta_{NT}^2)$,

(c) $N^{-1} \sum_{i=1}^{N} \| T^{-1} V_i' (\tilde{H} - H^0 \mathcal{R}) || || T^{-1} (\tilde{H} - H^0 \mathcal{R})' \varepsilon_i || = O_p(\delta_{NT}^4)$.

**Lemma B.3** Under Assumptions A to D, we have

$$\frac{1}{\sqrt{N_T}} \sum_{i=1}^{N} V_i \varphi_h M_{\tilde{H}} u_i = \frac{1}{\sqrt{N_T}} \sum_{i=1}^{N} V_i' M_{\varphi_h} M_{\tilde{H}} \varepsilon_i$$

$$- N^{-3/2} T^{-1/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i' \varepsilon_i \varphi_h (Y_{\varphi}^{-1})_0 \varphi_i \varphi_h' + O_p(\delta_{NT}^{-1})$$

27
Lemma B.4 Under Assumptions A to D, we have

\[ N^{-1/2}T^{-1/2} \sum_{i=1}^{N} \Gamma_{i}^{0} \mathbf{F}^{0} \mathbf{M}_{p} \mathbf{M}_{H} \mathbf{u}_{i} = -N^{-3/2}T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \Gamma_{\ell}^{0} \mathbf{V}' \mathbf{M}_{p} \mathbf{M}_{H} \mathbf{u}_{i} \]

\[ -N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \left( T^{-1} \mathbf{F}^{0} \right)^{-1} \mathbf{F}' \mathbf{V}' \mathbf{V}' \mathbf{M}_{p} \mathbf{M}_{H} \mathbf{u}_{i} + O_{p}(\delta^{-2}_{NT}) + O_{p}(N^{-1/2}T^{1/2} \delta^{-2}_{NT}) \]

Lemma B.5 Under Assumptions A to D, we have

\[ N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \left( T^{-1} \mathbf{F}^{0} \right)^{-1} \mathbf{F}' \mathbf{V}' \mathbf{V}' \mathbf{M}_{p} \mathbf{M}_{H} \mathbf{u}_{i} = O_{p}(T^{1/2} \delta^{-2}_{NT}) + O_{p}(N^{-1/2}T^{-1/2} \delta^{-1}_{NT}) \]

Lemma B.6 Under Assumptions A to D, we have

\[ -\frac{1}{N^{3/2}T^{1/2}} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \Gamma_{\ell}^{0} \mathbf{V}' \mathbf{M}_{p} \mathbf{M}_{H} \mathbf{u}_{i} = -\frac{1}{N^{3/2}T^{1/2}} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \Gamma_{\ell}^{0} \mathbf{V}' \mathbf{M}_{p} \mathbf{M}_{H} \mathbf{\varepsilon}_{i}^{0} \]

\[ -N^{-5/2}T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \Gamma_{\ell}^{0} \mathbf{V}' \mathbf{\varepsilon}_{h} \varphi_{\ell}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \varphi_{i}^{0} \]

\[ -N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \Gamma_{\ell}^{0} \mathbf{V}' \Sigma_{\varepsilon} \mathbf{H} \left( T^{-1} \mathbf{H}^{0} \mathbf{H}^{0} \right)^{-1} \left( \mathbf{Y}^{0} \right)^{-1} \varphi_{i}^{0} \]

\[ +O_{p}(T^{1/2} \delta^{-2}_{NT}) \]

Proof of Proposition 3.2. By Lemmas B.3, B.4, B.5 and B.6 and the fact that \( \mathbf{M}_{H}\mathbf{X}_{i} = \mathbf{M}_{H}\mathbf{V}_{i} \), we can derive that

\[ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{X}'_{i} \mathbf{M}_{p} \mathbf{M}_{H} \mathbf{u}_{i} = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}'_{i} \mathbf{\varepsilon}_{i} + b_{0FH} + b_{1FH} + b_{2FH} + o_{p}(1) \]

with

\[ b_{0FH} = -\frac{1}{N^{1/2}T^{3/2}} \sum_{i=1}^{N} \sum_{j=1}^{N} \left( \Gamma_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \Gamma_{j}^{0} + \varphi_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \varphi_{j}^{0} \right) \mathbf{V}'_{i} \mathbf{\varepsilon}_{i} \]

\[ b_{1FH} = -\frac{1}{N^{1/2}N^{2T^{1/2}}} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \left( \Gamma_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \Gamma_{\ell}^{0} \left( \mathbf{V}'_{i} \mathbf{\varepsilon}_{h} \right) \varphi_{\ell}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \varphi_{i}^{0} \right) \]

\[ b_{2FH} = -\frac{1}{T^{1/2}N^{3/2}} \sum_{i=1}^{N} \sum_{j=1}^{N} \Gamma_{i}^{0} \left( \mathbf{Y}^{0} \right)^{-1} \mathbf{V}'_{i} \Sigma_{\varepsilon} \mathbf{H} \left( \mathbf{H}^{0} \mathbf{H}^{0} \right)^{-1} \left( \mathbf{Y}^{0} \right)^{-1} \varphi_{i}^{0} \]

\[ \Sigma_{\varepsilon} = \frac{1}{T} \sum_{j=1}^{T} \mathbb{E} \left( \mathbf{\varepsilon}_{j} \mathbf{\varepsilon}_{j}' \right) \], where \( b_{0FH} = O_{p} \left( N^{-1/2} \right) \), \( b_{1FH} = O_{p} \left( N^{-1/2} \right) \) and \( b_{2FH} = O_{p} \left( T^{-1/2} \right) \). Hence, we have

\[ \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{X}'_{i} \mathbf{M}_{p} \mathbf{M}_{H} \mathbf{u}_{i} = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \mathbf{V}'_{i} \mathbf{\varepsilon}_{i} + o_{p}(1) . \]

In addition, it is easily shown that

\[ \frac{1}{NT} \sum_{i=1}^{N} \mathbf{X}'_{i} \mathbf{M}_{p} \mathbf{M}_{H} \mathbf{X}_{i} - \frac{1}{NT} \sum_{i=1}^{N} \mathbf{V}'_{i} \mathbf{V}_{i} = O_{p}(\delta^{-1}_{NT}) \]

This completes the proof. \( \square \)
Proof of Theorem 3.1. By Proposition 3.2 we have
\[ \sqrt{N} \left( \hat{\beta}_{2SIV} - \beta \right) = (N^{-1/2}T^{-1} \sum_{i=1}^{N} V_i' V_i) - N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i' \varepsilon_i + O_p(\sqrt{\delta_N T}) \].

Under Assumptions A-E and G-H, we have \( N^{-1}T^{-1} \sum_{i=1}^{N} V_i' V_i \xrightarrow{P} A \) and \( N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i' \varepsilon_i \xrightarrow{d} N(0, B) \) by Assumption F, together with continuous mapping theorem yield the required result. \( \Box \)

Appendix C  Lemmas and proofs of Proposition 4.1 and Theorem 4.1

Lemma C.1 Under Assumptions A to D, we have
\[ (a) \| T^{-1} \varepsilon_i (F_0^T - \hat{F} R)^{-1} \| = O_p(\delta_N^2 T) \]
\[ (b) \| T^{-1} V_i' (F - F_0^T) \| = O_p(\delta_N^2 T) \]

Lemma C.2 Under Assumptions A-E and G-H, we have
\[ T^{-1/2} X' M_{p} u_i = T^{-1/2} X' M_{p_0} u_i + O_p(T^{1/2} \delta_N^2 T) . \]

Proof of Proposition 4.1. It’s easy to show that
\[ \| T^{-1} X' M_{p} X_i - T^{-1} X' M_{p_0} X_i \| \leq \| T^{-1/2} X' \|^2 \| M_{p} - M_{p_0} \| = O_p(\delta_N^2) \]
with Lemma A.1(f), we have \( \hat{\beta}_p - \beta_i = O_p(T^{-1/2}) + O_p(\delta_N^2) \), we can derive that
\[ \sqrt{T} (\hat{\beta}_p - \beta_i) = (T^{-1} X' M_{p} X_i)^{-1} \times T^{-1/2} X' M_{p} u_i \]
\[ = (T^{-1} X' M_{p} X_i)^{-1} \times T^{-1/2} X' M_{p_0} u_i + O_p \left( \frac{1}{\sqrt{T}} \delta_N^2 \right) \]
\[ = (T^{-1} X' M_{p_0} X_i)^{-1} \times T^{-1/2} X' M_{p_0} u_i + O_p \left( \frac{1}{\sqrt{T}} \delta_N^2 \right) + O_p \left( \frac{1}{\sqrt{T}} \delta_N^2 \right) \]
which implies that Theorem. As \( \hat{u}_i = M_{p_0} u_i \), \( T^{-1/2} X' M_{p_0} u_i = T^{-1/2} V_i' \hat{u}_i = T^{-1/2} \sum_{t=1}^{T} v_{it} \hat{u}_{it} \), and the term
\[ T^{-1/2} X' M_{p_0} u_i \xrightarrow{d} N(0, \Omega) \]
where \( \Omega = T^{-1} \lim_{T \to \infty} \sum_{i=1}^{T} \sum_{t=1}^{T} \hat{u}_{is} \hat{u}_{it} \mathbb{E}(v_{is} v_{it}) \). This completes the proof. \( \Box \)

Lemma C.3 Under Assumptions A-E and G-H, we have
\[ (a) \sup_{1 \leq i \leq N} \| \Gamma_0' \| = O_p(\sqrt{N}) \]
\[ (b) \sup_{1 \leq i \leq N} \sum_{t=1}^{N} \| V_i' \Gamma_0' \| = \sup_{1 \leq i \leq N} \sum_{t=1}^{N} \sum_{t=1}^{T} V_i' \varepsilon_i V_i' \Gamma_0' = O_p(\sqrt{N}) \]
\[ (c) \sup_{1 \leq i \leq N} \sum_{t=1}^{N} \| V_i' \Gamma_0' \| = O_p(\sqrt{N}) \]
\[ (d) \sup_{1 \leq i \leq N} \sum_{t=1}^{N} \| V_i' \Gamma_0' \| = O_p(\sqrt{N}) \]

Lemma C.4 Under Assumptions A-E and G-H, we have
\[ (a) \sup_{1 \leq i \leq N} \| T^{-1} \varepsilon_i (F_0^T - \hat{F} R)^{-1} \| = O_p(\sqrt{N}) + O_p(\sqrt{N} T^{-1}) + O_p(\sqrt{N} T^{-1/2}) \]
\[ (b) \sup_{1 \leq i \leq N} \| T^{-1} V_i' (F - F_0^T) \| = O_p(\sqrt{N}) \]

Lemma C.5 Under Assumptions A-E and G-H, we have
\[ (a) \sum_{i=1}^{N} \| X' M_{p} u_i - X' M_{p_0} u_i \| = O_p(\sqrt{N}) \]
\[ (b) \sup_{1 \leq i \leq N} \| T^{-1} X' M_{p} X_i - T^{-1} X' M_{p_0} X_i \| = O_p(\sqrt{N}) \]
Proof of Theorem 4.1. Under Assumptions A-E and G-H, we have

\[
\sqrt{N} (\hat{\beta}_{MIV} - \beta) = N^{-1/2} \sum_{i=1}^{N} (\hat{\beta}_i - \beta) = N^{-1/2} \sum_{i=1}^{N} (\hat{\beta}_i - \beta_i) + N^{-1/2} \sum_{i=1}^{N} \epsilon_i
\]

where

\[
N^{-1/2} \sum_{i=1}^{N} (\hat{\beta}_i - \beta_i) = N^{-1/2} \sum_{i=1}^{N} (X_i^T \beta + X_i^T \epsilon_i)^{-1} X_i^T \beta_i u_i
\]

\[
= N^{-1/2} \sum_{i=1}^{N} (X_i^T \beta_i + X_i^T \epsilon_i)^{-1} X_i^T \beta_i u_i + N^{-1/2} \sum_{i=1}^{N} \left[ (X_i^T \beta_i + X_i^T \epsilon_i)^{-1} X_i^T \beta_i u_i - (X_i^T \beta_i + X_i^T \epsilon_i)^{-1} X_i^T \beta_i u_i \right]
\]

\[
= N^{-1/2} \sum_{i=1}^{N} (X_i^T \beta_i + X_i^T \epsilon_i)^{-1} X_i^T \beta_i u_i + N^{-1/2} \sum_{i=1}^{N} \left[ (X_i^T \beta_i + X_i^T \epsilon_i)^{-1} - (X_i^T \beta_i + X_i^T \epsilon_i)^{-1} \right] X_i^T \beta_i u_i
\]

\[
+ N^{-1/2} \sum_{i=1}^{N} (X_i^T \beta_i + X_i^T \epsilon_i)^{-1} (X_i^T \beta_i - X_i^T \beta_i) + N^{-1/2} \sum_{i=1}^{N} (X_i^T \beta_i + X_i^T \epsilon_i)^{-1} (X_i^T \beta_i - X_i^T \beta_i)
\]

\[
= D_1 + D_2 + D_3 + D_4
\]

We first consider the terms \(D_2, D_3,\) and \(D_4.\) Since

\[
T^{-1} X_i^T \beta_i + X_i^T \epsilon_i = T^{-1} V_i^T \beta_i + T^{-1} V_i^T \epsilon_i - T^{-1} V_i^T (F^0_0 F^0_0)^{-1} F^0_0 V_i
\]

\[
= T^{-1} E(V_i V_i) + (T^{-1} V_i^T V_i - T^{-1} E(V_i V_i)) - T^{-1} V_i^T (F^0_0 F^0_0)^{-1} F^0_0 V_i
\]

we have

\[
\sup_{1 \leq i \leq N} \|T^{-1} X_i^T \beta_i + X_i^T \epsilon_i - T^{-1} E(V_i V_i)\|
\]

\[
= T^{-1/2} \sup_{1 \leq i \leq N} \|T^{-1/2} (V_i^T V_i - E(V_i V_i))\| + \sup_{1 \leq i \leq N} \|T^{-1/2} V_i^T (F^0_0 F^0_0)^{-1} F^0_0 V_i\|
\]

\[
= O_p(N^{1/4} T^{-1/2}) + O_p(N^{1/2} T^{-1}).
\]

Furthermore, since

\[
(T^{-1} X_i^T \beta_i + X_i^T \epsilon_i)^{-1} - [T^{-1} E(V_i V_i)]^{-1}
\]

\[
= (T^{-1} X_i^T \beta_i + X_i^T \epsilon_i)^{-1} [T^{-1} E(V_i V_i) - T^{-1} X_i^T \beta_i + X_i^T \epsilon_i][T^{-1} E(V_i V_i)]^{-1}
\]

\[
= [T^{-1} E(V_i V_i)]^{-1} [T^{-1} E(V_i V_i) - T^{-1} X_i^T \beta_i + X_i^T \epsilon_i][T^{-1} E(V_i V_i)]^{-1}
\]

\[
+ [T^{-1} X_i^T \beta_i + X_i^T \epsilon_i]^{-1} [T^{-1} E(V_i V_i) - T^{-1} X_i^T \beta_i + X_i^T \epsilon_i][T^{-1} E(V_i V_i)]^{-1}
\]

we have

\[
\sup_{1 \leq i \leq N} \|(T^{-1} X_i^T \beta_i + X_i^T \epsilon_i)^{-1} - [T^{-1} E(V_i V_i)]^{-1}\|
\]

\[
\leq \sup_{1 \leq i \leq N} \|[T^{-1} E(V_i V_i)]^{-1}\|^2 \sup_{1 \leq i \leq N} \|T^{-1} E(V_i V_i) - T^{-1} X_i^T \beta_i + X_i^T \epsilon_i\|
\]

\[
+ \sup_{1 \leq i \leq N} \|[T^{-1} X_i^T \beta_i + X_i^T \epsilon_i]^{-1} - [T^{-1} E(V_i V_i)]^{-1}\| \sup_{1 \leq i \leq N} \|T^{-1} E(V_i V_i) - T^{-1} X_i^T \beta_i + X_i^T \epsilon_i\| \sup_{1 \leq i \leq N} \|T^{-1} E(V_i V_i)\|^{-1}
\]

\[
= \sup_{1 \leq i \leq N} \|T^{-1} E(V_i V_i) - T^{-1} X_i^T \beta_i + X_i^T \epsilon_i\| \cdot C^2
\]

\[
+ \sup_{1 \leq i \leq N} \|[T^{-1} X_i^T \beta_i + X_i^T \epsilon_i]^{-1} - [T^{-1} E(V_i V_i)]^{-1}\| \cdot [O_p(N^{1/4} T^{-1/2}) + O_p(N^{1/2} T^{-1})] \cdot C
\]

Since \(\sqrt{N}/T \to 0,\) we can see that the second term is \(\sup_{1 \leq i \leq N} \|[T^{-1} X_i^T \beta_i + X_i^T \epsilon_i]^{-1} - [T^{-1} E(V_i V_i)]^{-1}\|. \cdot O_p(1),\) which means that the first term dominates the second term, thus

\[
\sup_{1 \leq i \leq N} \|(T^{-1} X_i^T \beta_i + X_i^T \epsilon_i)^{-1} - [T^{-1} E(V_i V_i)]^{-1}\| = O_p(N^{1/4} T^{-1/2}) + O_p(N^{1/2} T^{-1})
\]

(C.1)
and
\[
\sup_{1 \leq i \leq N} \| (T^{-1} X_i' M_{p0} X_i) \|^{-1} = O_p(1) \tag{C.2}
\]
as \sup_{1 \leq i \leq N} \| (T^{-1} E(V_i V_i') )^{-1} \| \leq C_{\min}^{-1} \leq C. Similarly, by Lemma C.5 (b), we can show that
\[
\sup_{1 \leq i \leq N} \| (T^{-1} X_i' M_{p0} X_i) \|^{-1} - (T^{-1} X_i' M_{p0} X_i) \|^{-1} \| = O_p(N^{-1/2} \delta_{NT}^{-2}) \tag{C.3}
\]
With the above facts, we have
\[
\| D_2 \| \leq N^{-1/2} \sum_{i=1}^{N} \| T^{-1} X_i' M_{p0} u_i \| \cdot \sup_{1 \leq i \leq N} \| (T^{-1} X_i' M_{p0} X_i) \|^{-1} - (T^{-1} X_i' M_{p0} X_i) \|^{-1} \| = O_p(N T^{-1/2} \delta_{NT}^{-2})
\]
\[
\| D_3 \| \leq N^{-1/2} T^{-1} \sum_{i=1}^{N} \| X_i' M_{p0} u_i - X_i' M_{p0} u_i \| \cdot \sup_{1 \leq i \leq N} \| (T^{-1} X_i' M_{p0} X_i) \|^{-1} - (T^{-1} X_i' M_{p0} X_i) \|^{-1} \| = O_p(N \delta_{NT}^{-4})
\]
and
\[
\| D_4 \| \leq N^{-1/2} T^{-1} \sum_{i=1}^{N} \| X_i' M_{p0} u_i - X_i' M_{p0} u_i \| \cdot \sup_{1 \leq i \leq N} \| (T^{-1} X_i' M_{p0} X_i) \|^{-1} \| = O_p(N^{-1/2} \delta_{NT}^{-2}).
\]
Consider $D_1$. Since
\[
X_i' M_{p0} u_i = V_i' H^0 \varphi_i' - V_i' F^0 (F^0 F^0)'^{-1} F^0 H^0 \varphi_i + V_i' \varepsilon_i - V_i' F^0 (F^0 F^0)'^{-1} F^0 \varepsilon_i
\]
we have
\[
D_1 = N^{-1/2} \sum_{i=1}^{N} (X_i' M_{p0} X_i) \|^{-1} X_i' M_{p0} u_i
\]
\[
= N^{-1/2} \sum_{i=1}^{N} (X_i' M_{p0} X_i) \|^{-1} (E(V_i V_i')^{-1} - X_i' M_{p0} u_i) + N^{-1/2} \sum_{i=1}^{N} (E(V_i V_i')^{-1} - X_i' M_{p0} u_i)
\]
\[
- N^{-1/2} \sum_{i=1}^{N} (E(V_i V_i')^{-1} - X_i' M_{p0} u_i) + N^{-1/2} \sum_{i=1}^{N} (E(V_i V_i')^{-1} - X_i' M_{p0} u_i)
\]
\[
- N^{-1/2} \sum_{i=1}^{N} (E(V_i V_i')^{-1} - X_i' M_{p0} u_i)
\]
With (C.1), we can show that the first term is $O_p(N^{3/4} T^{-1}) + O_p(N T^{-3/2})$. It’s easy to show that the last term is $O_p(N^{1/2} T^{-1})$. For the second term, we have
\[
E \| N^{-1/2} \sum_{i=1}^{N} (E(V_i V_i')^{-1} - X_i' M_{p0} u_i) \|^{-1} \| V_i' H^0 \varphi_i \|^{-1} = E \| N^{-1/2} \sum_{i=1}^{N} (E(V_i V_i')^{-1} - X_i' M_{p0} u_i) \|^{-1} \| V_i' H^0 \varphi_i \|^{-1}
\]
\[
= \text{tr} \left( N^{-1/2} \sum_{i=1}^{N} T T \sum_{j=1}^{T} \sum_{a=1}^{T} \sum_{a=1}^{T} \| E(V_i V_i')^{-1} - (V_i' V_j') \| \| E(h_{i0}^0 \varphi_i' \varphi_j h_{j0}^0) \| \right)
\]
\[
\leq C N^{-1} \sum_{i=1}^{N} T T \sum_{j=1}^{T} \sum_{a=1}^{T} \| E(V_i V_i')^{-1} - (V_i' V_j') \| \| E(h_{i0}^0 \varphi_i' \varphi_j h_{j0}^0) \| \| \| \varphi_i' \| \| \varphi_j' \| \|
\]
\[
\leq C N^{-1} T^{-2} \sum_{i=1}^{N} T T \sum_{j=1}^{T} \sum_{a=1}^{T} \| \varphi_i' \| \| \varphi_j' \| \leq C T^{-1}
\]
by Lemma B.2. then the second term is $O_p(T^{-1/2})$. Analogously, the third term can be proved to be $O_p(T^{-1/2})$. Thus, $D_1 = O_p(N^{3/4} T^{-1}) + O_p(N T^{-3/2}) + O_p(T^{-1/2})$.
Finally consider \( \tilde{\beta} \). Since \( \tilde{\beta} - \tilde{\beta}_{MIV} = \beta - \beta - \left( \tilde{\beta}_{MIV} - \beta \right) + e_i \), we have

\[
\tilde{\Sigma}_\beta = \frac{1}{N - 1} \sum_{i=1}^{N} \left( \tilde{\beta}_i - \tilde{\beta}_{MIV} \right) \left( \tilde{\beta}_i - \tilde{\beta}_{MIV} \right)'
\]

Consequently, we obtain

\[
\sqrt{N}(\tilde{\beta}_{MIV} - \beta) = N^{-1/2} \sum_{i=1}^{N} e_i + o_p(1)
\]

and by a standard central limit theorem

\[
\sqrt{N}(\tilde{\beta}_{MIV} - \beta) \xrightarrow{d} N(0, \Sigma_\beta).
\]

Finally consider \( \hat{\Sigma}_\beta \). Since \( \hat{\beta} - \tilde{\beta}_{MIV} = \beta - \beta - \left( \hat{\beta}_{MIV} - \beta \right) + e_i \), we have

\[
\hat{\Sigma}_\beta = \frac{1}{N - 1} \sum_{i=1}^{N} \left( \hat{\beta}_i - \hat{\beta}_{MIV} \right) \left( \hat{\beta}_i - \hat{\beta}_{MIV} \right)'
\]

Combining the above terms, we have

\[
N^{-1/2} \sum_{i=1}^{N} (\hat{\beta}_i - \beta_i) = O_p(N^{3/4}T^{-1}) + O_p(NT^{-3/2}) + O_p(N^{1/2}\delta_{NT}^{-2}).
\]

Consequently,

\[
\sqrt{N}(\hat{\beta}_{MIV} - \beta) = N^{-1/2} \sum_{i=1}^{N} e_i + o_p(1)
\]

and by a standard central limit theorem

\[
\sqrt{N}(\hat{\beta}_{MIV} - \beta) \xrightarrow{d} N(0, \Sigma_\beta).
\]

Appendix D Lemmas and Proofs of Propositions 5.1 and 5.2

Lemma D.1 Under Assumptions A-E

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^{N} X_i' \Lambda H_0 H \varphi_i = - \frac{1}{\sqrt{N^{3/2}T^{1/2}}} \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij} X_j' \Lambda H_0 \varepsilon_i
\]

\[
- \frac{1}{\sqrt{N^{1/2}T^{3/2}}} \sum_{i=1}^{N} X_i' \Lambda H_0 \Sigma_\beta H \left( \frac{H' H}{T} \right)^{-1} (Y_0^{-1} \varphi_0
\]
where \( a_{ij} = \varphi_j^0 (Y_0^i) - 1 \varphi_j^0 \) and \( \Sigma_x = \frac{1}{N} \sum_{j=1}^N E (\varepsilon_j \varepsilon'_j) \).

**Lemma D.2** Under Assumptions A-E

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{M}_\mathbf{R} \varepsilon_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{M}_{H^0} \varepsilon_i \\
- \sqrt{\frac{T}{N^2}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{H}^0 \left( \frac{\mathbf{H}^0 \mathbf{H}^0}{T} \right)^{-1} (\mathbf{Y}_0^i)^{-1} \mathbf{\varphi}_i \mathbf{E} (\varepsilon'_j \varepsilon_i / T) \\
+ O_p \left( \sqrt{NT \delta_{N_T}^3} \right)
\]

where \( \mathcal{X}_i = \mathbf{X}_i - N^{-1} \sum_{j=1}^N a_{ij} \mathbf{X}_j \).

**Lemma D.3** Under Assumptions A-E

\[
\frac{1}{N^{1/2} T^{3/2}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{M}_\mathbf{H} \Sigma_i \mathbf{\hat{H}} \left( \frac{\mathbf{H}^0 \mathbf{H}^0}{T} \right)^{-1} (\mathbf{Y}_0^i)^{-1} \mathbf{\varphi}_i \\
= \sqrt{\frac{T}{N^2}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{M}_{H^0} \Sigma_i \mathbf{H}^0 \left( \frac{\mathbf{H}^0 \mathbf{H}^0}{T} \right)^{-1} (\mathbf{Y}_0^i)^{-1} \mathbf{\varphi}_i + O_p \left( \sqrt{NT \delta_{N_T}^3} \right).
\]

**Proof of Proposition 5.1** Under Assumptions A-E, by Lemmas D.1, D.2 and D.3, we have

\[
\frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{M}_\mathbf{R} \mathbf{u}_i = \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{M}_\mathbf{H} \mathbf{H}^0 \mathbf{\varphi}_i + \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{M}_\mathbf{R} \varepsilon_i \\
= \frac{1}{\sqrt{NT}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{M}_{H^0} \varepsilon_i \\
- \sqrt{\frac{T}{N^2}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{H}^0 \left( \frac{\mathbf{H}^0 \mathbf{H}^0}{T} \right)^{-1} (\mathbf{Y}_0^i)^{-1} \mathbf{\varphi}_i \mathbf{E} (\varepsilon'_j \varepsilon_i / T) \\
- \sqrt{\frac{T}{N^2}} \sum_{i=1}^N \mathcal{X}_i^T \mathbf{M}_{H^0} \Sigma_i \mathbf{H}^0 \left( \frac{\mathbf{H}^0 \mathbf{H}^0}{T} \right)^{-1} (\mathbf{Y}_0^i)^{-1} \mathbf{\varphi}_i + O_p \left( \sqrt{NT \delta_{N_T}^3} \right)
\]

as required. \( \square \)

**Proof of Proposition 5.2** Under Assumptions A-E, following Westerlund and Urbain (2015) we have

\[
N^{-1/2} T^{-1/2} \sum_{i=1}^N \mathbf{V}_i^T \mathbf{M}_\mathbf{G}_2 \varepsilon_i = N^{-1/2} T^{-1/2} \sum_{i=1}^N \mathbf{V}_i^T \mathbf{M}_{H^0} \varepsilon_i + O_p (N^{1/2} T^{1/2} \delta_{N_T}^{-3})
\]

\[
N^{-1/2} T^{-1/2} \sum_{i=1}^N \mathbf{V}_i^T \mathbf{M}_\mathbf{G}_2 \mathbf{H}^0 \varphi_i = -N^{-3/2} T^{-1/2} \sum_{i=1}^N \sum_{h=1}^N \mathbf{V}_i^T \mathbf{U}_h \mathbf{A}_h^0 (\mathbf{Y}_\Lambda^0)^{-1} \varphi_i + O_p (N^{1/2} T^{1/2} \delta_{N_T}^{-3})
\]

\[
N^{-1/2} T^{-1/2} \sum_{i=1}^N \mathbf{I}_i^0 \mathbf{F}_0^0 \mathbf{M}_\mathbf{G}_2 \varepsilon_i = -N^{-3/2} T^{-1/2} \sum_{i=1}^N \sum_{j=1}^N \mathbf{I}_i^0 \mathbf{U}_j \mathbf{A}_h^0 (\mathbf{Y}_\Lambda^0)^{-1} \mathbf{A}_h \mathbf{E} (\mathbf{U}_j \mathbf{u}_i) + O_p (N^{1/2} T^{1/2} \delta_{N_T}^{-3})
\]

\[
N^{1/2} T^{-1/2} \sum_{i=1}^N \mathbf{I}_i^0 \mathbf{F}_0^0 \mathbf{M}_\mathbf{G}_2 \mathbf{H}^0 \varphi_i = N^{-5/2} T^{-1/2} \sum_{i=1}^N \sum_{j=1}^N \sum_{h=1}^N \mathbf{I}_i^0 \mathbf{U}_j \mathbf{A}_h^0 (\mathbf{Y}_\Lambda^0)^{-1} \mathbf{A}_h \mathbf{E} (\mathbf{U}_j \mathbf{u}_i) \mathbf{A}_h (\mathbf{Y}_\Lambda^0)^{-1} \varphi_i \\
+ O_p (N^{1/2} T^{1/2} \delta_{N_T}^{-3})
\]
where $\mathbf{Y}_i^\Lambda = N^{-1} \sum_{t=1}^{N} \mathbf{A}_i^0 \mathbf{A}_i^0$, and

$$
N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \mathbf{V}' \mathbf{M}_2 \varepsilon_i = N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \mathbf{V}' \mathbf{M}_2 \mathbf{H}^0 \varepsilon_i + O_p(N^{1/2} T^{1/2} \delta_{NT}^{-3})
$$

$$
N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \mathbf{V}' \mathbf{M}_2 \mathbf{H}^0 \varphi_i^0 = -N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \mathbf{V}' \mathbf{A}^{-1} \varphi_i^0 + O_p(N^{1/2} T^{1/2} \delta_{NT}^{-3})
$$

$$
N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \mathbf{T}_i^0 \mathbf{F}_i^0 \mathbf{M}_2 \varepsilon_i = -N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \mathbf{T}_i^0 \mathbf{H}^0 \mathbf{F}_i^0 \varphi_i^0 + O_p(N^{1/2} T^{1/2} \delta_{NT}^{-3})
$$

$$
N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \mathbf{T}_i^0 \mathbf{F}_i^0 \mathbf{M}_2 \mathbf{H}^0 \varphi_i^0 = N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \mathbf{T}_i^0 \mathbf{F}_i^0 \mathbf{H}^0 \varphi_i^0 + O_p(N^{1/2} T^{1/2} \delta_{NT}^{-3})
$$

where $\mathbf{U} = N^{-1} \sum_{i=1}^{N} \mathbf{U}_i$ and $\hat{\Lambda}^\Lambda = \Lambda^\Lambda (\Lambda^\Lambda)^{-1}$ with $\Lambda = N^{-1} \sum_{i=1}^{N} \mathbf{A}_i^0$.

References


Online Supplemental Material to
“Two-Stage Instrumental Variable Estimation of Linear Panel Data Models with a Multifactor Error Structure”
by Guowei Cui, Milda Norkutė, Vasilis Sarafidis and Takashi Yamagata

1 Proofs of Lemmas

A Proofs of Lemmas in Appendix A

Proof of Lemma A.1. For the proofs of (a) to (d), and (f), see Proof of Lemma 4 in Supplemental Material, Norkutė et al. (2021). For (e), we decompose the left hand side term as

$$M_F - M_{F^0} = -T^{-1}\hat{F}(\hat{F} - F^0R) - T^{-1}(\hat{F} - F^0R)RF^0 - T^{-1}F^0\left(RR' - (T^{-1}F^0F^0)^{-1}\right)F^0$$

then it will be bounded in norm by

$$\|T^{-1/2}\hat{F}\| \|T^{-1/2}(\hat{F} - F^0R)\| + \|R\| \|T^{-1/2}F^0\| \|T^{-1/2}(\hat{F} - F^0R)\| + \|T^{-1/2}F^0\| \|RR' - (T^{-1}F^0F^0)^{-1}\| = O_p(\delta_{NT}^2)$$

with (a), (c), (d) and the facts that $T^{-1/2}\hat{F} = r_1$ and $E\|T^{-1/2}F^0\|^2 \leq C$ by Assumption C. This completes the proof. □

Proof of Lemma A.2. Consider (a). With the equation (A.1), we have

$$N^{-1}\sum_{i=1}^{N} \|\Gamma_i^0\| \|T^{-1}\varepsilon_i'(F^0 - \hat{F}R)^{-1}\|$$

$$\leq N^{-2}T^{-2}\sum_{i=1}^{N} \|\Gamma_i^0\| \|\sum_{\ell=1}^{N} \varepsilon_i' F_0'\Gamma_{0}\varepsilon_i\| \|\Xi^{-1}R^{-1}\| + N^{-2}T^{-2}\sum_{i=1}^{N} \|\Gamma_i^0\| \|\sum_{\ell=1}^{N} \varepsilon_i' V_i \varepsilon_i' \varepsilon_i V_i' F_0' \| \|\Xi^{-1}R^{-1}\|$$

$$+ N^{-2}T^{-2}\sum_{i=1}^{N} \|\Gamma_i^0\| \|\sum_{\ell=1}^{N} \varepsilon_i' V_i \varepsilon_i' F_0' \| \|\Xi^{-1}R^{-1}\|$$

Since $\Xi^{-1} = O_p(1)$ and $R^{-1} = O_p(1)$ by Lemma A.1 (c), we omit $\|\Xi^{-1}R^{-1}\|$, which is $O_p(1)$, in the following analysis. The first term is bounded in norm by

$$T^{-1/2}\left(N^{-1}\sum_{i=1}^{N} \|\Gamma_i^0\| \|T^{-1/2}\varepsilon_i'(F^0)\| \right) \sum_{\ell=1}^{N} \|\Gamma_{0}\| \|V_i' F_0'\|$$

With Assumptions A and C, we have

$$E\|T^{-1/2}\varepsilon_i'(F^0)\|^2 = T^{-1}\sum_{s=1}^{T} \sum_{t=1}^{T} E(\varepsilon_{is}\varepsilon_{it}) \varepsilon_i'(F^0) \varepsilon_i(F^0) \leq T^{-1}\sum_{s=1}^{T} \sum_{t=1}^{T} \sigma_{st} \sqrt{E\|\varepsilon_i'(F^0)\|^2 E\|\varepsilon_i'(F^0)\|^2} \leq C.$$ 

then $E(N^{-1}\sum_{i=1}^{N} \|\Gamma_i^0\| \|T^{-1/2}\varepsilon_i'(F^0)\|) \leq N^{-1}\sum_{i=1}^{N} \sqrt{E\|\Gamma_i^0\|^2 E\|T^{-1/2}\varepsilon_i'(F^0)\|^2} \leq C$ by Assumption D, which then implies that

$$N^{-1}\sum_{i=1}^{N} \|\Gamma_i^0\| \|T^{-1/2}\varepsilon_i'(F^0)\| = O_p(1) \quad (A.2)$$
By Lemma A.1 (h), we have

\[ N^{-1}T^{-1} \sum_{\ell=1}^{N} \bar{F} V_{\ell} \bar{\Gamma}_{\ell}^{00} = N^{-1}T^{-1} \sum_{\ell=1}^{N} R^{0} V_{\ell} \Gamma_{\ell}^{00} + N^{-1}T^{-1} \sum_{\ell=1}^{N} (\bar{F} - F^{0} R) V_{\ell} \Gamma_{\ell}^{00} \]

\[ = O_p(N^{-1/2}T^{-1/2}) + O_p(N^{-1}). \]

With the above two equations, the first term is \( O_p(N^{-1/2}T^{-1}) + O_p(N^{-1}T^{-1/2}) \). The second term is bounded in norm by

\[ N^{-3/2}T^{-1/2} \sum_{i=1}^{N} \| \bar{\Gamma}_{i}^{0} \| \| N^{-1/2}T^{-1/2} \sum_{i=1}^{N} \epsilon_{i} V_{i} \bar{\Gamma}_{i}^{00} \| \| T^{-1/2}F^{0} \| \| T^{-1/2}\bar{F} \| = O_p(N^{-1/2}T^{-1/2}) \]

where \( N^{-1}\sum_{i=1}^{N} \| \bar{\Gamma}_{i}^{0} \| \| N^{-1/2}T^{-1/2} \sum_{i=1}^{N} \epsilon_{i} V_{i} \bar{\Gamma}_{i}^{00} \| = O_p(1) \) can be proved by following the way in the proof of (C.4). Consider the third term. Easily, we can prove \( \mathbb{E} \| T^{-1/2} \epsilon_{i} \|^{2} \leq C \). By Cauchy-Schwartz inequality, we have

\[ \mathbb{E} \| N^{-1}T^{-1} \sum_{\ell=1}^{N} \epsilon_{i}^{\prime} \mathbb{E}(V_{i} V_{i}^{\prime}) F^{0} \|^{2} = \mathbb{E} \left[ \left( \sum_{s=1}^{T} \sum_{t=1}^{T} (N^{-1} \sum_{\ell=1}^{N} \epsilon_{i}^{\prime} \mathbb{E}(V_{i}^{s} V_{i}^{t} \ell)) \epsilon_{i} f_{i}^{0} \right)^{2} \right] \]

\[ \leq T^{-2} \sum_{s_{1}=1}^{T} \sum_{t_{1}=1}^{T} \sum_{s_{2}=1}^{T} \sum_{t_{2}=1}^{T} \| N^{-1} \sum_{\ell=1}^{N} \mathbb{E}(V_{i}^{s_{1}} V_{i}^{t_{1}} \ell) \| \| N^{-1} \sum_{\ell=1}^{N} \mathbb{E}(V_{i}^{s_{2}} V_{i}^{t_{2}} \ell) \| \mathbb{E} \| \epsilon_{i} \|^{2} \| \epsilon_{i} f_{i}^{0} \|^{2} \]

\[ \leq T^{-2} \sum_{s=1}^{T} \sum_{t=1}^{T} \sum_{s_{1}=1}^{T} \sum_{t_{1}=1}^{T} \sum_{s_{2}=1}^{T} \sum_{t_{2}=1}^{T} \| s_{1} t_{1} s_{2} t_{2} \| \| \epsilon_{i} \|^{2} \| \epsilon_{i} f_{i}^{0} \|^{2} \]

\[ \leq C \cdot (T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \| \epsilon_{i} \|^{2} \leq C^{2}, \]

by Assumptions A, B2, and C. With Assumption B5, we can follow the way of the proof of Lemma A.2(i) in Bai (2009) to show that \( \mathbb{E} \| N^{-1/2}T^{-1} \sum_{i=1}^{N} \epsilon_{i}^{\prime} \left[ V_{i} V_{i}^{\prime} - \mathbb{E} (V_{i} V_{i}^{\prime}) \right] F^{0} \|^{2} \leq C \). Using the similar argument of (C.4), with the above three moment conditions, we obtain

\[ N^{-1} \sum_{i=1}^{N} \| \bar{\Gamma}_{i}^{0} \| \| N^{-1/2} \epsilon_{i} \| = O_p(1) \]

\[ N^{-1} \sum_{i=1}^{N} \| \bar{\Gamma}_{i}^{0} \| \| N^{-1/2} \epsilon_{i} \| E(V_{i} V_{i}^{\prime}) F^{0} \| = O_p(1) \]

\[ N^{-1} \sum_{i=1}^{N} \| \bar{\Gamma}_{i}^{0} \| \| N^{-1/2} \epsilon_{i} \| [V_{i} V_{i}^{\prime} - \mathbb{E} (V_{i} V_{i}^{\prime})] F^{0} \| = O_p(1) \]

Thus, with Lemma A.1 (a), the third term is bounded in norm by

\[ N^{-1} \sum_{i=1}^{N} \| \bar{\Gamma}_{i}^{0} \| \| N^{-1/2} \epsilon_{i} \| \cdot \| N^{-1/2} \sum_{i=1}^{N} V_{i} V_{i}^{\prime} \| \| T^{-1/2} (\bar{F} - F^{0} R) \| \]

\[ + T^{-1} \cdot N^{-1} \sum_{i=1}^{N} \| \bar{\Gamma}_{i}^{0} \| \| N^{-1/2} \sum_{i=1}^{N} \epsilon_{i}^{\prime} E(V_{i} V_{i}^{\prime}) F^{0} \| \| R \| \]

\[ + N^{-1/2} T^{-1} \cdot N^{-1} \sum_{i=1}^{N} \| \bar{\Gamma}_{i}^{0} \| \| N^{-1/2} \sum_{i=1}^{N} \epsilon_{i}^{\prime} [V_{i} V_{i}^{\prime} - \mathbb{E} (V_{i} V_{i}^{\prime})] F^{0} \| \| R \| = O_p(\delta_{NT}^{-2}) \]

because

\[ \| N^{-1/2} T^{-1} \sum_{i=1}^{N} V_{i} V_{i}^{\prime} \| = O_p(\delta_{NT}^{-1}). \]
which suggest from

\[ \| N^{-1} T^{-1/2} \sum_{t=1}^{N} \mathbb{E} \left( V_t V_t' \right) \|^2 = T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \left| N^{-1} \sum_{i=1}^{N} \mathbb{E} \left( v_{is} v_{it} \right) \right|^2 \]

(A.6)

\[ \leq C N^{-1} T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} | \mathbb{E} \left( v_{is} v_{it} \right) | \leq C T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \sigma_{st} \leq C^2 , \]

and

\[ \mathbb{E} \left[ \left\| N^{-1/2} T^{-1} \sum_{t=1}^{N} \left[ V_t V_t' - \mathbb{E} (V_t V_t') \right] \right\|^2 \right] = T^{-2} \sum_{s=1}^{T} \sum_{t=1}^{T} \mathbb{E} \left[ \left| N^{-1/2} \sum_{i=1}^{N} \left( v_{is} v_{it} - \mathbb{E} (v_{is} v_{it}) \right) \right|^2 \right] \leq C , \]

(A.7)

given \( N^{-1} \sum_{i=1}^{N} | \mathbb{E} (v_{is} v_{it}) | \leq N^{-1} \sum_{i=1}^{N} | \mathbb{E} (v_{is} v_{it}) | \leq N^{-1} \sum_{i=1}^{N} \sqrt{\mathbb{E} | v_{is} |^2 \mathbb{E} | v_{it} |^2} \leq C \) and Assumption B. Collecting the above three terms, the claim holds.

Consider (b). Replacing \( \tilde{F} - F^0 R \) by its expression (A.1), we have

\[ N^{-1} \sum_{i=1}^{N} \left\| \phi_0^i \right\| \left\| T^{-1} V_i' (\tilde{F} - F^0 R) \right\| \]

\[ \leq N^{-2} - T^{-2} \sum_{i=1}^{N} \left\| \phi_0^i \right\| \sum_{\ell=1}^{N} V_i' F_0^\ell \Gamma_\ell^0 V_i' \tilde{F} \left\| \Xi^{-1} \right\| + N^{-2} - T^{-2} \sum_{i=1}^{N} \left\| \phi_0^i \right\| \sum_{\ell=1}^{N} V_i' V_\ell \Gamma_\ell^0 F_0^\ell \tilde{F} \left\| \Xi^{-1} \right\| \]

\[ + N^{-2} - T^{-2} \sum_{i=1}^{N} \left\| \phi_0^i \right\| \sum_{\ell=1}^{N} V_i' V_\ell \tilde{F} \left\| \Xi^{-1} \right\| \]

Ignoring \( \left\| \Xi^{-1} \right\| \) and following the arguments of the first term in the proof of (a), the first term is \( O_p(N^{-1/2} T^{-1}) + O_p(N^{-1/2} T^{-1/2}) + O_p(N^{-1/2} T^{-1/2} \delta_{N \ell}^2) \). The second term is bounded in norm by

\[ N^{-2} - T^{-1} \sum_{i=1}^{N} \left\| \phi_0^i \right\| \sum_{\ell=1}^{N} V_i' V_\ell \Gamma_\ell^0 \left\| \right\| T^{-1/2} F^0 \left\| \right\| T^{-1/2} \tilde{F} \]

\[ = N^{-2} - T^{-1} \sum_{i=1}^{N} \left\| \phi_0^i \right\| \sum_{\ell=1}^{N} V_i' V_\ell \Gamma_\ell^0 \left\| \times O_p(1) \right\|

\[ \leq N^{-2} - T^{-1} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \left\| \phi_0^i \right\| \left\| \mathbb{E} \left( V_i' V_\ell \right) \right\| \left\| \Gamma_\ell^0 \right\| + N^{-2} - T^{-1} \sum_{i=1}^{N} \left\| \phi_0^i \right\| \sum_{\ell=1}^{N} \left( V_i' V_\ell - \mathbb{E} \left( V_i' V_\ell \right) \right) \left\| \Gamma_\ell^0 \right\| \]

\[ \leq N^{-2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \left\| \phi_0^i \right\| \left\| \Gamma_\ell^0 \right\| + N^{-2} - T^{-1} \sum_{i=1}^{N} \left\| \phi_0^i \right\| \sum_{\ell=1}^{N} \left( V_i' V_\ell - \mathbb{E} \left( V_i' V_\ell \right) \right) \left\| \Gamma_\ell^0 \right\| \]

\[ = O_p(N^{-1}) + O_p(N^{-1/2} T^{-1/2}) , \]

since

\[ \mathbb{E} \left( \sum_{i=1}^{N} \sum_{\ell=1}^{N} \left\| \phi_0^i \right\| \left\| \Gamma_\ell^0 \right\| \right) \leq N^{-2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \left\| \phi_0^i \right\| \sqrt{\mathbb{E} \left( \mathbb{E} \left( \Gamma_\ell^0 \right) \right)^2} \leq C N^{-2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \left\| \phi_0^i \right\| \left\| \Gamma_\ell^0 \right\| \leq C^2 N^{-1} \]

by Assumption B2 and

\[ N^{-3/2} T^{-1/2} \sum_{i=1}^{N} \left\| \phi_0^i \right\| \sum_{\ell=1}^{N} \left( V_i' V_\ell - \mathbb{E} \left( V_i' V_\ell \right) \right) \left\| \Gamma_\ell^0 \right\| = O_p(1) \]
given \( \mathbb{E}\|N^{-1/2}T^{-1/2} \sum_{\ell=1}^{N} (\mathbf{V}_t \mathbf{V}_t - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t)) \mathbf{F}_0 \|^2 \leq C \) by Assumption B4. With (A.5) and Lemma A.1 (a) and (d), the third term is bounded in norm by

\[
N^{-1} \sum_{i=1}^{N} \| \varphi_i \| \| T^{-1/2} \mathbf{V}_t \| \| N^{-1/2} \sum_{\ell=1}^{N} \mathbf{V}_t \mathbf{V}_t \| \| T^{-1/2} (\mathbf{F} - \mathbf{F}_0) \| \\
+ N^{-1} T^{-1} \sum_{i=1}^{N} \| \varphi_i \| \| N^{-1} \sum_{\ell=1}^{N} \mathbf{V}_t \mathbf{V}_t \mathbb{E} (\mathbf{V}_t \mathbf{V}_t) \mathbf{F}_0 \| \| \mathbf{R} \| \\
+ N^{-3/2} T^{-1/2} \sum_{i=1}^{N} \| \varphi_i \| \| T^{-1/2} \mathbf{V}_t \| \| N^{-1/2} T^{-1} \sum_{\ell=1}^{N} [\mathbf{V}_t \mathbf{V}_t - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t)] \mathbf{F}_0 \| \| \mathbf{R} \| = O_p(\delta_N^{-2})
\]

because \( \mathbb{E}\|N^{-1/2}T^{-1/2} \sum_{\ell=1}^{N} \mathbf{V}_t \mathbf{V}_t \mathbb{E} (\mathbf{V}_t \mathbf{V}_t) \mathbf{F}_0 \| ^2 \leq C \), which can be proved by following the way of the proof of (C.6), and

\[
\mathbb{E}\left[ \frac{1}{\sqrt{N}} \sum_{i=1}^{N} (\mathbf{V}_t \mathbf{V}_t - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t)) \mathbf{F}_0 \| ^2 = T^{-1} \sum_{s=1}^{T} \mathbb{E}\left[ \frac{1}{\sqrt{N}} \sum_{t=1}^{T} \mathbb{E} (\mathbf{V}_s \mathbf{V}_t - \mathbb{E}(\mathbf{V}_s \mathbf{V}_t)) \mathbf{F}_0 \| ^2 \leq C
\]

by Assumption B2. Combining the above three terms, (b) holds.

Consider (c). In the proof of (b), we only require \( \mathbb{E}\| \varphi_i \| ^2 \leq C \) with respect to \( \varphi_i \). Then, we can follow the argument in the proof of (b) to show that

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1} \mathbf{V}_i (\mathbf{F} - \mathbf{F}_0) \| \| T^{-1/2} \mathbf{F}_0 \| = O_p(\delta_N^{-2})
\]

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1} \mathbf{V}_i (\mathbf{F} - \mathbf{F}_0) \| \| N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \mathbf{V}_t \mathbf{V}_t \mathbb{E} (\mathbf{F}_0) \| = O_p(\delta_N^{-2})
\]

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1} \mathbf{V}_i (\mathbf{F} - \mathbf{F}_0) \| \| T^{-1/2} \mathbf{F}_0 \| = O_p(\delta_N^{-2})
\]

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1} \mathbf{V}_i (\mathbf{F} - \mathbf{F}_0) \| \| N^{-1} T^{-1} \sum_{\ell=1}^{N} \mathbf{V}_t \mathbf{V}_t - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t) \mathbf{F}_0 \| = O_p(\delta_N^{-2})
\]

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1} \mathbf{V}_i (\mathbf{F} - \mathbf{F}_0) \| \| N^{-1/2} T^{-1} \sum_{\ell=1}^{N} \mathbf{V}_t \mathbf{V}_t - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t) \| \mathbf{F}_0 \| = O_p(\delta_N^{-2})
\]

With the above equations, the proof of (c) is analogous to that of (a), in which we replace \( \| \mathbf{F}_0 \| \) by \( \| T^{-1} \mathbf{V}_i (\mathbf{F} - \mathbf{F}_0) \| \). This completes the proof. □

**Proof of Lemma A.3.** As \( \mathbf{M}_T \mathbf{F} = 0 \), the term on the left hand is equal to \( N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \mathbf{V}_t (\mathbf{F}_0 - \mathbf{F}_0 \mathbf{F}_0^{-1}) \mathbf{M}_T \mathbf{F} \mathbf{u}_t \). With the equation (A.1), it can be decomposed as

\[
- N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{F}_t R^{-1} \mathbf{V}_t (\mathbf{F}_0 - \mathbf{F}_0 \mathbf{F}_0^{-1}) \mathbf{M}_T \mathbf{F} \mathbf{u}_t = - N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{F}_t R^{-1} \mathbf{V}_t (\mathbf{F}_0 - \mathbf{F}_0 \mathbf{F}_0^{-1}) \mathbf{V}_t \mathbf{M}_T \mathbf{F} \mathbf{u}_t
\]

\[
- N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{F}_t R^{-1} \mathbf{V}_t (\mathbf{F}_0 - \mathbf{F}_0 \mathbf{F}_0^{-1}) \mathbf{V}_t \mathbf{M}_T \mathbf{F} \mathbf{u}_t = - N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{F}_t R^{-1} \mathbf{V}_t (\mathbf{F}_0 - \mathbf{F}_0 \mathbf{F}_0^{-1}) \mathbf{V}_t \mathbf{M}_T \mathbf{F} \mathbf{u}_t
\]

\[
- N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{F}_t R^{-1} \mathbf{V}_t (\mathbf{F}_0 - \mathbf{F}_0 \mathbf{F}_0^{-1}) \mathbf{V}_t \mathbf{M}_T \mathbf{F} \mathbf{u}_t = - N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbf{F}_t R^{-1} \mathbf{V}_t (\mathbf{F}_0 - \mathbf{F}_0 \mathbf{F}_0^{-1}) \mathbf{V}_t \mathbf{M}_T \mathbf{F} \mathbf{u}_t
\]
We consider $A_1$. It’s easy to show that $N^{-1} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \varphi_i^0 \right\| = O_p(1)$, then,

$$N^{-1/2}T^{-1/2} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \left\| (F^0 - \hat{F}R^{-1})' u_i \right\|$$

$$\leq N^{1/2}T^{1/2} \left( N^{-1} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \varphi_i^0 \right\| T^{-1}(F^0 - \hat{F}R^{-1})' H^0 \right\| + N^{-1} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| T^{-1}(F^0 - \hat{F}R^{-1})' \varepsilon_i \right\|$$

$$= O_p(N^{1/2}T^{1/2}\delta_{NT}^{-2})$$

(A.9)

by Lemma A.1(b), A.2(a) and $u_i = H_0^0 \varphi_i^0 + \varepsilon_i$. In addition, we have

$$N^{-1/2}T^{-1/2} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \left\| \hat{F}' u_i \right\|$$

$$\leq \sqrt{NT} \left( N^{-1} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \varphi_i^0 \right\| \left\| T^{-1/2} \hat{F}' \right\| \left\| T^{-1/2} H^0 \right\| + N^{-1} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \left\| T^{-1/2} \varepsilon_i \right\| \left\| T^{-1/2} \hat{F}' \right\| \right)$$

(A.10)

$$= O_p(N^{1/2}T^{1/2})$$

by Assumption C and $\left\| T^{-1/2} \hat{F}' \right\| = \sqrt{T}$. Since $M_F F = M_{\hat{F}} (F^0 - \hat{F}R^{-1})$, $N^{-1/2}T^{-1/2} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \left\| (F^0 - \hat{F}R^{-1})' u_i \right\| = O_p(N^{1/2}T^{1/2}\delta_{NT}^{-2})$ and $N^{-1/2}T^{-1/2} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \left\| \hat{F}' u_i \right\| = O_p(N^{1/2}T^{1/2})$, $A_1$ is bounded in norm by

$$N^{-1/2}T^{-1/2} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \left\| (F^0 - \hat{F}R^{-1})' M_{\hat{F}} u_i \right\| \cdot \left\| N^{-1} T^{-1} \sum_{\ell=1}^N \hat{F}' V_i \Gamma_\ell^0 \right\| \left\| R^{-1} \right\| \left\| \Xi^{-1} \right\|$$

$$= N^{-1/2}T^{-1/2} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \left\| (F^0 - \hat{F}R^{-1})' M_{\hat{F}} u_i \right\| \cdot [O_p(N^{-1}) + O_p(N^{-1/2}T^{-1/2})]$$

$$\leq N^{-1/2}T^{-1/2} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \left\| (F^0 - \hat{F}R^{-1})' u_i \right\| \cdot [O_p(N^{-1}) + O_p(N^{-1/2}T^{-1/2})]$$

$$+ N^{-1/2}T^{-1/2} \sum_{i=1}^N \left\| \Gamma_i^0 \right\| \left\| \hat{F}' u_i \right\| \cdot \left\| T^{-1}(F^0 - \hat{F}R^{-1})' \hat{F}' \right\| \cdot [O_p(N^{-1}) + O_p(N^{-1/2}T^{-1/2})]$$

$$= O_p(N^{-1/2}T^{1/2}\delta_{NT}^{-2}) + O_p(\delta_{NT}^{-2})$$

with (C.5) and Lemma A.1(b). With the definition of $R$, $A_2$ and $A_3$ can be reformulated as

$$A_2 = -N^{-3/2}T^{-1/2} \sum_{i=1}^N \sum_{\ell=1}^N T_i^0 (\mathbf{y})^{-1} T_i^0 V_i M_{\hat{F}} u_i$$

$$A_3 = -N^{-3/2}T^{-1/2} \sum_{i=1}^N \sum_{\ell=1}^N T_i^0 (\mathbf{y})^{-1} (T^{-1} \hat{F}' F^0)^{-1} \hat{F}' V_i V_i' M_{\hat{F}} u_i$$

Combining the above three terms, we can complete the proof. □

Proof of Lemma A.4. Note that $M_{F^0} - M_{\hat{F}} = P_{\hat{F}} - P_{F^0}$ and $P_{\hat{F}} = T^{-1} \hat{F} \hat{F}'$. We can derive
that

\[-N^{-3/2}T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^0(\mathbf{Y}_0)^{-1} \Gamma_i^0 \mathbf{V}_i^t \mathbf{M}_\mathbf{p} u_i - (-N^{-3/2}T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^0(\mathbf{Y}_0)^{-1} \Gamma_i^0 \mathbf{V}_i^t \mathbf{M}_\mathbf{p} u_i) \]

\[=N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^0(\mathbf{Y}_0)^{-1} \Gamma_i^0 \mathbf{V}_i^t (\mathbf{F} - \mathbf{F}_0) \mathbf{R}^t \mathbf{F}_0^t u_i + N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^0(\mathbf{Y}_0)^{-1} \Gamma_i^0 \mathbf{V}_i^t \mathbf{F}_0^t \mathbf{R} (\mathbf{F} - \mathbf{F}_0) u_i \]

\[+ N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^0(\mathbf{Y}_0)^{-1} \Gamma_i^0 \mathbf{V}_i^t (\mathbf{F} - \mathbf{F}_0) (\mathbf{F} - \mathbf{F}_0) u_i \]

\[+ N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^0(\mathbf{Y}_0)^{-1} \Gamma_i^0 \mathbf{V}_i^t \mathbf{F}_0 (\mathbf{R} - (T^{-1} \mathbf{F}_0^t \mathbf{F}_0)^{-1}) \mathbf{F}_0^t u_i \]

\[= \mathbb{B}_1 + \mathbb{B}_2 + \mathbb{B}_3 + \mathbb{B}_4 \]

We first consider the last three terms. Consider the term $\mathbb{B}_2$. By Lemmas A.1 (c) and (A.9), $\mathbb{B}_2$ is bounded in norm by

\[N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \|_1 \| T^{-1}(\mathbf{F} - \mathbf{F}_0 \mathbf{R}) u_i \| \cdot N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_i^0 \mathbf{V}_i^t \mathbf{F}_0 \| \| \mathbf{Y}_0^{-1} \| \mathbf{R} \| = O_p(\delta_{NT}^2) \]

given the fact that $N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_i^0 \mathbf{V}_i^t \mathbf{F}_0 = O_p(1)$. $\mathbb{B}_3$ is bounded in norm by

\[N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \|_1 \| T^{-1}(\mathbf{F} - \mathbf{F}_0 \mathbf{R}) u_i \| \cdot N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_i^0 \mathbf{V}_i^t \mathbf{F}_0 \| \| \mathbf{Y}_0^{-1} \| \mathbf{R} \| = O_p(\delta_{NT}^2) \cdot \left[ O_p(N^{-1/2} T^{1/2}) + O_p(T^{1/2} \delta_{NT}^{-2}) \right] = O_p(N^{-1/2} T^{1/2} \delta_{NT}^{-2}) + O_p(T^{1/2} \delta_{NT}^{-4}) \]

by Lemmas A.1 (f) and (A.9). $\mathbb{B}_4$ is bounded in norm by

\[N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \|_1 \| T^{-1/2} u_i \| \cdot N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_i^0 \mathbf{V}_i^t \mathbf{F}_0 \| \| \mathbf{R} \| (T^{-1} \mathbf{F}_0^t \mathbf{F}_0)^{-1} \| \| T^{-1/2} \mathbf{F}_0 \| \| \mathbf{Y}_0^{-1} \| = O_p(\delta_{NT}^2) \]

by Lemma A.1 (d), and $N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \|_1 \| T^{-1/2} u_i \| = O_p(1)$ which can be proved similar to (A.10). $\mathbb{B}_1$ is decomposed as

\[\mathbb{B}_1 = N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^0(\mathbf{Y}_0)^{-1} \Gamma_i^0 \mathbf{V}_i^t (\mathbf{F} \mathbf{R}^{-1} - \mathbf{F}_0)(T^{-1} \mathbf{F}_0^t \mathbf{F}_0)^{-1} \mathbf{F}_0^t u_i \]

\[+ N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^0(\mathbf{Y}_0)^{-1} \Gamma_i^0 \mathbf{V}_i^t (\mathbf{F} \mathbf{R}^{-1} - \mathbf{F}_0)(\mathbf{R} - (T^{-1} \mathbf{F}_0^t \mathbf{F}_0)^{-1}) \mathbf{F}_0^t u_i \]

\[= \mathbb{B}_{1,1} + \mathbb{B}_{1,2} \]

Since $N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \|_1 \| T^{-1} \mathbf{F}_0^t u_i \| = O_p(1)$, which can be proved by following the argument in (A.10), the term $\mathbb{B}_{1,2}$ is bounded in norm by

\[N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \|_1 \| T^{-1} \mathbf{F}_0^t u_i \| \cdot N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_i^0 \mathbf{V}_i^t (\mathbf{F} \mathbf{R}^{-1} - \mathbf{F}_0)(\mathbf{R} - (T^{-1} \mathbf{F}_0^t \mathbf{F}_0)^{-1}) \| \| \mathbf{Y}_0^{-1} \| = O_p(N^{-1/2} T^{1/2} \delta_{NT}^{-2}) + O_p(T^{1/2} \delta_{NT}^{-4}) \]

by Lemmas A.1 (d), (f).
We consider the term $\mathcal{B}_{1.1}$. By (A.1), $\mathcal{B}_{1.1}$ is decomposed as

$$
N^{-5/2}T^{-5/2} \sum_{i=1}^{N} \sum_{t=1}^{N} \sum_{h=1}^{N} \Gamma^{0}_{t,i}(Y^{0})^{-1} \Gamma^{0}_{t,i} V^{0}_{t,i} F^{0}_{t,i} \hat{V}^{0}_{h} F^{0}(T^{-1} F^{0} \hat{F})^{-1} \Gamma^{0}_{t,i} \Gamma^{0}_{t,i} (T^{-1} F^{0} \hat{F})^{-1} (T^{-1} F^{0} F^{0})^{-1} F^{0} u_{i} \\
+ N^{-5/2}T^{-5/2} \sum_{i=1}^{N} \sum_{t=1}^{N} \sum_{h=1}^{N} \Gamma^{0}_{t,i}(Y^{0})^{-1} \Gamma^{0}_{t,i} V^{0}_{t,i} \hat{V}^{0}_{h} (\hat{F} - F^{0} R) (T^{-1} F^{0} \hat{F})^{-1} (Y^{0})^{-1} (T^{-1} F^{0} F^{0})^{-1} F^{0} u_{i} \\
+ N^{-5/2}T^{-5/2} \sum_{i=1}^{N} \sum_{t=1}^{N} \sum_{h=1}^{N} \Gamma^{0}_{t,i}(Y^{0})^{-1} \Gamma^{0}_{t,i} V^{0}_{t,i} \hat{V}^{0}_{h} (T^{-1} F^{0} F^{0})^{-1} (Y^{0})^{-1} (T^{-1} F^{0} F^{0})^{-1} F^{0} u_{i} \\
= \mathcal{B}_{1.1} + \mathcal{B}_{1.1.2} + \mathcal{B}_{1.1.3} + \mathcal{B}_{1.1.4}
$$

The term $\mathcal{B}_{1.1.1}$ is bounded in norm by

$$
N^{-3/2} T^{-1/2} \sum_{i=1}^{N} \left\| \Gamma^{0}_{t,i} \right\| \left\| T^{-1/2} u_{i} \right\| \cdot \left\| N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \left\| \Gamma^{0}_{t,i} V^{0}_{t,i} F^{0} \right\| ^{2} \right. \\
\times \left. \left\| (T^{-1} F^{0} \hat{F})^{-1} (T^{-1} F^{0} F^{0})^{-1} (Y^{0})^{-1} \right\| \left\| (Y^{0})^{-1} \right\| ^{2} \right) = O_{p}(N^{-1/2} T^{-1/2})
$$

by Lemma A.1 (e). Similarly, we can show the term $\mathcal{B}_{1.1.2}$ is bounded in norm by

$$
N^{-3/2} T^{-1/2} \sum_{i=1}^{N} \left\| \Gamma^{0}_{t,i} \right\| \left\| T^{-1/2} u_{i} \right\| \cdot \left\| N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \left\| \Gamma^{0}_{t,i} V^{0}_{t,i} (\hat{F} - F^{0} R) \right\| \right. \\
\times \left. \left\| (T^{-1} F^{0} \hat{F})^{-1} (T^{-1} F^{0} F^{0})^{-1} (Y^{0})^{-1} \right\| \left\| (Y^{0})^{-1} \right\| ^{2} \right) = O_{p}(N^{-1}) + O_{p}(N^{-1/2} \delta_{N T}^{-2})
$$

by Lemma A.1 (f). $\mathcal{B}_{1.1.3}$ is the leading term, which is reformulated as

$$
N^{-1/2} T^{-1/2} \cdot N^{-1} \sum_{i=1}^{N} \Gamma^{0}_{t,i}(Y^{0})^{-1} (T^{-1} T^{-1} T^{-1} \sum_{\ell=1}^{N} \left\| \Gamma^{0}_{t,i} V^{0}_{t,i} \hat{V}^{0}_{h} \right\| \left\| \Gamma^{0}_{t,i} \hat{V}^{0}_{h} (\hat{F} - F^{0} R) \right. \right. \\
\left. \times \left. \left\| (T^{-1} F^{0} \hat{F})^{-1} (T^{-1} F^{0} F^{0})^{-1} (Y^{0})^{-1} \right\| \left\| (Y^{0})^{-1} \right\| ^{2} \right) \right) = O_{p}(1)
$$

For the term $\mathcal{B}_{1.1.4}$, it is bounded in norm by

$$
N^{-1} \sum_{i=1}^{N} \left\| \Gamma^{0}_{t,i} \right\| \left\| T^{-1/2} u_{i} \right\| \cdot \left\| (T^{-1} F^{0} \hat{F})^{-1} (T^{-1} F^{0} F^{0})^{-1} (T^{-1} F^{0} F^{0})^{-1} \right\| \left\| (Y^{0})^{-1} \right\| ^{2} \\
\times \left\| N^{-3/2} T^{-3/2} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \left\| \Gamma^{0}_{t,i} V^{0}_{t,i} \hat{V}^{0}_{h} \hat{V}^{0}_{h} \right\| \left\| \Gamma^{0}_{t,i} \hat{V}^{0}_{h} (\hat{F} - F^{0} R) \right. \right. \\
\left. \times \left. \left\| (T^{-1} F^{0} \hat{F})^{-1} (T^{-1} F^{0} F^{0})^{-1} (Y^{0})^{-1} \right\| \left\| (Y^{0})^{-1} \right\| ^{2} \right) \right) = O_{p}(1)
$$

Comparing the first term is bounded in norm by

$$
N^{-1/2} \left\| N^{-1/2} T^{-1} \sum_{h=1}^{N} (V_{h} V_{h}^{0} - E(V_{h} V_{h}^{0}) F^{0}) \right\| \left\| R \right. \right. \\
\left. \times \left. \left\| N^{-1} T^{-1} \sum_{h=1}^{N} V_{h} V_{h}^{0} \right\| \left\| T^{-1/2} (\hat{F} - F^{0} R) \right. \right. \\
= O_{p}(1)
$$

the first term is bounded in norm by

$$
N^{-1/2} \left[ N^{-1/2} T^{-1} \sum_{h=1}^{N} (V_{h} V_{h}^{0} - E(V_{h} V_{h}^{0}) F^{0}) \right] \left\| R \right. \right. \\
\left. \times \left. \left\| N^{-1} T^{-1} \sum_{h=1}^{N} V_{h} V_{h}^{0} \right\| \left\| T^{-1/2} (\hat{F} - F^{0} R) \right. \right. \\
= O_{p}(1)
$$

the first term is bounded in norm by

$$
N^{-1/2} \left[ N^{-1/2} T^{-1} \sum_{h=1}^{N} (V_{h} V_{h}^{0} - E(V_{h} V_{h}^{0}) F^{0}) \right] \left\| R \right. \right. \\
\left. \times \left. \left\| N^{-1} T^{-1} \sum_{h=1}^{N} V_{h} V_{h}^{0} \right\| \left\| T^{-1/2} (\hat{F} - F^{0} R) \right. \right. \\
= O_{p}(1)
$$
=O_p(N^{-1/2}) + O_p(T^{1/2}\delta^{-2}_{NT})

by Lemmas A.1 (a), (A.5) and (A.8). The second term is bounded in norm by

\[
T^{-1/2} \left\| T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} (N^{-1/2} \sum_{\ell=1}^{N} \Gamma_\ell^0 V_{ts}) \cdot f_t^0 \text{tr}(N^{-1} \sum_{h=1}^{N} \Sigma_{hh,ts}) \right\|
\]

= \leq kT^{-3/2} \sum_{s=1}^{T} \sum_{t=1}^{T} (N^{-1/2} \sum_{\ell=1}^{N} \Gamma_\ell^0 V_{ts}) \|f_t^0\| \bar{\tau}_{ts} = O_p(T^{-1/2})

since tr(A) ≤ k\|A\| for any k × k matrix A, Assumption B2 and

\[
\mathbb{E}(T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} (N^{-1/2} \sum_{\ell=1}^{N} \Gamma_\ell^0 V_{ts}) \|f_t^0\| \bar{\tau}_{ts})
\]

= T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \bar{\tau}_{ts} \sqrt{\mathbb{E}(f_t^0)^2 \mathbb{E}(N^{-1/2} \sum_{\ell=1}^{N} \Gamma_\ell^0 V_{ts})^2} \leq CT^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \bar{\tau}_{ts} \sqrt{N^{-1} \sum_{\ell=1}^{N} \sum_{j=1}^{N} \text{tr}(\Sigma_{\ell j,ss} f_t^0 \Gamma_\ell^0)}
\]

= kCT^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \bar{\tau}_{ts} \sqrt{N^{-1} \sum_{\ell=1}^{N} \sum_{j=1}^{N} \text{tr}(\Sigma_{\ell j,ss} f_t^0 \Gamma_\ell^0)} \leq C

by Assumptions B2, C and D. Then \( \mathbb{E}_{1.1.4} = O_p(T^{1/2}\delta^{-2}_{NT}) + O_p(\delta^{-1}_{NT}) \). Combining the above terms, we complete the proof. □

**Proof of Lemma A.5.** By substracting and adding terms, we have

\[
N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \Gamma_i^0 (Y^0) - 1 (T^{-1} \bar{F}'F^0) - 1 \bar{F}'V_i V_i' M_{\bar{F}} u_i
\]

- \( N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \Gamma_i^0 (Y^0) - 1 (T^{-1} F^0'F^0) - 1 F^0'V_i V_i' M_{F^0} u_i\)

= \( N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \Gamma_i^0 (Y^0) - 1 (T^{-1} \bar{F}'F^0) - 1 \bar{F}' (V_i V_i' - E(V_i V_i')) M_{\bar{F}} u_i\)

+ \( N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \Gamma_i^0 (Y^0) - 1 (T^{-1} F^0'F^0) - 1 F^0' (V_i V_i' - E(V_i V_i')) M_{F^0} u_i\)

= C_1 + C_2 + C_3

8
Consider $\mathbf{C}_1$. As $\mathbf{M}_F = \mathbf{I}_T - T^{-1} \hat{\mathbf{F}} \hat{\mathbf{F}}'$, it is bounded in norm by

$$N^{-3/2}T^{-3/2} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| \hat{\mathbf{F}}' \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \mathbf{M}_F \mathbf{u}_i \| \cdot \|(\mathbf{F}_0')^{-1}\| \cdot \|(T^{-1} \hat{\mathbf{F}}' \mathbf{F}_0')^{-1}\|$$

$$\leq N^{-3/2}T^{-3/2} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| \hat{\mathbf{F}}' \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \mathbf{u}_i \cdot O_p(1)$$

$$+ N^{-3/2}T^{-5/2} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| \hat{\mathbf{F}}' \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \hat{\mathbf{F}}' \mathbf{u}_i \cdot O_p(1)$$

$$\leq N^{-3/2}T^{-3/2} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| \varphi_0^0 \| \cdot \| \hat{\mathbf{F}}' \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \mathbf{H}_0^0 \cdot O_p(1)$$

$$+ N^{-3/2}T^{-3/2} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| \varphi_0^0 \| \cdot \| \hat{\mathbf{F}}' \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \mathbf{u}_i \cdot O_p(1)$$

$$+ N^{-3/2}T^{-5/2} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| \varphi_0^0 \| \cdot \| \hat{\mathbf{F}}' \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \mathbf{H}_0' \cdot \hat{\mathbf{F}} \cdot O_p(1)$$

Consider the first term. Following the argument in the proof of Lemma A.2(i) in Bai (2009), we can show that $\|N^{-1/2}T^{-1} \sum_{\ell=1}^N \mathbf{F}_0^0 (\mathbf{V}_\ell \mathbf{V}_\ell' - \mathbb{E}(\mathbf{V}_\ell \mathbf{V}_\ell')) \mathbf{H}_0^0 \| = O_p(1)$. In addition, similar to (A.8), we can show that $\|N^{-1/2}T^{-1} \sum_{\ell=1}^N (\mathbf{V}_\ell \mathbf{V}_\ell' - \mathbb{E}(\mathbf{V}_\ell \mathbf{V}_\ell')) \mathbf{H}_0^0 \| = O_p(1)$. The first term is bounded in norm by

$$T^{-1/2} \cdot N^{-1} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| \varphi_0^0 \| \cdot \| N^{-1/2}T^{-1} \mathbf{F}_0^0 \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \mathbf{H}_0^0 \| \cdot \| \mathbf{R} \|$$

$$+ N^{-1} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| \varphi_0^0 \| \cdot \| T^{-1/2} (\hat{\mathbf{F}} - \mathbf{F}_0^0 \mathbf{R}) \| \| N^{-1/2}T^{-1} \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \mathbf{H}_0^0 \| = O_p(\delta_{NT}^{-1})$$

by Lemma A.1 (a) and (c). Similar to the proof of the first term, the second term is bounded in norm by

$$T^{-1/2} \cdot N^{-1} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| N^{-1/2}T^{-1} \mathbf{F}_0^0 \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \mathbf{u}_i \| \cdot \| \mathbf{R} \|$$

$$+ N^{-1} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| N^{-1/2}T^{-1} \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \mathbf{u}_i \| \cdot \| T^{-1/2} (\hat{\mathbf{F}} - \mathbf{F}_0^0 \mathbf{R}) \| = O_p(\delta_{NT}^{-1})$$

by Lemma A.1 (a). The third term is bounded in norm by

$$T^{-1/2} \cdot N^{-1} \sum_{i=1}^N \| \mathbf{F}_i^0 \| \| T^{-1/2} \mathbf{u}_i \| \cdot \| T^{-1/2} \hat{\mathbf{F}} \| \| N^{-1/2}T^{-1} \hat{\mathbf{F}} \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \hat{\mathbf{F}} \|$$

$$= O_p(T^{-1/2}) \cdot \| N^{-1/2}T^{-1} \hat{\mathbf{F}} \| \sum_{t=1}^N (\mathbf{V}_t \mathbf{V}_t' - \mathbb{E}(\mathbf{V}_t \mathbf{V}_t')) \hat{\mathbf{F}} \| = O_p(\delta_{NT}^{-1}) + O_p(T^{1/2} \delta_{NT}^{-2})$$

9
because

\[ \|N^{-1/2}T^{-1}\tilde{F}'\sum_{\ell=1}^{N} (V_{\ell}V'_{\ell} - \mathbb{E}(V_{\ell}V'_{\ell})) \tilde{F}\| \]

\[ \leq \|N^{-1/2}T^{-1}F^0\sum_{\ell=1}^{N} (V_{\ell}V'_{\ell} - \mathbb{E}(V_{\ell}V'_{\ell})) F^0\|\|R\|^2 \]

\[ + 2T^{1/2} \cdot \|N^{-1/2}T^{-1}F^0\sum_{\ell=1}^{N} (V_{\ell}V'_{\ell} - \mathbb{E}(V_{\ell}V'_{\ell})) \|\|T^{-1/2}(\tilde{F} - F^0R)\|\|R\| \]

\[ + T \cdot \|N^{-1/2}T^{-1}\sum_{\ell=1}^{N} (V_{\ell}V'_{\ell} - \mathbb{E}(V_{\ell}V'_{\ell})) \|\|T^{-1/2}(\tilde{F} - F^0R)\|^2 \]

\[ = O_p(T^{1/2}\delta_{NT}^{-1}) + O_p(T\delta_{NT}^{-2}) \]

by (A.7), (A.8) and Lemmas A.1 (a), (c), and the fact that \(\|N^{-1/2}T^{-1}\sum_{\ell=1}^{N} F^0 (V_{\ell}V'_{\ell} - \mathbb{E}(V_{\ell}V'_{\ell})) F^0\| = O_p(1)\) by following the argument in the proof of Lemma A.2(i) in Bai (2009). Collecting the above terms, \(C_1 = O_p(\delta_{NT}^{-1}) + O_p(T^{1/2}\delta_{NT}^{-2})\).

Consider \(C_2\) and \(C_3\). Note that by Assumptions B2, we have

\[ \|\mathbb{E}(V_{\ell}V'_{\ell})(M_{F}u_\ell)\| \leq \mu_{\max}(\mathbb{E}(V_{\ell}V'_{\ell})) \|M_{F}u_\ell\| \leq C\|u_\ell\| \]

\[ \|\mathbb{E}(V_{\ell}V'_{\ell})(M_{F} - M_{F^0})u_\ell\| \leq \mu_{\max}(\mathbb{E}(V_{\ell}V'_{\ell})) \|M_{F} - M_{F^0}\|\|u_\ell\| = O_p(\delta_{NT}^{-1})\|u_\ell\| \]

In addition,

\[ \|((T^{-1/2}\tilde{F}'F^0)^{-1}(T^{-1/2}\tilde{F}')(T^{-1/2}F^0) - (T^{-1/2}\tilde{F}'F^0)^{-1}(T^{-1/2}\tilde{F}')(T^{-1/2}F^0))M_{F^0}\| \]

\[ = \|((T^{-1/2}\tilde{F}'F^0)^{-1}(T^{-1/2}(\tilde{F} - F^0R))(T^{-1/2}F^0) - (T^{-1/2}\tilde{F}'F^0)^{-1}(T^{-1/2}(\tilde{F} - F^0R))(T^{-1/2}F^0))\| = O_p(\delta_{NT}^{-1}) \]

then \(C_2\) is bounded in norm by

\[ \sqrt{N} T N^{-2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \|\tilde{p}\|\|\mathbb{E}(V_{\ell}V'_{\ell})(M_{F}u_\ell)\|\|((T^{-1/2}\tilde{F}'F^0)^{-1}(T^{-1/2}\tilde{F}')(T^{-1/2}F^0) - (T^{-1/2}\tilde{F}'F^0)^{-1}(T^{-1/2}\tilde{F}')(T^{-1/2}F^0))\|\|\|Y^{-1}\| \]

\[ \leq \sqrt{N} T N^{-2} \sum_{i=1}^{N} \|\tilde{p}\|\|T^{-1/2}u_\ell\|\|((T^{-1/2}\tilde{F}'F^0)^{-1}(T^{-1/2}\tilde{F}')(T^{-1/2}F^0) - (T^{-1/2}\tilde{F}'F^0)^{-1}(T^{-1/2}\tilde{F}')(T^{-1/2}F^0))\|\|\|Y^{-1}\| \]

\[ = O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-1}) \]

and \(C_3\) is bounded in norm by

\[ \sqrt{N} T N^{-2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \|\tilde{p}\|\|T^{-1/2}\mathbb{E}(V_{\ell}V'_{\ell})(M_{F} - M_{F^0})u_\ell\|\|((T^{-1/2}F^0F^0)^{-1}(T^{-1/2}\tilde{F}')(T^{-1/2}F^0) - (T^{-1/2}F^0F^0)^{-1}(T^{-1/2}\tilde{F}')(T^{-1/2}F^0))\|\|\|Y^{-1}\| \]

\[ = O_p(N^{1/2}T^{-1/2}\delta_{NT}^{-1}) \]

by the above three facts. This completes the proof. \(\square\)

**Proof of Lemma A.6.** Since \(M_{F} - M_{F^0} = -T^{-1}(\tilde{F} - F^0R)R'F^0 - T^{-1}F^0R(\tilde{F} - F^0R)' - \)
we can prove that

We first consider the last four terms. Following the argument in the proof of Lemma A.2(b),

\[ F_0 = \|\mathbf{R}^{-1} - \mathbf{F}_0\|^{1/2} \sum_{i=1}^N V_i' \mathbf{M}_i \mathbf{M}_F u_i \]

\[ = -N^{-1/2} T^{-3/2} \sum_{i=1}^N V_i' (\mathbf{FR}^{-1} - \mathbf{F}_0)(T^{-1} \mathbf{F}_0^0 \mathbf{F}_0^0)^{-1} \mathbf{F}_0^0 u_i \]

\[ = -N^{-1/2} T^{-3/2} \sum_{i=1}^N V_i' (\mathbf{FR}^{-1} - \mathbf{F}_0)(\mathbf{RR}' - (T^{-1} \mathbf{F}_0^0 \mathbf{F}_0^0)^{-1}) \mathbf{F}_0^0 u_i \]

\[ = -N^{-1/2} T^{-3/2} \sum_{i=1}^N V_i' (\mathbf{FR}^{-1} - \mathbf{F}_0)(\mathbf{RR}' - (T^{-1} \mathbf{F}_0^0 \mathbf{F}_0^0)^{-1}) \mathbf{F}_0^0 u_i \]

\[ = D_1 + D_2 + D_3 + D_4 + D_5 \]

We first consider the last four terms. Following the argument in the proof of Lemma A.2(b),

\[ \sum_{i=1}^N \|T^{-1} V_i' (\mathbf{FR}^{-1} - \mathbf{F}_0)\| \|\varphi_i^0\| = O_p(\delta^{-2}_{N_f}) \]

\[ \sum_{i=1}^N \|T^{-1} V_i' (\mathbf{FR}^{-1} - \mathbf{F}_0)\| \|\varphi_i^0\| = O_p(\delta^{-2}_{N_f}) \]

\[ = O_p(N^{1/2} T^{1/2} \delta^{-2}_{N_f}) \]

Following the argument in the proof of Lemma A.2(a), we derive \( \sum_{i=1}^N \|T^{-1/2} V_i' \mathbf{F}_0^0\| T^{-1} (\mathbf{FR}^{-1} - \mathbf{F}_0) \varepsilon_i\| = O_p(\delta^{-2}_{N_f}) \). Then, \( D_3 \) is bounded in norm by

\[ N^{1/2} \sum_{i=1}^N \|T^{-1/2} V_i' \mathbf{F}_0^0\| \|\varphi_i^0\| T^{-1} (\mathbf{FR}^{-1} - \mathbf{F}_0) \varepsilon_i\| = O_p(\delta^{-2}_{N_f}) \]

by Lemmas A.1 (b), (c). Since \( \sum_{i=1}^N \|T^{-1} V_i' (\mathbf{FR}^{-1} - \mathbf{F}_0)\| \|\varphi_i^0\| = O_p(\delta^{-2}_{N_f}) \), \( D_4 \) is bounded in norm by

\[ N^{1/2} \sum_{i=1}^N \|T^{-1} V_i' (\mathbf{FR}^{-1} - \mathbf{F}_0)\| \|\varphi_i^0\| T^{-1} (\mathbf{FR}^{-1} - \mathbf{F}_0) \varepsilon_i\| = O_p(\delta^{-2}_{N_f}) \]

by Lemmas A.1(b) and A.2(c). \( D_5 \) is bounded in norm by

\[ N^{1/2} \sum_{i=1}^N \|T^{-1/2} V_i' \mathbf{F}_0^0\| \|\varphi_i^0\| T^{-1} (\mathbf{FR}^{-1} - \mathbf{F}_0) \varepsilon_i\| = O_p(\delta^{-2}_{N_f}) \]

\[ + N^{1/2} \sum_{i=1}^N \|T^{-1/2} V_i' \mathbf{F}_0^0\| \|\varphi_i^0\| T^{-1} (\mathbf{FR}^{-1} - \mathbf{F}_0) \varepsilon_i\| = O_p(\delta^{-2}_{N_f}) \]
by Lemma A.1(d).

Now consider the term $\mathcal{D}_1$. With (A.1), we decompose the term $\mathcal{D}_1$ as follows

\[
- N^{-3/2} p^{3/2} \sum_{i=1}^N \sum_{h=1}^N V_i'(V_h, \Gamma_h^0) (T^{-1}F^0)' - (T^{-1}F^0)' - 1 F^0 u_i
\]

\[
- N^{-3/2} p^{5/2} \sum_{i=1}^N \sum_{h=1}^N V_i'F^0T^0V_j^0(V_h, \Phi(T^{-1}F^0)' - (T^{-1}F^0)' - 1 F^0 u_i
\]

\[
- N^{-3/2} p^{5/2} \sum_{i=1}^N \sum_{h=1}^N V_i'(V_h, V_j^0) V_j^0(V_h, \Phi(T^{-1}F^0)' - (T^{-1}F^0)' - 1 F^0 u_i
\]

\[
- N^{-3/2} p^{5/2} \sum_{i=1}^N \sum_{h=1}^N (V_i'(V_h, V_j^0) V_j^0(V_h, \Phi(T^{-1}F^0)' - (T^{-1}F^0)' - 1 F^0 u_i
\]

\[
= \mathcal{D}_{1,1} + \mathcal{D}_{1,2} + \mathcal{D}_{1,3} + \mathcal{D}_{1,4}
\]

We consider the last three terms $\mathcal{D}_{1,2}, \mathcal{D}_{1,3},$ and $\mathcal{D}_{1,4}$. $\mathcal{D}_{1,2}$ is bounded in norm by

\[
T^{-1} / N^{-1} \sum_{i=1}^N \|T^{-1/2} V_i'(F^0)\| \|T^{-1/2} u_i\| \cdot \|N^{-1/2} T^{-1/2} \sum_{i=1}^N \Gamma_i^0 V_i \Phi\| \times \|T^{-1} F^0\|^{-1} \|T^{-1} F^0\|^{-1} \|T^{-1} F^0\|^{-1} \|T^{-1} F^0\|^{-1}
\]

\[
= O_p(T^{-1/2}) \cdot \|N^{-1/2} T^{-1/2} \sum_{i=1}^N \Gamma_i^0 V_i \Phi\| = O_p(\delta_{N/T}^{-1})
\]

by (C.5). As $\|E(T^{-1} V_i'(V_h))\| \leq \bar{r}_{ih}$ by Assumption B2, $\mathcal{D}_{1,3}$ is bounded in norm by

\[
N^{-3/2} p^{3/2} \sum_{i=1}^N \sum_{h=1}^N \|T^{-1/2} u_i\| \|E(V_i'(V_h)) \| \|V_i \Phi\| \cdot \|T^{-1} F^0\|^{-1} \|T^{-1} F^0\|^{-1} \|T^{-1} F^0\|^{-1} \|T^{-1} F^0\|
\]

\[
= N^{-3/2} p^{3/2} \sum_{i=1}^N \sum_{h=1}^N \|T^{-1/2} u_i\| \|E(V_i'(V_h)) \| \|V_i \Phi\| \cdot O_p(1)
\]

\[
\leq N^{-3/2} p^{3/2} \sum_{i=1}^N \sum_{h=1}^N (\|T^{-1/2} H\| \|\varphi_0'\| + \|T^{-1/2} \varphi_1'\|) \|T^{-1} E(V_i'(V_h)) \| \|T^{-1/2} V_i F^0\| \|R\| \cdot O_p(1)
\]

\[
+ N^{-3/2} p^{1/2} \sum_{i=1}^N \sum_{h=1}^N (\|T^{-1/2} H\| \|\varphi_0'\| + \|T^{-1/2} \varphi_1'\|) \|T^{-1} E(V_i'(V_h)) \| \|T^{-1/2} V_i F^0\| \|T^{-1/2}(\Phi - F^0 R)\| \cdot O_p(1)
\]

\[
\leq N^{-3/2} p^{3/2} \sum_{i=1}^N \sum_{h=1}^N \bar{r}_{ih} \|\varphi_0'\| \|T^{-1/2} V_i F^0\| \cdot O_p(1) + N^{-3/2} p^{3/2} \sum_{i=1}^N \sum_{h=1}^N \bar{r}_{ih} \|T^{-1/2} \varphi_1'\| \|T^{-1/2} V_i F^0\| \cdot O_p(1)
\]

\[
+ N^{-3/2} p^{1/2} \sum_{i=1}^N \sum_{h=1}^N \bar{r}_{ih} \|\varphi_0'\| \|T^{-1/2} V_i F^0\| \cdot O_p(\delta^{-1}_{N/T}) + N^{-3/2} p^{1/2} \sum_{i=1}^N \sum_{h=1}^N \bar{r}_{ih} \|T^{-1/2} \varphi_1'\| \|T^{-1/2} V_i F^0\| \cdot O_p(\delta^{-1}_{N/T})
\]

The first term is $O_p(N^{-1/2})$ since $E(N^{-1} \sum_{i=1}^N \sum_{h=1}^N \bar{r}_{ih} \|\varphi_0'\| \|T^{-1/2} V_i F^0\|)$ is bounded by

\[
N^{-1} \sum_{i=1}^N \sum_{h=1}^N \bar{r}_{ih} \|\varphi_0'\| \|T^{-1/2} V_i F^0\|^2 \leq C N^{-1} \sum_{i=1}^N \sum_{h=1}^N \bar{r}_{ih} \leq C
\]

by Assumption B2, C and D. Similarly, we can show that the second term is $O_p(N^{-1/2})$, while the third and the fourth terms both are $O_p(N^{-1/2} T^{1/2} \delta^{-1}_{N/T})$. Thus, $\mathcal{D}_{1,3}$ is $O_p(N^{-1/2})+O_p(N^{-1/2} T^{1/2} \delta^{-1}_{N/T})$. 

12
\( \mathcal{D}_{1.4} \) is decomposed as following

\[ \begin{align*}
&-N^{-3/2}T^{-5/2} \sum_{i=1}^{N} \sum_{h=1}^{N} (V'_i V_h - E(V'_i V_h)) V'_i F_0^0 R (T^{-1} F_0^0 \hat{F})^{-1} (T^{-1} F_0^0 \hat{F})^{-1} F_0^0 H^0 \varphi_i^0 \\
&-N^{-3/2}T^{-5/2} \sum_{i=1}^{N} \sum_{h=1}^{N} (V'_i V_h - E(V'_i V_h)) V'_h F_0^0 R (T^{-1} F_0^0 \hat{F})^{-1} (T^{-1} F_0^0 \hat{F})^{-1} F_0^0 H^0 \varphi_i^0 \\
&-N^{-3/2}T^{-5/2} \sum_{i=1}^{N} \sum_{h=1}^{N} (V'_i V_h - E(V'_i V_h)) V'_h (\hat{F} - F_0^0 R) (T^{-1} F_0^0 \hat{F})^{-1} (T^{-1} F_0^0 \hat{F})^{-1} F_0^0 H^0 \varphi_i^0 \\
&-N^{-3/2}T^{-5/2} \sum_{i=1}^{N} \sum_{h=1}^{N} (V'_i V_h - E(V'_i V_h)) V'_h (\hat{F} - F_0^0 R) (T^{-1} F_0^0 \hat{F})^{-1} (T^{-1} F_0^0 \hat{F})^{-1} F_0^0 H^0 \varphi_i^0 \\
= &\mathbb{D}_{1.4.1} + \mathbb{D}_{1.4.2} + \mathbb{D}_{1.4.3} + \mathbb{D}_{1.4.4}
\end{align*} \]

Consider \( \mathbb{D}_{1.4.1} \). As vec \((ABC) = (C' \otimes A)\)vec(B) for any comfortable matrices \(A, B\) and \(C\), \( \mathbb{D}_{1.4.1} \) is bounded in norm by

\[ \begin{align*}
&\left\| \text{vec}(N^{-3/2}T^{-5/2} \sum_{h=1}^{N} (V'_i V_h - E(V'_i V_h)) \cdot V'_i F_0^0 R (T^{-1} F_0^0 \hat{F})^{-1} (T^{-1} F_0^0 \hat{F})^{-1} F_0^0 H^0 \varphi_i^0) \right\| \\
= &\left\| N^{-3/2}T^{-5/2} \sum_{h=1}^{N} \left( \sum_{i=1}^{N} \varphi_i^0 \otimes (V'_i V_h - E(V'_i V_h)) \right) \right\| \times \text{vec} \left( I_k \cdot \text{vec}(V'_i F_0^0) \right) \\
= &\left\| N^{-3/2}T^{-5/2} \sum_{h=1}^{N} \left( \sum_{i=1}^{N} \varphi_i^0 \otimes (V'_i V_h - E(V'_i V_h)) \right) \times \left( (R (T^{-1} F_0^0 \hat{F})^{-1} (T^{-1} F_0^0 \hat{F})^{-1} F_0^0 H^0)^T \right) \otimes I_k \cdot \text{vec}(V'_i F_0^0) \right\| \\
\leq &N^{-3/2}T^{-3/2} \sum_{h=1}^{N} \left\| \sum_{i=1}^{N} \varphi_i^0 \otimes (V'_i V_h - E(V'_i V_h)) \right\| \cdot \| \text{vec}(V'_i F_0^0) \| \\
&\times \| R \| \| (T^{-1} F_0^0 \hat{F})^{-1} \| \| (T^{-1} F_0^0 \hat{F})^{-1} \| \| T^{-1} F_0^0 H^0 \| \| I_k \| \\
= &O_p(T^{-1/2}) \cdot N^{-1} \sum_{h=1}^{N} \sum_{i=1}^{N} \| N^{-1/2}T^{-1/2} \sum_{i=1}^{T} \varphi_i^0 \otimes (v_{ih} v'_{ih} - E(v_{ih} v'_{ih})) \| T^{-1/2} \sum_{s=1}^{T} v_{hs} f_s^0 \| \\
= &O_p(T^{-1/2})
\end{align*} \]

by Lemma A.1 (c) and because

\[ \mathbb{E} \left( N^{-1} \sum_{h=1}^{N} \left\| N^{-1/2}T^{-1/2} \sum_{i=1}^{T} \varphi_i^0 \otimes (v_{ih} v'_{ih} - E(v_{ih} v'_{ih})) \right\| T^{-1/2} \sum_{s=1}^{T} v_{hs} f_s^0 \| \right) \]

\[ \leq N^{-1} \sum_{h=1}^{N} \sqrt{\mathbb{E} \left( N^{-1/2}T^{-1/2} \sum_{i=1}^{T} \varphi_i^0 \otimes (v_{ih} v'_{ih} - E(v_{ih} v'_{ih})) \right)^2 \mathbb{E} \left( T^{-1/2} \sum_{s=1}^{T} v_{hs} f_s^0 \right)^2} \leq C \]

by Assumption B4 and \( \mathbb{E}\|T^{-1/2} \sum_{s=1}^{T} v_{hs} f_s^0 \|^2 \leq C \), which can be proved easily by Assumption B2. Similarly, we can show that \( \mathbb{D}_{1.4.2} \) is \( O_p(T^{-1/2}) \), while \( \mathbb{D}_{1.4.3} \) and \( \mathbb{D}_{1.4.4} \) both are \( O_p(\delta_{NT}^{-1}) \).
Collecting $D_{1.4.1}$ to $D_{1.4.4}$, we have $O_p(\delta_{NT}^{-1})$. Thus

$$N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i^T(M - M_{0})u_i$$

$$= - N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i^T V_h \Gamma_h^{0}(\mathbf{Y}^0)^{-1}(T^{-1}F^{0}F^{0})^{-1}F^{0}u_i + O_p(N^{1/2}T^{1/2}\delta_{NT}^{-3})$$

This completes the proof. □

**Proof of Lemma A.7.** Denote

$$a_1 = - \frac{1}{NT} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i^T V_h \Gamma_h^{0}(\mathbf{Y}^0)^{-1}(T^{-1}F^{0}F^{0})^{-1}F^{0}u_i$$

$$a_2 = - \frac{1}{NT} \sum_{i=1}^{N} \Gamma_i^{0}(\mathbf{Y}^0)^{-1}(T^{-1}F^{0}F^{0})^{-1}F^{0}\Sigma M_{0}u_i$$

with $X_i = X_i - N^{-1} \sum_{\ell=1}^{N} X_i \Gamma_h^{0}(\mathbf{Y}^0)^{-1}\Gamma_i^{0}$, $V_i = V_i - N^{-1} \sum_{\ell=1}^{N} V_i \Gamma_h^{0}(\mathbf{Y}^0)^{-1}\Gamma_i^{0}$, $\mathbf{Y}^0 = N^{-1} \sum_{i=1}^{N} \Gamma_i^{0}\Gamma_i^{0}$. In addition, $\Sigma = N^{-1} \sum_{\ell=1}^{N} E(V_i V_i^T)$.

Here we investigate the stochastic order of $a_1$ and $a_2$. $a_1$ is decomposed as

$$- \frac{1}{NT} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i^T V_h \Gamma_h^{0}(\mathbf{Y}^0)^{-1}(T^{-1}F^{0}F^{0})^{-1}F^{0}u_i$$

$$= - \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{h=1}^{N} E(V_i V_h)^{0}(\mathbf{Y}^0)^{-1}(T^{-1}F^{0}F^{0})^{-1}F^{0}\varphi_1^{0}$$

$$+ \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{h=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^{0}(\mathbf{Y}^0)^{-1}\Gamma_i^{0} E(V_i V_h) \Gamma_h^{0}(\mathbf{Y}^0)^{-1}(T^{-1}F^{0}F^{0})^{-1}F^{0}\varphi_1^{0}$$

$$- \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{h=1}^{N} (V_i^T V_h - E(V_i V_h)) \Gamma_h^{0}(\mathbf{Y}^0)^{-1}(T^{-1}F^{0}F^{0})^{-1}F^{0}\varphi_1^{0}$$

$$+ \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{h=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^{0}(\mathbf{Y}^0)^{-1}\Gamma_i^{0} (V_i^T V_h - E(V_i V_h)) \Gamma_h^{0}(\mathbf{Y}^0)^{-1}(T^{-1}F^{0}F^{0})^{-1}F^{0}\varphi_1^{0}$$

$$- \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i^T V_h \Gamma_h^{0}(\mathbf{Y}^0)^{-1}(T^{-1}F^{0}F^{0})^{-1}F^{0}\varphi_1^{0}$$

$$+ \frac{1}{NT^2} \sum_{i=1}^{N} \sum_{h=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^{0}(\mathbf{Y}^0)^{-1}\Gamma_i^{0} V_i^T V_h \Gamma_h^{0}(\mathbf{Y}^0)^{-1}(T^{-1}F^{0}F^{0})^{-1}F^{0}\varphi_1^{0}$$

As $\|E(V_i V_h)\| \leq T\bar{\tau}_{ih}$ by Assumption B2, the first term is bounded in norm by

$$N^{-1}T^{-1} \sum_{i=1}^{N} \sum_{h=1}^{N} \|E(V_i V_h)\| \varphi_1^{0} \|\Gamma_h^{0}\| \cdot \|\mathbf{Y}^0\|^{-1}\|\|T^{-1}F^{0}F^{0}\|^{-1}\|T^{-1}F^{0}\|$$

$$\leq N^{-1} \sum_{i=1}^{N} \sum_{h=1}^{N} \bar{\tau}_{ih} \varphi_1^{0} \|\Gamma_h^{0}\| \cdot O_p(1) = O_p(1)$$

with $N^{-1} \sum_{i=1}^{N} \sum_{h=1}^{N} \bar{\tau}_{ih} E(\varphi_1^{0}) \|\Gamma_h^{0}\| \leq N^{-1} \sum_{i=1}^{N} \sum_{h=1}^{N} \bar{\tau}_{ih} E(\|\varphi_1^{0}\|)^2 E(\|\Gamma_h^{0}\|)^2 \leq C N^{-1} \sum_{i=1}^{N} \sum_{h=1}^{N} \bar{\tau}_{ih} \leq C^2$ by Assumption B2. Similarly, we can show that the second term is $O_p(1)$. The third term is
bounded in norm by
\[
\left\| \text{vec}\left( N^{-1}T^{-2} \sum_{i=1}^{N} \sum_{h=1}^{N} (V_i^h V_h - E(V_i^h V_h)) \Gamma_0^0 (Y_0)^{-1} (T^{-1}F_0^0 F_0^{-1} - 1) F_0^0 H_0 \varphi_i^0 \right) \right\| \\
= T^{-1/2} \cdot \left\| N^{-1}T^{-1/2} \sum_{i=1}^{N} \sum_{h=1}^{N} \varphi_i^0 \otimes \left( (V_i^h V_h - E(V_i^h V_h)) \Gamma_0^0 \right) \text{vec}\left( (Y_0)^{-1} (T^{-1}F_0^0 F_0^{-1} - 1) T^{-1}F_0^0 H_0 \right) \right\| \\
\leq T^{-1/2} \cdot \left\| N^{-1}T^{-1/2} \sum_{i=1}^{N} \sum_{h=1}^{N} \sum_{i=1}^{T} \varphi_i^0 \otimes \left( (V_i V_h - E(V_i V_h)) \Gamma_0^0 \right) \right\| \cdot \| (Y_0)^{-1} \| (T^{-1}F_0^0 F_0^{-1} - 1) \| T^{-1}F_0^0 H_0 \| \\
= O_p(T^{-1/2})
\]

Similarly, we can show that the forth term is \( O_p(T^{-1/2}) \). The fifth term is bounded in norm by
\[
\left\| \frac{1}{N T^2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i^h \Gamma_0^0 (Y_0)^{-1} (T^{-1}F_0^0 F_0^{-1} - 1) F_0^0 \varphi_i^0 \right\| + \left\| \frac{1}{N T^2} \sum_{i=1}^{N} \sum_{h=1}^{N} (V_i^h V_h - E(V_i^h V_h)) \Gamma_0^0 (Y_0)^{-1} (T^{-1}F_0^0 F_0^{-1} - 1) F_0^0 \varphi_i^0 \right\|
\]

Similar to the argument in the proof of the first term, the former term is \( O_p(T^{-1/2}) \). Similar to the argument in the proof of the third term, the latter is \( O_p(T^{-1}) \). Then the fifth term is \( O_p(T^{-1/2}) \).

The sixth term is bounded in norm by
\[
T^{-1/2} N^{-1} \sum_{i=1}^{N} \sum_{h=1}^{N} \| \Gamma_0^0 \| \| T^{-1/2} F_0^0 \varphi_i^0 \| \| (Y_0)^{-1} \|^{1/2} \| T^{-1} F_0^0 F_0^{-1} - 1 \| \| T^{-1} F_0^0 H_0 \| = O_p(T^{-1/2})
\]

thus \( a_1 = O_p(1) \) and
\[
a_1 = - \frac{1}{N T^2} \sum_{i=1}^{N} \sum_{h=1}^{N} E(V_i^h V_h) \Gamma_0^0 (Y_0)^{-1} (T^{-1}F_0^0 F_0^{-1} - 1) F_0^0 H_0 \varphi_i^0 \\
+ \frac{1}{N^2 T^2} \sum_{i=1}^{N} \sum_{h=1}^{N} \sum_{i=1}^{N} \Gamma_0^0 (Y_0)^{-1} \Gamma_0^0 E(V_i V_h) \Gamma_0^0 (Y_0)^{-1} (T^{-1}F_0^0 F_0^{-1} - 1) F_0^0 H_0 \varphi_i^0 + O_p(T^{-1/2}).
\]

Next, we have
\[
a_2 = - \frac{1}{N T} \sum_{i=1}^{N} \Gamma_0^0 (Y_0)^{-1} (T^{-1}F_0^0 F_0^{-1} - 1) F_0^0 \Sigma F_0 \varphi_i^0 \\
- \frac{1}{N T} \sum_{i=1}^{N} \Gamma_0^0 (Y_0)^{-1} (T^{-1}F_0^0 F_0^{-1} - 1) F_0^0 \Sigma e_i + \frac{1}{N T} \sum_{i=1}^{N} \Gamma_0^0 (Y_0)^{-1} (T^{-1}F_0^0 F_0^{-1} - 1) F_0^0 \Sigma F_0 (F_0^0 F_0^{-1}) F_0^0 \varphi_i.
\]

The first term is bounded in norm by
\[
N^{-1} \sum_{i=1}^{N} \| \Gamma_0^0 \| \| \varphi_i^0 \| \cdot \| (Y_0)^{-1} \| \| (T^{-1}F_0^0 F_0^{-1} - 1) \| \cdot \| T^{-1} F_0^0 \Sigma F_0 \| \| H_0 \| = T^{-1} \| F_0^0 \Sigma M_{F_0} H_0 \| \cdot O_p(1)
\]

\[
\leq T^{-1} \| F_0^0 \Sigma H_0 \| \cdot O_p(1) + T^{-1} \| F_0^0 \Sigma F_0 \| \| (T^{-1}F_0^0 F_0^{-1} - 1) \| \| T^{-1} F_0^0 H_0 \| \cdot O_p(1)
\]

\[
= T^{-1} \| F_0^0 \Sigma H_0 \| \cdot O_p(1) + T^{-1} \| F_0^0 \Sigma F_0 \| \cdot O_p(1) = O_p(1)
\]

because \( T^{-1} \| F_0^0 \Sigma H_0 \| = O_p(1) \) and \( \| F_0^0 \Sigma F_0 \| = O_p(1) \), where the former holds because
\[
T^{-1} \| F_0^0 \Sigma H_0 \| = N^{-1} T^{-1} \| E \| \sum_{s=1}^{T} \sum_{t=1}^{T} \sum_{i=1}^{N} f_s^0 f_t^0 E(V_{is} V_{it}) \| \leq N^{-1} T^{-1} \| E \| \sum_{s=1}^{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \| f_s^0 \| \| f_t^0 \| \| E \| \| f_s^0 \| \| f_t^0 \| \| \bar{r}_{st} \|
\]

\[
\leq N^{-1} T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \sum_{i=1}^{N} \sqrt{E \| f_s^0 \|^2 \| f_t^0 \|^2} \| \bar{r}_{st} \| \leq C T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \| \bar{r}_{st} \| \leq C
\]
and the latter also holds whose proof is similar to that of the former. The second is equal to

\[-N^{-2}T^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} T_i^0 (Y_i^0)^{-1} (T^{-1} F_0^0 F_0^0)^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} f_i^0 \varepsilon (v_{is}^t v_{it}) \varepsilon_{it} \]

\[= -N^{-2}T^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{t=1}^{T} \mathbb{E} (v_{is}^t v_{it}) \sum_{i=1}^{N} \Gamma_i^0 \varepsilon_{it} (Y_i^0)^{-1} (T^{-1} F_0^0 F_0^0)^{-1} f_i^0 \]

which is bounded in norm by

\[
N^{-2}T^{-1} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{r}_{st} \| \sum_{i=1}^{N} \Gamma_i^0 \varepsilon_{it} \| \| f_i^0 \| \cdot \| (Y_i^0)^{-1} \| \| (T^{-1} F_0^0 F_0^0)^{-1} \|
\leq N^{-1/2} \sum_{i=1}^{N} \sum_{s=1}^{T} \tilde{r}_{st} \| N^{-1/2} \sum_{i=1}^{N} \Gamma_i^0 \varepsilon_{it} \| \| f_i^0 \| \cdot O_p(1) = O_p(N^{-1/2})
\]

because

\[
T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{r}_{st} \mathbb{E} \| N^{-1/2} \sum_{i=1}^{N} \Gamma_i^0 \varepsilon_{it} \| \leq T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{r}_{st} \mathbb{E} \| N^{-1/2} \sum_{i=1}^{N} \Gamma_i^0 \varepsilon_{it} \| ^2 \leq CT^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{r}_{st} \leq C \text{ by Assumption B2, A1, C and D.}
\]

The third term is bounded in norm by

\[
T^{-1/2} \sum_{i=1}^{N} \| \| f_i^0 \| \| T^{-1/2} F_0^0 \varepsilon_{i} \| \cdot \| (Y_i^0)^{-1} \| \| T^{-1} F_0^0 F_0^0)^{-1} \| \| T^{-1} F_0^0 F_0^0 \| = O_p(T^{-1/2})
\]

Then \( a_2 = O_p(1) \) and

\[
a_2 = -\frac{1}{NT} \sum_{i=1}^{N} \Gamma_i^0 (Y_i^0)^{-1} (T^{-1} F_0^0 F_0^0)^{-1} F_0^0 \Sigma F_0^0 \varphi_i^0 + O_p(\delta_1^{-1} N T^-1).
\]

This completes the proof.

**B Proofs of Lemmas in Appendix B**

**Proof of Lemma B.1.** With \( \tilde{\beta}_1 SIL - \beta = O_p(N^{-1/2} T^{-1/2}) \), we can follow the argument in the proof of Proposition A.1(ii), Lemma A.3 and Lemma A.4(iii) to prove this lemma. Thus, we omitted the details. □

**Proof of Lemma B.2.** Consider (a). With (B.1), we have

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1/2} V_i \| \| T^{-1} (\tilde{H} - H^0 \mathcal{R}) \varepsilon_i \|
\]

\[
\leq N^{-2} T^{-2} \sum_{i=1}^{N} \| T^{-1/2} V_i \| \| \sum_{\ell=1}^{N} \varepsilon_i \langle x_i \rangle (\beta - \tilde{\beta}_1 SIL) (\beta - \tilde{\beta}_1 SIL) \langle x_i \rangle \tilde{H} \tilde{E}^{-1} \|
\]

\[
+ N^{-2} T^{-2} \sum_{i=1}^{N} \| T^{-1/2} V_i \| \| \sum_{\ell=1}^{N} \varepsilon_i \langle x_i \rangle (\beta - \tilde{\beta}) \langle u_i \rangle \tilde{H} \tilde{E}^{-1} \|
\]

\[
+ N^{-2} T^{-2} \sum_{i=1}^{N} \| T^{-1/2} V_i \| \| \sum_{\ell=1}^{N} \varepsilon_i \langle x_i \rangle (\beta - \tilde{\beta}_1 SIL) \langle x_i \rangle \tilde{H} \tilde{E}^{-1} \|
\]

\[
+ N^{-2} T^{-2} \sum_{i=1}^{N} \| T^{-1/2} V_i \| \| \sum_{\ell=1}^{N} \varepsilon_i \langle x_i \rangle (\beta - \tilde{\beta}_1 SIL) \langle x_i \rangle \tilde{H} \tilde{E}^{-1} \|
\]

\[
+ N^{-2} T^{-2} \sum_{i=1}^{N} \| T^{-1/2} V_i \| \| \sum_{\ell=1}^{N} \varepsilon_i \langle x_i \rangle (\beta - \tilde{\beta}_1 SIL) \langle x_i \rangle \tilde{H} \tilde{E}^{-1} \|
\]

The first term is bounded in norm by

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1/2} V_i \| \| T^{-1/2} \varepsilon_i \| \cdot N^{-1} \sum_{\ell=1}^{N} \| T^{-1/2} x_\ell \| \| T^{-1/2} \tilde{H} \| \| \tilde{E}^{-1} \| \| \beta - \tilde{\beta}_1 SIL \|^2 = O_p(\| \beta - \tilde{\beta}_1 SIL \|^2)
\]
thus, the first term is $O_p(\delta_{NT}^{-4})$. Similarly, we can prove that the second and the third terms both are $O_p(\|\beta - \hat{\beta}_{1 junk}\|) = O_p(N^{-1/2}T^{-1/2}).$ Following the argument in the proof of Lemma A.2(b), the last three terms are $O_p(\delta_{NT}^{-2}).$ Consequently, $N^{-1} \sum_{i=1}^{N} \|T^{-1/2}V_i\|\|T^{-1}(\hat{H} - H^0 \mathcal{R})'\| = O_p(\delta_{NT}^{-2}).$

Consider \((b).\) With \((B.1),\) \(N^{-1} \sum_{i=1}^{N} \|\varphi_i^0\|\|T^{-1}V_i'(\hat{H} - H^0 \mathcal{R})\|\) can be decomposed into six terms. The three terms involved of $\beta - \hat{\beta}_{1 junk}$ are $O_p(\|\beta - \hat{\beta}_{1 junk}\|) = O_p(N^{-1/2}T^{-1/2}).$ The remaining three terms can be proved to be $O_p(\delta_{NT}^{-2})$ by following the argument in the proof of Lemma A.2(a), then we have \((b).\)

Following the way in the proof of Lemma A.2(c), we can prove \((c).\) This completes the proof.

**Proof of Lemma B.3.** Since

$$M_{\hat{\beta}}M_{\hat{\beta}} - M_{\beta 0}M_{H^0} = M_{\beta 0}(M_{\hat{\beta}} - M_{H^0}) + (M_{\hat{\beta}} - M_{\beta 0})M_{H^0} + (M_{\hat{\beta}} - M_{\beta 0})(M_{\hat{\beta}} - M_{H^0})$$

we have

$$N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i'M_{\hat{\beta}}M_{\hat{\beta}}u_i - N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i'M_{\beta 0}M_{H^0}e_i$$

$$= N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i'M_{\beta 0}(M_{\hat{\beta}} - M_{H^0})u_i + N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i'(M_{\hat{\beta}} - M_{\beta 0})M_{H^0}u_i$$

$$+ N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i'(M_{\hat{\beta}} - M_{\beta 0})(M_{\hat{\beta}} - M_{H^0})u_i$$

$$= F_1 + F_2 + F_3$$

Now we consider the term $F_1.$ Since $M_{\hat{\beta}} - M_{H^0} = -T^{-1}(\hat{H} - H^0 \mathcal{R})\mathcal{R}'H^0 - T^{-1}H^0 \mathcal{R}(\hat{H} - H^0 \mathcal{R})' - T^{-1}(\hat{H} - H^0 \mathcal{R})(\hat{H} - H^0 \mathcal{R})' - T^{-1}H^0(\mathcal{R} \mathcal{R}' - (T^{-1}H^0H^0)^{-1})H^0,$ we have

$$N^{-1/2}T^{-1/2} \sum_{i=1}^{N} V_i'M_{\beta 0}(M_{\hat{\beta}} - M_{H^0})u_i$$

$$= - N^{-1/2}T^{-3/2} \sum_{i=1}^{N} V_i'(\hat{H} - H^0)(T^{-1}H^0H^0)^{-1}H^0u_i$$

$$- N^{-1/2}T^{-3/2} \sum_{i=1}^{N} V_i'(\hat{H} - H^0)(\mathcal{R} \mathcal{R}' - (T^{-1}H^0H^0)^{-1})H^0u_i$$

$$- N^{-1/2}T^{-3/2} \sum_{i=1}^{N} V_i'H^0(\hat{H} - H^0)'u_i - N^{-1/2}T^{-3/2} \sum_{i=1}^{N} V_i'(\hat{H} - H^0 \mathcal{R})(\hat{H} - H^0 \mathcal{R})'u_i$$

$$- N^{-1/2}T^{-3/2} \sum_{i=1}^{N} V_i'H^0(\mathcal{R} \mathcal{R}' - (T^{-1}H^0H^0)^{-1})H^0u_i$$

$$+ N^{-1/2}T^{-3/2} \sum_{i=1}^{N} V_i'F^0(F^0)^{-1}F^0(\hat{H} - H^0 \mathcal{R})\hat{H}u_i$$

$$+ N^{-1/2}T^{-3/2} \sum_{i=1}^{N} V_i'F^0(F^0)^{-1}F^0H^0(\mathcal{R} \mathcal{R}' - (T^{-1}H^0H^0)^{-1})H^0u_i$$

$$+ N^{-1/2}T^{-3/2} \sum_{i=1}^{N} V_i'F^0(F^0)^{-1}F^0H^0(\mathcal{R} \mathcal{R}' - (T^{-1}H^0H^0)^{-1})H^0u_i$$

$$= F_{1.1} + F_{1.2} + F_{1.3} + F_{1.4} + F_{1.5} + F_{1.6} + F_{1.7} + F_{1.8}$$
We first consider the terms $F_{1.2}$ to $F_{1.8}$. Note that $u_i = H^0\varphi_i^0 + \varepsilon_i$, $F_{1.2}$ is bounded in norm by

$$
N^{1/2}T^{1/2} \cdot N^{-1} \sum_{i=1}^N \|T^{-1/2} V_i^i(\hat{\mathbf{H}} - H^0)\|\|\varphi_i^0\| \cdot \|T^{-1}H^{0'}H^0\|\|\mathcal{R}\mathcal{R}' - (T^{-1}H^{0'}H^0)^{-1}\|\|\mathcal{R}^{-1}\|
$$

$$
+ N^{1/2} \cdot N^{-1} \sum_{i=1}^N \|T^{-1/2} V_i^i(\hat{\mathbf{H}} - H^0)\|\|T^{-1/2}H^{0'}\varepsilon_i\|\cdot \|\mathcal{R}\mathcal{R}' - (T^{-1}H^{0'}H^0)^{-1}\|\|\mathcal{R}^{-1}\|
$$

$$
= O_p(N^{1/2}T^{1/2}\delta_{NT}^{-4})
$$

by Lemma B.1(c), (f) and Lemma B.2 (b) and $N^{-1} \sum_{i=1}^N \|T^{-1}V_i(\hat{\mathbf{H}} - H^0)\|\|T^{-1/2}H^{0'}\varepsilon_i\| = O_p(\delta_{NT}^{-2})$, which can be proved similar to the argument in the proof of Lemma B.2 (b). Similar to the proof of Lemma B.2(a), we have $N^{-1} \sum_{i=1}^N \|T^{-1/2}V_i^i H^0\|\|T^{-1}(\hat{\mathbf{H}} - H^0)\varepsilon_i\|\cdot \|\mathcal{R}\| = O_p(N^{1/2}\delta_{NT}^{-2})$.

by Lemmas B.1(b), (c). $F_{1.4}$ is bounded in norm by

$$
N^{1/2}T^{1/2} \cdot N^{-1} \sum_{i=1}^N \|T^{-1}V_i(\hat{\mathbf{H}} - H^0)\|\|\varphi_i^0\| \cdot \|T^{-1}(\hat{\mathbf{H}} - H^0)\|\|\mathcal{R}\|
$$

$$
+ N^{1/2}T^{1/2} \cdot N^{-1} \sum_{i=1}^N \|T^{-1}V_i(\hat{\mathbf{H}} - H^0)\|\|T^{-1}(\hat{\mathbf{H}} - H^0)\|\|\varepsilon_i\|\cdot \|\mathcal{R}\| = O_p(N^{1/2}T^{1/2}\delta_{NT}^{-4})
$$

by Lemmas B.1(b) and Lemmas B.2 (b), (c). $F_{1.5}$ is bounded in norm by

$$
N^{1/2} \cdot N^{-1} \sum_{i=1}^N \|T^{-1/2}V_i^i H^0\|\|\varphi_i^0\| \cdot \|T^{-1}H^{0'}H^0\|\|\mathcal{R}\mathcal{R}' - (T^{-1}H^{0'}H^0)^{-1}\|
$$

$$
+ N^{1/2}T^{-1/2} \cdot N^{-1} \sum_{i=1}^N \|T^{-1/2}V_i^i H^0\|\|T^{-1/2}H^{0'}\varepsilon_i\|\cdot \|\mathcal{R}\mathcal{R}' - (T^{-1}H^{0'}H^0)^{-1}\| = O_p(N^{1/2}\delta_{NT}^{-2})
$$

by Lemma B.1(f). Similarly, we can prove that $F_{1.6}$, $F_{1.7}$ and $F_{1.8}$ both are $O_p(N^{1/2}\delta_{NT}^{-2})$. Consider the term $F_{1.1}$. With (B.1) and the definition of $\mathcal{R}$, we have

$$
\hat{\mathbf{H}}\mathcal{R}^{-1} - H^0
$$

$$
= -N^{-1}T^{-1} \sum_{i=1}^N X_i(\beta - \hat{\beta}_{1SIV})(\beta - \hat{\beta}_{1SIV})'X_i'\hat{\mathbf{H}}(T^{-1}H^{0'}\hat{\mathbf{H}})^{-1}(Y_i^0)^{-1}
$$

$$
+ N^{-1}T^{-1} \sum_{i=1}^N X_i(\beta - \hat{\beta}_{1SIV})u_i'\hat{\mathbf{H}}(T^{-1}H^{0'}\hat{\mathbf{H}})^{-1}(Y_i^0)^{-1}
$$

$$
+ N^{-1}T^{-1} \sum_{i=1}^N u_i(\beta - \hat{\beta}_{1SIV})'X_i'\hat{\mathbf{H}}(T^{-1}H^{0'}\hat{\mathbf{H}})^{-1}(Y_i^0)^{-1} + N^{-1}T^{-1} \sum_{i=1}^N H^0\varphi_i^0\varepsilon_i'\hat{\mathbf{H}}(T^{-1}H^{0'}\hat{\mathbf{H}})^{-1}(Y_i^0)^{-1}
$$

$$
+ N^{-1} \sum_{i=1}^N \varepsilon_i\varphi_i^0(Y_i^0)^{-1} + N^{-1}T^{-1} \sum_{i=1}^N \varepsilon_i\varepsilon_i'\hat{\mathbf{H}}(T^{-1}H^{0'}\hat{\mathbf{H}})^{-1}(Y_i^0)^{-1}
$$
we can decompose the term \( F_{1.1} \) as follows

\[
- N^{-1/2} T^{-3/2} \sum_{i=1}^{N} V_i' (\hat{H} \mathcal{R}^{-1} - \mathbf{H}^0) (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \mathbf{H}^0 u_i
\]

\[
= - N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i' \varepsilon_h \varphi^0_h (\Upsilon^0_\varphi)^{-1} (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \mathbf{H}^0 u_i
\]

\[
- N^{-3/2} T^{-5/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i' \varepsilon_h \varepsilon_i' \hat{H} (T^{-1} \mathbf{H}^0 \hat{H})^{-1} (\Upsilon^0_\varphi)^{-1} (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \mathbf{H}^0 u_i
\]

\[
- N^{-3/2} T^{-5/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i' \varepsilon_h \hat{H} (T^{-1} \mathbf{H}^0 \hat{H})^{-1} (\Upsilon^0_\varphi)^{-1} (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \mathbf{H}^0 u_i
\]

\[
= F_{1.1.1} + F_{1.1.2} + F_{1.1.3} + F_{1.1.4} + F_{1.1.5} + F_{1.1.6}
\]

We consider the last five terms \( F_{1.1.2} \) to \( F_{1.1.6} \). \( F_{1.1.2} \) is bounded in norm by

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1/2} V_i' \mathbf{H}^0 \| \| T^{-1/2} u_i \| \cdot \| N^{-1/2} T^{-1} \sum_{h=1}^{N} \varphi^0_h \varepsilon_h \hat{H} \| (T^{-1} \mathbf{H}^0 \hat{H})^{-1} \| (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \| (T^{-1/2} \mathbf{H}^0)^{-1} \| (\Upsilon^0_\varphi)^{-1} \|
\]

\[
= O_p(1) \cdot \| N^{-1/2} T^{-1} \sum_{h=1}^{N} \varphi^0_h \varepsilon_h \hat{H} \|
\]

\[
\leq O_p(T^{-1/2}) \cdot \| N^{-1/2} T^{-1} \sum_{h=1}^{N} \varphi^0_h \varepsilon_h \mathbf{H}^0 \| \| \mathcal{R} \| + O_p(1) \cdot \| N^{-1/2} T^{-1/2} \sum_{h=1}^{N} \varphi^0_h \varepsilon_h \| \| \hat{H} - \mathbf{H}^0 \mathcal{R} \| = O_p(\delta^{-1}_{NT})
\]

by Lemmas B.1(a). As \( E (V_i' \varepsilon_h) = 0 \), we can follow the argument in the proof of D.1.4, we can prove that \( F_{1.1.3} = O_p(\delta^{-1}_{NT}) \). \( F_{1.1.4} \) is bounded in norm by

\[
N^{1/2} T^{1/2} \| \beta - \hat{\beta}_{1SIV} \|^2 \cdot N^{-1} \sum_{i=1}^{N} \| T^{-1/2} V_i \| \| T^{-1/2} u_i \|
\]

\[
\times N^{-1} \sum_{h=1}^{N} \| T^{-1/2} \mathbf{X}_h \|^2 \| (\Upsilon^0_\varphi)^{-1} \| (T^{-1} \mathbf{H}^0 \hat{H})^{-1} \| (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \| (T^{-1/2} \hat{H}) \| (T^{-1/2} \mathbf{H}^0) \|
\]

\[
= O_p(N^{-1/2} T^{-1/2})
\]

With definitions of \( u_h \) and \( u_i \), \( F_{1.1.5} \) is bounded in norm by

\[
\| \text{vec} \left( N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i' \mathbf{X}_h (\beta - \hat{\beta}_{1SIV}) \varphi^0_h (\Upsilon^0_\varphi)^{-1} \varphi^0 \right) \|
\]

\[
+ \left\| N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i' \mathbf{X}_h (\beta - \hat{\beta}_{1SIV}) \varepsilon_h \hat{H} (T^{-1} \mathbf{H}^0 \hat{H})^{-1} (\Upsilon^0_\varphi)^{-1} \varphi^0 \right\|
\]

\[
+ \left\| N^{-3/2} T^{-5/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i' \mathbf{X}_h (\beta - \hat{\beta}_{1SIV}) u_h \hat{H} (T^{-1} \mathbf{H}^0 \hat{H})^{-1} (\Upsilon^0_\varphi)^{-1} (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \mathbf{H}^0 \varepsilon_i \right\|
\]

19
The first term is equal to \( \|(N^{-1/2}T^{-1/2} \sum_{i=1}^{N} \Phi_i \otimes V_i) \text{vec}(N^{-1} \sum_{h=1}^{N} \mathbf{x}_h(\beta - \hat{\beta}_{1SIV})\Phi_{i}(\mathbf{Y}_\varphi)^{-1})\| \), then is bounded in norm by

\[
\|N^{-1/2}T^{-1/2} \sum_{i=1}^{N} \Phi_i \otimes V_i\| \cdot N^{-1} \sum_{h=1}^{N} \|T^{-1/2}\mathbf{x}_h\| \|\Phi_i\| \| (\mathbf{Y}_\varphi)^{-1}\| \cdot T^{-1/2}\|\beta - \hat{\beta}_{1SIV}\| = O_p(T^{1/2}\|\beta - \hat{\beta}_{1SIV}\|) = O_p(N^{-1/2})
\]

The second term is bounded in norm by

\[
N^{1/2}\|\beta - \hat{\beta}_{1SIV}\| \cdot N^{-1} \sum_{i=1}^{N} \|T^{-1/2}\mathbf{v}_i\| \|\Phi_i\| \cdot \| (\mathbf{Y}_\varphi)^{-1}\| \cdot \|T^{-1/2}\mathbf{x}_h\| \|T^{-1/2}\epsilon_i\hat{H}\| = O_p(N^{1/2}\|\beta - \hat{\beta}_{1SIV}\|) \cdot N^{-1} \sum_{h=1}^{N} \|T^{-1/2}\mathbf{x}_h\| \|T^{-1/2}\epsilon_i\hat{H}\|
\]

\[
\leq O_p(N^{1/2}\|\beta - \hat{\beta}_{1SIV}\|) \cdot N^{-1} \sum_{h=1}^{N} \|T^{-1/2}\mathbf{x}_h\| \|T^{-1/2}\epsilon_i\mathbf{H}^0\| \|\mathbf{R}\|
\]

\[
+ O_p(N^{1/2}\|\beta - \hat{\beta}_{1SIV}\|) \cdot T^{1/2}N^{-1} \sum_{h=1}^{N} \|T^{-1/2}\mathbf{x}_h\| \|T^{-1/2}\epsilon_i\mathbf{H}^0\| \|T^{-1/2}(\hat{H} - \mathbf{H}^0\mathbf{R})\| = O_p(T^{-1/2}) + O_p(\delta_{NT}^{-1})
\]

by Lemma B.1(a), (c). The third term is bounded in norm by

\[
N^{1/2}\|\beta - \hat{\beta}_{1SIV}\| \cdot N^{-1} \sum_{i=1}^{N} \|T^{-1/2}\mathbf{v}_i\| \|T^{-1/2}\mathbf{H}^0\| \|\epsilon_i\| \cdot N^{-1} \sum_{h=1}^{N} \|T^{-1/2}\mathbf{x}_h\| \|T^{-1/2}\mathbf{u}_h\| \cdot \|T^{-1/2}\hat{H}\| (T^{-1}\mathbf{H}^0\hat{H})^{-1} \| (T^{-1}\mathbf{H}^0\mathbf{H}^0)^{-1}\| = O_p(N^{1/2}\|\beta - \hat{\beta}_{1SIV}\|) = O_p(T^{-1/2})
\]

Then \( \mathbb{F}_{1.5} = O_p(\delta_{NT}^{-1}) \).

With the definitions of \( \mathbf{u}_h \) and \( \mathbf{v}_i \), \( \mathbb{F}_{1.6} \) is bounded in norm by

\[
\|N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} \mathbf{v}_i\mathbf{H}^0\Phi_i(\beta - \hat{\beta}_{1SIV})\mathbf{X}_h\hat{H}(T^{-1}\mathbf{H}^0\hat{H})^{-1}(\mathbf{Y}_\varphi)^{-1}\mathbf{V}_i\| \]

\[
+ \|N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} \mathbf{v}_i\epsilon_h(\beta - \hat{\beta}_{1SIV})\mathbf{X}_h\hat{H}(T^{-1}\mathbf{H}^0\hat{H})^{-1}(\mathbf{Y}_\varphi)^{-1}\epsilon_i\| \]

\[
+ \|N^{-3/2}T^{-5/2} \sum_{i=1}^{N} \sum_{h=1}^{N} \mathbf{v}_i\mathbf{u}_h(\beta - \hat{\beta}_{1SIV})\mathbf{X}_h\hat{H}(T^{-1}\mathbf{H}^0\hat{H})^{-1}(\mathbf{Y}_\varphi)^{-1}(T^{-1}\mathbf{H}^0\mathbf{H}^0)^{-1}\mathbf{H}^0\epsilon_i\| \]

\[
\leq N^{1/2}\|\beta - \hat{\beta}_{1SIV}\| \cdot N^{-1} \sum_{i=1}^{N} \|T^{-1/2}\mathbf{v}_i\| \|\epsilon_i\| \cdot N^{-1} \sum_{h=1}^{N} \|\Phi_i\| \|T^{-1/2}\mathbf{x}_h\| \|T^{-1/2}\hat{H}\| (T^{-1}\mathbf{H}^0\hat{H})^{-1} \| (\mathbf{Y}_\varphi)^{-1}\|
\]

\[
+ N^{-1/2}\|\beta - \hat{\beta}_{1SIV}\| \cdot N^{-2} \sum_{h=1}^{N} \sum_{i=1}^{N} \|\Phi_i\| \|T^{-1/2}\mathbf{v}_i\| \|T^{-1/2}\epsilon_i\mathbf{X}_h\| \|T^{-1/2}\hat{H}\| (T^{-1}\mathbf{H}^0\hat{H})^{-1} \| (\mathbf{Y}_\varphi)^{-1}\|
\]

\[
+ N^{1/2}\|\beta - \hat{\beta}_{1SIV}\| \cdot N^{-1} \sum_{i=1}^{N} \|T^{-1/2}\mathbf{v}_i\| \|T^{-1/2}\mathbf{H}^0\| \|\epsilon_i\| \cdot N^{-1} \sum_{h=1}^{N} \|T^{-1/2}\mathbf{u}_h\| \|T^{-1/2}\mathbf{x}_h\|
\]

\[
\times \|T^{-1/2}\hat{H}\| (T^{-1}\mathbf{H}^0\hat{H})^{-1} \| (\mathbf{Y}_\varphi)^{-1}\| (T^{-1}\mathbf{H}^0\mathbf{H}^0)^{-1}\| = O_p(N^{1/2}\|\beta - \hat{\beta}_{1SIV}\|) = O_p(T^{-1/2})
\]
Combining the above terms and noting that $F_1 = -N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i \varepsilon h \varphi_i^0 (y_i^0)^{-1} (T^{-1} H^0 H^0)^{-1} H^0 u_i = O_p(T^{-1/2})$, we can show that

$$F_1 = -N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{h=1}^{N} V_i \varepsilon h \varphi_i^0 (y_i^0)^{-1} \varphi_i^0 + O_p(\delta^{-1}_N).$$

Consider the term $F_2$. Since $M_F - M_{F0} = -T^{-1} F^0 (F^0 H^0)^{-1} (F R^{-1} F^0)^{-1} F^0$, we have

$$F_2 = \sum_{i=1}^{N} \sum_{h=1}^{N} V_i \varepsilon h (F - F^0 R)^{-1} F^0 (H^0)^{-1} \varphi_i^0.$$

For the terms $F_{2.2}$ to $F_{2.5}$, we can easily show that

$$\|F_{2.2}\| \leq \frac{\sqrt{N}}{N} \sum_{i=1}^{N} \bigg| \frac{V_i F^0}{\sqrt{T}} \bigg| \|\varepsilon h (F - F^0 R)^{-1} F^0 (H^0)^{-1} \varphi_i^0 \| \leq O_p \left( \sqrt{N} \delta^{-4}_N \right),$$

$$\|F_{2.3}\| \leq \frac{N^{1/2}}{N} \sum_{i=1}^{N} \bigg| \frac{V_i \varepsilon h (F - F^0 R)}{\sqrt{T}} \bigg| \|\varepsilon h (F - F^0 R)^{-1} F^0 (H^0)^{-1} \varphi_i^0 \| \leq O_p \left( \sqrt{N} \delta^{-2}_N \right),$$

$$\|F_{2.4}\| \leq \frac{N^{1/2} T^{1/2}}{N} \sum_{i=1}^{N} \|T^{-1} V_i \varepsilon h (F - F^0 R)^{-1} (F - F^0 R) \varphi_i^0 \| \cdot \|T^{-1} (F - F^0 R)^{-1} H^0 \| \|T^{-1} (F - F^0 R)^{-1} H^0 \| \leq O_p \left( N^{1/2} \delta^{-2}_N \right).$$
$$= O_p(N^{1/2}T^{1/2} \delta_{NT}^{-1}),$$

$$\|F_{2.5}\| \leq \sqrt{\frac{N}{T}} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V_i^T \tilde{F}}{\sqrt{T}} \right\| \left\| RR - (T^{-1} F^0 \tilde{F})^{-1} \right\| \left\| \frac{F^0_i \epsilon_i}{\sqrt{T}} \right\|$$

$$+ \sqrt{\frac{N}{T}} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V_i^T \tilde{F}}{\sqrt{T}} \right\| \left\| RR - (T^{-1} F^0 \tilde{F})^{-1} \right\| \left\| \frac{H^0_i \epsilon_i}{\sqrt{T}} \right\|$$

$$= O_p \left( \sqrt{\frac{N}{T}} \delta_{NT}^{-2} \right),$$

by Lemmas B.1 (b), Lemma B.2 (i). Then we have

$$F_2 = O_p(N^{1/2} \delta_{NT}^{-2} + O_p(N^{1/2}T^{1/2} \delta_{NT}^{-4}).$$

Using

$$\tilde{F}^{-1} - F^0 = N^{-1}T^{-1} \sum_{i=1}^{N} F^0_i T \tilde{V}_i F \left( \frac{F_i^0 \tilde{F}}{T} \right)^{-1} (\Upsilon^0)^{-1} + N^{-1} \sum_{i=1}^{N} V_i T \tilde{V}_j (\Upsilon^0)^{-1}$$

$$+ N^{-1} T^{-1} \sum_{i=1}^{N} V_i V_i F \left( \frac{F^0_0 \tilde{F}}{T} \right)^{-1} (\Upsilon^0)^{-1}$$

$$F_{2.1} = -N^{-1/2} T^{-3/2} \sum_{i=1}^{N} V_i^T F \left( \frac{F^0_0 \tilde{F}}{T} \right)^{-1} N^{-1} T^{-1} \sum_{j=1}^{N} (\Upsilon^0)^{-1} \left( \frac{F^0_0 \tilde{F}}{T} \right)^{-1} \tilde{F} V_j \Gamma_j^0 F^0_0 M_0 \epsilon_i$$

$$- N^{-1/2} T^{-3/2} \sum_{i=1}^{N} V_i^T F \left( \frac{F^0_0 \tilde{F}}{T} \right)^{-1} N^{-1} \sum_{j=1}^{N} (\Upsilon^0)^{-1} \Gamma_j^0 V_j M_0 \epsilon_i$$

$$- N^{-1/2} T^{-3/2} \sum_{i=1}^{N} V_i^T F \left( \frac{F^0_0 \tilde{F}}{T} \right)^{-1} N^{-1} T^{-1} \sum_{j=1}^{N} (\Upsilon^0)^{-1} \left( \frac{F^0_0 \tilde{F}}{T} \right)^{-1} \tilde{F} V_j \Gamma_j^0 \epsilon_i$$

$$= F_{2.1.1} + F_{2.1.2} + F_{2.1.3}.$$
By a similar derivation for Lemma A.5, we can show that

\[ h_{\epsilon, q} \leq O \left( \frac{\|F^{0/\epsilon, q}\|}{\sqrt{T}} \right) = O_p \left( \sqrt{\frac{N}{T}} \right) + O_p \left( \sqrt{\frac{N}{T}} \right). \]

By a similar derivation for Lemma A.5, we can show that

\[ F_{2,1,2} = -N^{-1/2}T^{-3/2} \sum_{i=1}^{N} V_i^{T} \left( \frac{F^{0/\epsilon, q}}{T} \right)^{-1} N^{-1} T^{-1} \sum_{j=1}^{N} (Y_j^0)^{-1} \left( \frac{F^{0/\epsilon, q}}{T} \right)^{-1} F^{0/\epsilon, q} (V_j V_j') M^{0/\epsilon, q}, \]

and

\[ F_{2,1,3} = -N^{-1/2}T^{-3/2} \sum_{i=1}^{N} V_i^{T} \left( \frac{F^{0/\epsilon, q}}{T} \right)^{-1} (Y_j^0)^{-1} \left( \frac{F^{0/\epsilon, q}}{T} \right)^{-1} N^{-1} T^{-1} \sum_{j=1}^{N} F^{0/\epsilon, q} (V_j V_j') \epsilon_i \]

and

\[ F_{2,1,3} = O_p \left( \sqrt{\frac{N}{T}} \right) = O_p \left( \sqrt{\frac{N}{T}} \right). \]

hence, \( F_{2,1,3} = O_p \left( \sqrt{\frac{N}{T}} \right) + O_p \left( \sqrt{\frac{N}{T}} \right). \) Therefore, \( F_2 = O_p \left( \sqrt{\frac{N}{T}} \right). \)

Now we consider the term \( F_3. \) We have

\[ F_3 = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T \left( M_{\tilde{F}} - M_{F_0} \right) \left( M_{\tilde{H}} - M_{H_0} \right) u_i \]

and

\[ = \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T \left( \tilde{F} - F_0 R \right) R \left( \tilde{H} - H_0 R \right) R' H_0 u_i, \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} (\tilde{F} - F^0 R) R' F^{0'} T^{-1} H^0 \mathcal{R} (\tilde{H} - H^0 \mathcal{R})' u_i \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} (\tilde{F} - F^0 R) R' F^{0'} T^{-1} (\tilde{H} - H^0 \mathcal{R}) (\tilde{H} - H^0 \mathcal{R})' u_i \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} (\tilde{F} - F^0 R) R' F^{0'} T^{-1} H^0 \left( \mathcal{R} \mathcal{R}' - (T^{-1} H^0 H^0)^{-1} \right) H^0 u_i \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} F^0 R (\tilde{F} - F^0 R)' T^{-1} (\tilde{H} - H^0 \mathcal{R}) \mathcal{R}' H^0 u_i \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} F^0 R (\tilde{F} - F^0 R)' T^{-1} H^0 \mathcal{R} (\tilde{H} - H^0 \mathcal{R})' u_i \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} F^0 R (\tilde{F} - F^0 R)' T^{-1} (\tilde{H} - H^0 \mathcal{R}) (\tilde{H} - H^0 \mathcal{R})' u_i \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} F^0 R (\tilde{F} - F^0 R)' T^{-1} H^0 \left( \mathcal{R} \mathcal{R}' - (T^{-1} H^0 H^0)^{-1} \right) H^0 u_i \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} F^0 \left( \mathcal{R} \mathcal{R}' - (T^{-1} F^{0'} F^0)^{-1} \right) F^{0'} T^{-1} (\tilde{H} - H^0 \mathcal{R}) \mathcal{R}' H^0 u_i \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} F^0 \left( \mathcal{R} \mathcal{R}' - (T^{-1} F^{0'} F^0)^{-1} \right) F^{0'} T^{-1} H^0 \mathcal{R} (\tilde{H} - H^0 \mathcal{R})' u_i \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} F^0 \left( \mathcal{R} \mathcal{R}' - (T^{-1} F^{0'} F^0)^{-1} \right) F^{0'} T^{-1} (\tilde{H} - H^0 \mathcal{R}) (\tilde{H} - H^0 \mathcal{R})' u_i \]
\[ + \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} F^0 \left( \mathcal{R} \mathcal{R}' - (T^{-1} F^{0'} F^0)^{-1} \right) F^{0'} T^{-1} H^0 \left( \mathcal{R} \mathcal{R}' - (T^{-1} H^0 H^0)^{-1} \right) H^0 u_i \]

\[ = F_{3,1,1} + F_{3,1,2} + F_{3,1,3} + F_{3,1,4} + F_{3,2,1} + F_{3,2,2} + F_{3,2,3} + F_{3,2,4} + F_{3,3,1} + F_{3,3,2} + F_{3,3,3} + F_{3,3,4} + F_{3,4,1} + F_{3,4,2} + F_{3,4,3} + F_{3,4,4}. \]
\[
\|F_{3.1,1}\| \leq \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| T^{-1} V'_i (\hat{F} - F^0 R) R' F^0 T^{-1} H^0 (\hat{H} - H^0 \mathcal{R}) \right\| \| \mathcal{R}' \| \left\| T^{-1} F^0 (\hat{H} - H^0 \mathcal{R}) \right\| \| \mathcal{R}' \| \left\| H^0 u_i \right\| \left\| H^0 u_i \right\|
\]
\[= O_p \left( \sqrt{NT \delta_{NT}^{-3}} \right).\]

by Lemma A.2 (b), Lemma B.1 (b) and Lemma B.2 (d).

\[
\|F_{3.1,2}\| = \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V'_i T^{-1} (\hat{F} - F^0 R) R' F^0 T^{-1} H^0 (\hat{H} - H^0 \mathcal{R})' u_i \right\|
\]
\[\leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V'_i T^{-1} (\hat{F} - F^0 R) R' F^0 T^{-1} H^0 (\hat{H} - H^0 \mathcal{R})' \varphi_i \right\|
\[+ \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V'_i T^{-1} (\hat{F} - F^0 R) R' F^0 T^{-1} H^0 (\hat{H} - H^0 \mathcal{R})' \varepsilon_i \right\|
\]
\[\leq \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V'_i (\hat{F} - F^0 R)}{T} \right\| \left\| \frac{F^0 H^0}{T} \right\| \left\| \mathcal{R}' \right\| \left\| (\hat{H} - H^0 \mathcal{R})' \right\| \left\| \varphi_i \right\|
\]
\[+ \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V'_i (\hat{F} - F^0 R)}{T} \right\| \left\| \frac{F^0 H^0}{T} \right\| \left\| \mathcal{R}' \right\| \left\| (\hat{H} - H^0 \mathcal{R})' \right\| \left\| \varepsilon_i \right\|
\]
\[= O_p \left( \sqrt{NT \delta_{NT}^{-3}} \right).\]

by Lemma A.2 (b), Lemma B.1 (b) and Lemma B.2 (c).

\[
\|F_{3.1,3}\| = \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V'_i T^{-1} (\hat{F} - F^0 R) R' F^0 T^{-1} (\hat{H} - H^0 \mathcal{R}) (\hat{H} - H^0 \mathcal{R})' u_i \right\|
\]
\[\leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V'_i T^{-1} (\hat{F} - F^0 R) R' F^0 T^{-1} (\hat{H} - H^0 \mathcal{R}) (\hat{H} - H^0 \mathcal{R})' \varphi_i \right\|
\[+ \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V'_i T^{-1} (\hat{F} - F^0 R) R' F^0 T^{-1} (\hat{H} - H^0 \mathcal{R}) (\hat{H} - H^0 \mathcal{R})' \varepsilon_i \right\|
\]
\[\leq \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V'_i (\hat{F} - F^0 R)}{T} \right\| \left\| \mathcal{R}' \right\| \left\| \frac{F^0 (\hat{H} - H^0 \mathcal{R})}{T} \right\| \left\| (\hat{H} - H^0 \mathcal{R})' \right\| \left\| \varphi_i \right\|
\]
\[+ \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V'_i (\hat{F} - F^0 R)}{T} \right\| \left\| \mathcal{R}' \right\| \left\| \frac{F^0 (\hat{H} - H^0 \mathcal{R})}{T} \right\| \left\| (\hat{H} - H^0 \mathcal{R})' \right\| \left\| \varepsilon_i \right\|
\]
\[= O_p \left( \sqrt{NT \delta_{NT}^{-3}} \right).\]

by Lemma A.2 (b), Lemma B.1 (b) and Lemma B.2 (c).

\[
\|F_{3.1,4}\| = \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V'_i T^{-1} (\hat{F} - F^0 R) R' F^0 T^{-1} H^0 \left( \mathcal{R} \mathcal{R}' - (T^{-1} H^0 H^0)^{-1} \right) H^0 u_i \right\|
\]

25
\[
\leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} (\tilde{F} - F^0 R) R^0 T^{-1} H^0 \left( R R' - (T^{-1} H^0 H^0)^{-1} \right) H^0 H^0 \varphi_i^0 \right\|
\]
\[
+ \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} (\tilde{F} - F^0 R) R^0 T^{-1} H^0 \left( R R' - (T^{-1} H^0 H^0)^{-1} \right) H^0 \varepsilon_i \right\|
\]
\[
\leq \sqrt{NT} \left\| \frac{1}{N} \sum_{i=1}^{N} \frac{V_i^T (F - F^0 R)}{T} \right\| \left\| R' \right\| \left\| F^0 H^0 \right\| \left\| R R' - (T^{-1} H^0 H^0)^{-1} \right\| \left\| H^0 H^0 \right\| \left\| \varphi_i^0 \right\|
\]
\[
+ \sqrt{N} \left( \frac{N}{T} \sum_{i=1}^{N} \right) \left\| \frac{V_i^T (F - F^0 R)}{T} \right\| \left\| R' \right\| \left\| F^0 H^0 \right\| \left\| R R' - (T^{-1} H^0 H^0)^{-1} \right\| \left\| H^0 \varepsilon_i \right\|
\]
\[
= O_p \left( \frac{\sqrt{N}}{\delta N T} \right)
\]

by Lemma A.2 (b), Lemma B.1 (d) and Lemma B.2 (d).

\[
\left\| F_{3.2.1} \right\| = \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} F^0 R (\tilde{F} - F^0 R)' T^{-1} (\tilde{H} - H^0 R) R' H^0 u_i \right\|
\]
\[
\leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} F^0 R (\tilde{F} - F^0 R)' T^{-1} (\tilde{H} - H^0 R) R' H^0 \varphi_i^0 \right\|
\]
\[
+ \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} F^0 R (\tilde{F} - F^0 R)' T^{-1} (\tilde{H} - H^0 R) R' H^0 \varepsilon_i \right\|
\]
\[
\leq \sqrt{N} \left\| \frac{1}{\sqrt{T}} \sum_{i=1}^{N} V_i^T F^0 \right\| \left\| R \right\| \left\| (\tilde{F} - F^0 R) / \sqrt{T} \right\| \left\| (\tilde{H} - H^0 R) / \sqrt{T} \right\| \left\| R' \right\| \left\| H^0 H^0 \right\| \left\| \varphi_i^0 \right\|
\]
\[
+ \sqrt{N} \left( \frac{1}{\sqrt{T}} \sum_{i=1}^{N} \right) \left\| \frac{V_i^T F^0}{\sqrt{T}} \right\| \left\| R \right\| \left\| (\tilde{F} - F^0 R) / \sqrt{T} \right\| \left\| (\tilde{H} - H^0 R) / \sqrt{T} \right\| \left\| R' \right\| \left\| H^0 \varepsilon_i \right\|
\]
\[
= O_p \left( \frac{\sqrt{N}}{\delta N T} \right)
\]

by Lemma A.1 (b), Lemma A.2 (d) and Lemma B.1 (b).
\[ \|F_{3,2.3}\| = \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} F^0 \mathbf{R} (\hat{F} - F^0 \mathbf{R})^T T^{-1} (\hat{H} - H^0 \mathbf{R}) (\hat{H} - H^0 \mathbf{R})' H^0 \varphi_i \right\| \\
+ \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} F^0 \mathbf{R} (\hat{F} - F^0 \mathbf{R})^T T^{-1} (\hat{H} - H^0 \mathbf{R}) (\hat{H} - H^0 \mathbf{R})' \varepsilon_i \right\| \\
\leq \sqrt{N} \frac{1}{N} \sum_{i=1}^{N} \left\| V_i^T F^0 \right\| \left\| \mathbf{R} \right\| \left\| (\hat{F} - F^0 \mathbf{R})' \sqrt{T} \right\| \left\| (\hat{H} - H^0 \mathbf{R})' \sqrt{T} \right\| \left\| \frac{H^0 \mathbf{R}}{T} \right\| \| \varphi_i \| \\
+ \sqrt{N} \frac{1}{N} \sum_{i=1}^{N} \left\| V_i^T F^0 \right\| \left\| \mathbf{R} \right\| \left\| (\hat{F} - F^0 \mathbf{R})' \sqrt{T} \right\| \left\| (\hat{H} - H^0 \mathbf{R})' \sqrt{T} \right\| \left\| \frac{H^0 \mathbf{R}}{T} \right\| \| \varphi_i \| \\
= O_p \left( \sqrt{NT} \right) \\
\text{by Lemma B.1 (b) and Lemma B.2 (e)}. \]

\[ \|F_{3,2.4}\| \leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} F^0 \mathbf{R} (\hat{F} - F^0 \mathbf{R})^T T^{-1} \mathbf{H}^0 \left( \mathcal{R} \mathcal{R}' - (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \right) \mathbf{H}^0 \mathbf{H}^0 \varphi_i \right\| \\
+ \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} F^0 \mathbf{R} (\hat{F} - F^0 \mathbf{R})^T T^{-1} \mathbf{H}^0 \left( \mathcal{R} \mathcal{R}' - (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \right) \mathbf{H}^0 \varepsilon_i \right\| \\
\leq \sqrt{N} \frac{1}{N} \sum_{i=1}^{N} \left\| V_i^T F^0 \right\| \left\| \mathbf{R} \right\| \left\| (\hat{F} - F^0 \mathbf{R})' \frac{\mathbf{H}^0}{T} \right\| \left\| \mathcal{R} \mathcal{R}' - (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \right\| \left\| \frac{\mathbf{H}^0 \mathbf{H}^0}{T} \right\| \| \varphi_i \| \\
+ \sqrt{N} \frac{1}{N} \sum_{i=1}^{N} \left\| V_i^T F^0 \right\| \left\| \mathbf{R} \right\| \left\| (\hat{F} - F^0 \mathbf{R})' \frac{\mathbf{H}^0}{T} \right\| \left\| \mathcal{R} \mathcal{R}' - (T^{-1} \mathbf{H}^0 \mathbf{H}^0)^{-1} \right\| \left\| \frac{\mathbf{H}^0 \varepsilon_i}{\sqrt{T}} \right\| \\
= O_p \left( \sqrt{NT} \right) \\
\text{by Lemma A.1 (a), Lemma B.1 (a) (b) and Lemma B.2 (e)}. \]

\[ \|F_{3,3.1}\| \leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} (\hat{F} - F^0 \mathbf{R}) (\hat{F} - F^0 \mathbf{R})^T T^{-1} (\hat{H} - H^0 \mathbf{R}) \mathcal{R} H^0 \mathbf{H}^0 \varphi_i \right\| \\
+ \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^T T^{-1} (\hat{F} - F^0 \mathbf{R}) (\hat{F} - F^0 \mathbf{R})^T T^{-1} (\hat{H} - H^0 \mathbf{R}) \mathcal{R} H^0 \mathbf{H}^0 \varepsilon_i \right\| \\
\leq \sqrt{N} \frac{1}{N} \sum_{i=1}^{N} \left\| V_i (\hat{F} - F^0 \mathbf{R})' \sqrt{T} \right\| \left\| (\hat{F} - F^0 \mathbf{R})' \sqrt{T} \right\| \left\| (\hat{H} - H^0 \mathbf{R})' \sqrt{T} \right\| \left\| \mathcal{R} \right\| \left\| \frac{H^0 \mathbf{H}^0}{T} \right\| \| \varphi_i \| \\
+ \sqrt{N} \frac{1}{N} \sum_{i=1}^{N} \left\| V_i (\hat{F} - F^0 \mathbf{R})' \sqrt{T} \right\| \left\| (\hat{F} - F^0 \mathbf{R})' \sqrt{T} \right\| \left\| (\hat{H} - H^0 \mathbf{R})' \sqrt{T} \right\| \left\| \mathcal{R} \right\| \left\| \frac{H^0 \varepsilon_i}{\sqrt{T}} \right\| \\
= O_p \left( \sqrt{NT} \right) \]
by Lemma A.1 (a), Lemma A.2 (b), Lemma B.1 (a) (b) and Lemma B.2 (d).

\[
\|F_{3.3.2}\| \leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} (\hat{F} - F^0 R)(\hat{F} - F^0 R)' T^{-1} H^0 R (\hat{H} - H^0 R)' H^0 \varphi_i \right\| + \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} (\hat{F} - F^0 R)(\hat{F} - F^0 R)' T^{-1} H^0 R (\hat{H} - H^0 R)' \epsilon_i \right\|
\]

\[
\leq \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V_i (\hat{F} - F^0 R)}{T} \right\| \left\| \frac{(\hat{F} - F^0 R)' H^0}{T} \right\| \left\| R \right\| \left\| \frac{\hat{H} - H^0 R}{T} \right\| \left\| H^0 \varphi_i \right\|
\]

\[
+ \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V_i (\hat{F} - F^0 R)}{T} \right\| \left\| \frac{(\hat{F} - F^0 R)' H^0}{T} \right\| \left\| R \right\| \left\| \frac{\hat{H} - H^0 R}{T} \right\| \left\| \epsilon_i \right\|
\]

= \text{O}_p \left( \frac{\sqrt{NT}}{\delta_{NT}} \right)

by Lemma A.2 (b), Lemma B.1 (b) and Lemma B.2 (c).

\[
\|F_{3.3.3}\| \leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} (\hat{F} - F^0 R)(\hat{F} - F^0 R)' T^{-1} (\hat{H} - H^0 R)(\hat{H} - H^0 R)' H^0 \varphi_i \right\| + \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} (\hat{F} - F^0 R)(\hat{F} - F^0 R)' T^{-1} (\hat{H} - H^0 R)(\hat{H} - H^0 R)' \epsilon_i \right\|
\]

\[
\leq \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V_i (\hat{F} - F^0 R)}{T} \right\| \left\| \frac{(\hat{F} - F^0 R)' H^0}{T} \right\| \left\| \frac{\hat{H} - H^0 R}{T} \right\| \left\| \hat{H} - H^0 R\right\| \left\| \varphi_i \right\|
\]

\[
+ \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V_i (\hat{F} - F^0 R)}{T} \right\| \left\| \frac{(\hat{F} - F^0 R)' H^0}{T} \right\| \left\| \frac{\hat{H} - H^0 R}{T} \right\| \left\| \epsilon_i \right\|
\]

= \text{O}_p \left( \frac{\sqrt{NT}}{\delta_{NT}} \right)

by Lemma A.1 (a), Lemma A.2 (b), Lemma B.1 (a) (b) and Lemma B.2 (c).

\[
\|F_{3.3.4}\| \leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} (\hat{F} - F^0 R)(\hat{F} - F^0 R)' T^{-1} H^0 \left( R R' - (T^{-1} H^0 H^0)^{-1} \right) H^0 H^0 \varphi_i \right\| + \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} (\hat{F} - F^0 R)(\hat{F} - F^0 R)' T^{-1} H^0 \left( R R' - (T^{-1} H^0 H^0)^{-1} \right) H^0 \epsilon_i \right\|
\]

\[
\leq \sqrt{NT} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V_i (\hat{F} - F^0 R)}{T} \right\| \left\| \frac{(\hat{F} - F^0 R)' H^0}{T} \right\| \left\| \frac{\hat{R} R' - (T^{-1} H^0 H^0)^{-1}}{T} \right\| \left\| H^0 H^0 \varphi_i \right\|
\]

\[
+ \sqrt{N} \frac{1}{N} \sum_{i=1}^{N} \left\| \frac{V_i (\hat{F} - F^0 R)}{T} \right\| \left\| \frac{(\hat{F} - F^0 R)' H^0}{T} \right\| \left\| \hat{R} R' - (T^{-1} H^0 H^0)^{-1} \right\| \left\| \frac{H^0 \epsilon_i}{T} \right\|
\]

= \text{O}_p \left( \frac{\sqrt{NT}}{\delta_{NT}} \right)

by Lemma A.2 (b), Lemma B.1 (b) (d) and Lemma B.2 (d).

\[
\|F_{3.4.1}\| \leq \left\| \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i T^{-1} F^0 \left( R R' - (T^{-1} F^0 F^0)^{-1} \right) F^0 T^{-1} (\hat{H} - H^0 R) R H^0 H^0 \varphi_i \right\|
\]

28
by Lemma A.1 (d), Lemma B.1 (b) and Lemma B.2 (e).

\[
\|F_{3.4.2}\| \leq \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^* T^{-1} F_0 \left( R R' - (T^{-1} F_0^0 F_0^{-1}) \right) F_0^0 T^{-1} (\hat{H} - H^0) R' H^0 \varphi_i
\]

by Lemma A.1 (d), Lemma B.1 (b) and Lemma B.2 (e).

\[
\|F_{3.4.3}\| \leq \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} V_i^* T^{-1} F_0 \left( R R' - (T^{-1} F_0^0 F_0^{-1}) \right) F_0^0 T^{-1} (\hat{H} - H^0) R' (\hat{H} - H^0) H^0 \varphi_i
\]

by Lemma A.1 (d), Lemma B.1 (b) and Lemma B.3 (e).
by Lemma A.1 (d), Lemma B.1 (d) and Lemma B.2 (f).

Combining the above terms from $F_{3.1.1}$ to $F_{3.4.4}$, we derive that $F_3 = O_p \left( \sqrt{NT\delta_{NT}} \right)$. This completes the proof. □

**Proof of Lemma B.4.** The proof can be completed following the argument in the proof of Lemma A.3. □

**Proof of Lemma B.5.** Follow the way of the proof of Lemma A.5, we can show that

$$N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{00} (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \tilde{F}^0 \varepsilon_i \sum_{\ell} \Gamma_{i}^{00} \cdot (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \cdot F^0 \varepsilon_i$$

The first term is equal to

$$\text{vec} (N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{00} (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \tilde{F}^0 \varepsilon_i)$$

$$= \text{vec} (N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{00} \cdot (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \cdot F^0 \varepsilon_i)$$

$$- \text{vec} (N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{00} \cdot (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \cdot F^0 \varepsilon_i)$$

$$- \text{vec} (N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{00} \cdot (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \cdot F^0 \varepsilon_i)$$

$$+ \text{vec} (N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_{i}^{00} \cdot (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \cdot F^0 \varepsilon_i)$$

$$= N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \cdot \text{vec} ((\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1})$$

$$- N^{-1/2}T^{-3/2} \sum_{i=1}^{N} (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \cdot \text{vec} ((\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1})$$

$$- N^{-1/2}T^{-3/2} \sum_{i=1}^{N} (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \cdot \text{vec} ((\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1})$$

$$+ N^{-1/2}T^{-3/2} \sum_{i=1}^{N} (\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1} \cdot \text{vec} ((\gamma^0)^{-1} (T^{-1} F^0 F^0)^{-1})$$

$$\times (T^{-1} F^0 F^0)^{-1} (T^{-1} F^0 F^0) (T^{-1} H^0 H^0)^{-1}$$
\[
= N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} (\varepsilon_i^r (V_i V_i^T) F_0^0) \otimes (\Gamma_i^0) \cdot O_p(1) + N^{-1/2} T^{-3/2} \sum_{i=1}^{N} (\varepsilon_i^r F_0^0) \otimes (\Gamma_i^0) \cdot O_p(1)
\]

\[
+ N^{-1/2} T^{-3/2} \sum_{i=1}^{N} (\varepsilon_i H^0) \otimes (\Gamma_i^0) \cdot O_p(1)
\]

\[
= N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} \sum_{t=1}^{T} \mathbb{E} (\varepsilon_i^r V_i V_{i t}) \varepsilon_{i t} f_t^0 \otimes \Gamma_i^0 \cdot O_p(1) + N^{-1/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} (\varepsilon_t f_t^0) \otimes \Gamma_i^0 \cdot O_p(1)
\]

\[
\leq O_p(T^{-1/2})
\]

because

\[
\mathbb{E} \| N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{s=1}^{T} (\varepsilon_i^r V_i) \varepsilon_{i t} f_t^0 \otimes \Gamma_i^0 \|^2
\]

\[
\leq k T^{-1} \cdot N^{-2} \sum_{i=1}^{N} \sum_{s=1}^{T} \sum_{t=1}^{T} \sum_{t=1}^{T} \tilde{s}_{i t} \cdot T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \tilde{s}_{i t} \cdot T^{-1} \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{s}_{i t} \leq C T^{-1}
\]

by Assumption A2 and B2, C and D, and

\[
T^{-1} \cdot N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \varepsilon_{i t} f_t^0 \otimes \Gamma_i^0 = O_p(T^{-1}),
\]

Thus, we derive that

\[
N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{t=1}^{T} \Gamma_i^0 (\Upsilon_0)^{-1} (T^{-1} F_0 V_i^T) - T^{-1} F_0 V_i^T V_i M_0 M_0^* u_i = O_p(T^{1/2} \delta^{-2}) + O_p(N^{1/2} T^{1/2} \delta^{-1} N_T)
\]

This completes the proof. \(\square\)

**Proof of Lemma B.6.** We can follow the way of the proof of Lemma A.4 to show that

\[
- N^{-3/2} T^{-1/2} \sum_{i=1}^{N} \Gamma_i^0 (\Upsilon_0)^{-1} \Gamma_i^0 V_i M_0 M_0^* u_i
\]

\[
= -N^{-3/2} T^{-1/2} \sum_{i=1}^{N} \Gamma_i^0 (\Upsilon_0)^{-1} \Gamma_i^0 V_i M_0 M_0^* u_i
\]

\[
+ N^{-1/2} T^{1/2} \cdot N^{-1} \sum_{i=1}^{N} \Gamma_i^0 (\Upsilon_0)^{-1} (N^{-1} T^{-1} \sum_{t=1}^{T} \sum_{h=1}^{N} \Gamma_i^0 V_i V_h (\Gamma_h^0) (\Upsilon_0)^{-1} (F_0^0 F_0^0) - T^{-1} F_0^0 M_0^* u_i
\]

\[
+ O_p(T^{1/2} \delta^{-2})
\]

Consider the second term on the right hand side in (B.2). Denoting

\[
Q = (\Upsilon_0)^{-1} (N^{-1} T^{-1} \sum_{t=1}^{T} \sum_{h=1}^{N} \Gamma_i^0 V_i V_h (\Gamma_h^0) (\Upsilon_0)^{-1} (F_0^0 F_0^0) / T)^{-1},
\]

31
by a similar derivation of Lemma B.5, we can show that

\[ N^{-1} \cdot N^{-1/2} T^{-1/2} \sum_{i=1}^{N} \Gamma_i^{0r} QF^{0r} M_{H^0} u_i = O_p(N^{-1}) \]

Then, we consider the first term on the right hand side in (B.2). Noting that \( M_{H^0} u_i = M_{H^0} \varepsilon_i \) and \( M_{H^0} - M_H = T^{-1} \hat{H}^T - P_{H^0} \), We can derive that

\[
\begin{align*}
- N^{-3/2} T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^{0r} (T^0)^{-1} \Gamma_i^{0r} V_i^{0r} M_{F_0} M_{H^0} u_i \\
= N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^{0r} (T^0)^{-1} \Gamma_i^{0r} V_i^{0r} M_{F_0} \hat{H}^T u_i - N^{-3/2} T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^{0r} (T^0)^{-1} \Gamma_i^{0r} V_i^{0r} M_{F_0} P_{H^0} u_i \\
= N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^{0r} (T^0)^{-1} \Gamma_i^{0r} V_i^{0r} M_{F_0} (\hat{H} \mathcal{R}^{-1} - H^0)(T^{-1} H^0 H^0)^{-1} H^0 u_i \\
+ N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^{0r} (T^0)^{-1} \Gamma_i^{0r} V_i^{0r} M_{F_0} (\hat{H} \mathcal{R}^{-1} - H^0)(\mathcal{R} \mathcal{R}' - (T^{-1} H^0 H^0)^{-1} H^0) u_i \\
+ N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^{0r} (T^0)^{-1} \Gamma_i^{0r} V_i^{0r} M_{F_0} H^0 \mathcal{R}(\hat{H} - H^0) \mathcal{R}' u_i \\
+ N^{-3/2} T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \Gamma_i^{0r} (T^0)^{-1} \Gamma_i^{0r} V_i^{0r} M_{F_0} H(\mathcal{R} \mathcal{R}' - (T^{-1} H^0 H^0)^{-1}) H^0 u_i \\
= G_1 + G_2 + G_3 + G_4 + G_5
\end{align*}
\]

With the facts that

\[ N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_{i}^{0r} V_{i}^{0r} M_{F_0} H^0 = N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_{i}^{0r} V_{i}^{0r} H^0 = N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_{i}^{0r} V_{i}^{0r} F(0) F^0 = O_p(1) \]

and

\[
\begin{align*}
\|N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_{i}^{0r} V_{i}^{0r} M_{F_0} (\hat{H} - H^0) \mathcal{R}\| & \\
\leq \|N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_{i}^{0r} V_{i}^{0r} (\hat{H} - H^0) \mathcal{R}\| + \|N^{-1/2} T^{-1/2} \sum_{\ell=1}^{N} \Gamma_{i}^{0r} V_{i}^{0r} F^0\| \|T^{-1} F(0) F^0 = O_p(1)\| \|T^{-1} F(0) (\hat{H} - H^0) \mathcal{R}\| \\
= O_p(N^{-1/2} T^{1/2}) + O_p(1)\|T^{-1} F(0) (\hat{H} - H^0) \mathcal{R}\| \\
= O_p(1)\|T^{-1/2} \delta_{NT}^{-1}\| + O_p(T^{1/2} \delta_{NT}^{-1})
\end{align*}
\]

by Lemma B.1(b). We can closely follow the arguments in the proofs of \( E_{1,2}, E_{2,3} \) and \( E_{4} \), to show that \( E_2 = O_p(N^{-1/2} T^{1/2} \delta_{NT}^{-2}) + O_p(T^{1/2} \delta_{NT}^{-4}) \), \( E_3 = O_p(\delta_{NT}^{-2}) \), \( E_4 = O_p(N^{-1/2} T^{1/2} \delta_{NT}^{-2}) + O_p(T^{1/2} \delta_{NT}^{-4}) \) and \( E_5 = O_p(\delta_{NT}^{-4}) \).
With (B.1), $G_1$ is decomposed as follows

$$
N^{-5/2}T^{-5/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma_i^{0}(Y^{0})^{-1} \Gamma_{\ell}^{0} V_{\ell} M_{F_{\ell}} H_{\ell}^{0} \varphi_{h}^{0} \epsilon_{h}^{0} H_{\ell}^{0} R(T^{-1} H^{0} H^{0})^{-1} (Y_{\varphi}^{0})^{-1} (T^{-1} H^{0} H^{0})^{-1} H^{0} u_{i}
$$

$$
+ N^{-5/2}T^{-5/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma_i^{0}(Y^{0})^{-1} \Gamma_{\ell}^{0} V_{\ell} M_{F_{\ell}} H_{\ell}^{0} \varphi_{h}^{0} \epsilon_{h}^{0} (\hat{H} - H^{0} R)(T^{-1} H^{0} H^{0})^{-1} (Y_{\varphi}^{0})^{-1} (T^{-1} H^{0} H^{0})^{-1} H^{0} u_{i}
$$

$$
+ N^{-5/2}T^{-5/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma_i^{0}(Y^{0})^{-1} \Gamma_{\ell}^{0} V_{\ell} M_{F_{\ell}} \varphi_{h}^{0} \epsilon_{h}^{0} (\hat{H} - H^{0} R)(T^{-1} H^{0} H^{0})^{-1} (Y_{\varphi}^{0})^{-1} (T^{-1} H^{0} H^{0})^{-1} H^{0} u_{i}
$$

$$
+ N^{-5/2}T^{-5/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma_i^{0}(Y^{0})^{-1} \Gamma_{\ell}^{0} V_{\ell} M_{F_{\ell}} X_{h}(\beta - \tilde{\beta}_{1SIV})(\beta - \tilde{\beta}_{1SIV})' X_{h}^{0} H^{0}
$$

$$
\times \left(\frac{H^{0} H^{0}}{T}\right)^{-1} (Y_{\varphi}^{0})^{-1} \left(\frac{H^{0} H^{0}}{T}\right)^{-1} H^{0} u_{i}
$$

$$
+ N^{-5/2}T^{-5/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma_i^{0}(Y^{0})^{-1} \Gamma_{\ell}^{0} V_{\ell} M_{F_{\ell}} X_{h}(\beta - \tilde{\beta}_{1SIV}) u_{h} H^{0} H^{0} (T^{-1} H^{0} H^{0})^{-1} (Y_{\varphi}^{0})^{-1} (T^{-1} H^{0} H^{0})^{-1} H^{0} u_{i}
$$

$$
+ N^{-5/2}T^{-5/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma_i^{0}(Y^{0})^{-1} \Gamma_{\ell}^{0} V_{\ell} M_{F_{\ell}} u_{h}(\beta - \tilde{\beta}_{1SIV}) X_{h}^{0} H^{0} H^{0} (T^{-1} H^{0} H^{0})^{-1} (Y_{\varphi}^{0})^{-1} (T^{-1} H^{0} H^{0})^{-1} H^{0} u_{i}
$$

$$
= G_{1.1} + G_{1.2} + G_{1.3} + G_{1.4} + G_{1.5} + G_{1.6} + G_{1.7}
$$

Following the argument in the proof of $\mathbb{B}_{1.1}$, we can show that $G_{1.1} + G_{1.2} = O_{p}(T^{1/2} \delta_{NT}^{-2}) + O_{p}(N^{-1/2})$. In a similar manner, it can be shown that

$$
G_{1.4} = N^{-3/2}T^{-3/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma_i^{0}(Y^{0})^{-1} \Gamma_{\ell}^{0} V_{\ell} H_{\ell}^{0} (T^{-1} H^{0} H^{0})^{-1} (Y_{\varphi}^{0})^{-1} \varphi_{i}^{0}
$$

$$
+ O_{p}(T^{1/2} \delta_{NT}^{-2}) + O_{p}(T^{-1/2}) + O_{p}(N^{-1/2}),
$$

the first term of which is $O_{p}(T^{-1/2})$. As $\| M_{F_{\ell}} X_{h} \| \leq \| X_{h} \|$, $G_{1.6}$ is bounded in norm by

$$
T^{1/2} \| \beta - \tilde{\beta}_{1SIV} \| \cdot N \sum_{i=1}^{N} \| \Gamma_i^{0} \| \| T^{-1/2} u_{i} \| \cdot N \sum_{h=1}^{N} \| T^{-1/2} X_{h} \| \| T^{-1/2} u_{h} \|
$$

$$
\cdot \| N^{-1/2}T^{-1/2} \sum_{\ell=1}^{N} \Gamma_{\ell}^{0} V_{\ell} \| \| T^{-1/2} H_{\ell} \| \| T^{-1/2} H_{\ell} \| \| (Y^{0})^{-1} \| \| (T^{-1} H^{0} H^{0})^{-1} \| \| (Y_{\varphi}^{0})^{-1} \| \| (T^{-1} H^{0} H^{0})^{-1} \|
$$

$$
= O_{p}(T^{1/2} \| \beta - \tilde{\beta}_{1SIV} \|) = O_{p}(N^{-1/2}).
$$

A similar derivation yields that $G_{1.7} = O_{p}(T^{1/2} \| \beta - \tilde{\beta}_{1SIV} \|) = O_{p}(N^{-1/2})$. Also $G_{1.5} = O_{p}(N^{1/2}T^{1/2} \| \beta - \tilde{\beta}_{1SIV} \|^{2}) = O_{p}(N^{-1/2}T^{-1/2})$. Finally

$$
G_{1.3} = N^{-5/2}T^{-1/2} \sum_{i=1}^{N} \sum_{\ell=1}^{N} \sum_{h=1}^{N} \Gamma_i^{0}(Y^{0})^{-1} \Gamma_{\ell}^{0} V_{\ell} \varphi_{h}^{0} \epsilon_{h}^{0} (Y_{\varphi}^{0})^{-1} \varphi_{i}^{0} + O_{p}(T^{-1/2}),
$$

33
which is bounded in norm by
\[ N^{-1/2} \sum_{i=1}^{N} \||\Gamma_i^0||\|\varphi_i^0\| \cdot \||\nu(0)^{-1}||\|\nu(0)^{-1}\| \cdot ||N^{-1}T^{-1/2} \sum_{t=1}^{N} \Gamma_0^0 V_t M P^0 \varepsilon_h \varphi_h^0|| \]

\[ = N^{-1/2} O_p(1) \cdot ||N^{-1}T^{-1/2} \sum_{t=1}^{N} \sum_{h=1}^{N} \Gamma_0^0 V_t \varepsilon_h \varphi_h^0|| \]

\[ = O_p(N^{-1/2}) \]

because
\[ E\left( E\left( ||N^{-1}T^{-1/2} \sum_{t=1}^{N} \sum_{h=1}^{N} \sum_{t=1}^{T} \text{vec}(\Gamma_0^0 V_{t}\varepsilon_{t} \varphi_h^0||^2) \right| \{\Gamma_i^0, \varphi_i^0\}_{i=1}^{N}\right) \]

\[ = E\left( \text{tr}(N^{-2}T^{-1} \sum_{t_{1}=1}^{N} \sum_{t_{2}=1}^{N} \sum_{h_{1}=1}^{N} \sum_{h_{2}=1}^{N} \sum_{t_{1}=1}^{T} \sum_{t_{2}=1}^{T} (\varphi_{h_{1}}^0 \otimes \Gamma_{t_{1}}^0) E(V_{t_{1}} \varepsilon_{t_{1}} \varepsilon_{h_{2}} \varepsilon_{t_{2}}) E(\varepsilon_{h_{1}} \varepsilon_{t_{1}} \varepsilon_{h_{2}} \varepsilon_{t_{2}}) (\varphi_{h_{2}}^0 \otimes \Gamma_{t_{2}}^0)) \right) \]

\[ = CN^{-2}T^{-1} \sum_{t_{1}=1}^{N} \sum_{t_{2}=1}^{N} \sum_{h_{1}=1}^{N} \sum_{h_{2}=1}^{N} \sum_{t_{1}=1}^{T} \sum_{t_{2}=1}^{T} \|E(V_{t_{1}} \varepsilon_{t_{1}} V_{t_{2}}^t \varepsilon_{h_{2}} \varepsilon_{t_{2}})\| \|E(\varepsilon_{h_{1}} \varepsilon_{t_{1}} \varepsilon_{h_{2}} \varepsilon_{t_{2}})\| \|\varphi_{h_{1}}^0\| ||\varphi_{h_{2}}^0\| \|\Gamma_{t_{1}}^0\| ||\Gamma_{t_{2}}^0\| \]

\[ \leq C^2 \cdot N^{-1} \sum_{t_{1}=1}^{N} \sum_{t_{2}=1}^{N} |\ell_{1} \ell_{2} \cdot N^{-1}T^{-1} \sum_{h_{1}=1}^{N} \sum_{h_{2}=1}^{N} \sum_{t_{1}=1}^{T} \sum_{t_{2}=1}^{T} \|\sigma_{h_{1}} \sigma_{h_{2}} \varepsilon_{t_{1}} \varepsilon_{t_{2}}\| |\|

\[ \leq C \]

by Assumption A2, B2 and D. Collecting the above terms, we can derive that

\[ N^{-3/2}T^{-1/2} \sum_{i=1}^{N} N^{-1} \sum_{t=1}^{T} \Gamma_i^0 (\nu(0)^{-1}) \Gamma_i^0 V_t M P^0 M H u_i = -N^{-3/2}T^{-1/2} \sum_{i=1}^{N} N^{-1} \sum_{t=1}^{T} \Gamma_i^0 (\nu(0)^{-1}) \Gamma_i^0 V_t M P^0 M H \varepsilon_i \]

\[ -N^{-5/2}T^{-1/2} \sum_{i=1}^{N} N^{-1} \sum_{t=1}^{T} \sum_{h=1}^{N} \Gamma_i^0 (\nu(0)^{-1}) \Gamma_i^0 V_t \varepsilon_h \varepsilon_h \varphi_h^0 (\nu(0)^{-1}) \varphi_i \]

\[ -N^{-3/2}T^{-3/2} \sum_{i=1}^{N} N^{-1} \sum_{t=1}^{T} \sum_{h=1}^{N} \Gamma_i^0 (\nu(0)^{-1}) \Gamma_i^0 V_t \varepsilon_h \Sigma_h (T^{-1} H^0 H^0)^{-1} (\nu(0)^{-1}) \varphi_i \]

\[ + O_p(T^{1/2} \delta_{NT}^{-2}) \]

Consequently, with (B.3), we complete the proof. \( \square \)

C Proofs of Lemmas in Appendix C

Proof of Lemma C.1. Consider (a). With the equation (A.1), we have

\[ ||T^{-1} \varepsilon_i^t (\nu(0)^{-1} - \hat{\nu} F R^{-1})|| \]

\[ \leq N^{-1}T^{-2} \sum_{t=1}^{N} \varepsilon_i^t (\nu(0)^{-1} V_t^t F)^{t} ||\Sigma^{-1} R^{-1}|| + N^{-1}T^{-2} \sum_{t=1}^{N} \varepsilon_i^t V_t (\nu(0)^{-1} F)^{t} ||\Sigma^{-1} R^{-1}|| \]

\[ + N^{-1}T^{-2} \sum_{t=1}^{N} \varepsilon_i^t V_i \varepsilon_i^t F ||\Sigma^{-1} R^{-1}|| . \]
Since $\Xi^{-1} = O_p(1)$ and $R^{-1} = O_p(1)$, we omit $\|\Xi^{-1}R^{-1}\|$ in the following analysis. The first term is bounded in norm by

$$T^{-1/2} \cdot \|T^{-1/2} \varepsilon_i^2 F^i_0\| \cdot \left\| N^{-1} T^{-1} \sum_{t=1}^N \Gamma_t \varepsilon_i V_t F^i_0 \right\|$$

With Assumptions A and C, we have $E\|T^{-1/2} \varepsilon_i^2 F^i_0\|^2 = T^{-1} \sum_{t=1}^T E\|\varepsilon_i\|^2 \|f^i_t\|^2 \leq C$, which then implies that

$$\|T^{-1/2} \varepsilon_i^2 F^i_0\| = O_p(1). \quad \text{(C.4)}$$

By Lemma A.1 (h), we have

$$N^{-1} T^{-1} \sum_{t=1}^N \tilde{F} \varepsilon_i V_t \Gamma_t = N^{-1} T^{-1} \sum_{t=1}^N R \varepsilon_i V_t \Gamma_t + N^{-1} T^{-1} \sum_{t=1}^N (\tilde{F} - F^0) / V_t \Gamma_t \quad \text{(C.5)}$$

$$= O_p(N^{-1/2} T^{-1/2}) + O_p(N^{-1}) + O_p(N^{-1/2} \delta_{NT}^2).$$

With (C.4) and (C.5), the first term is $O_p(N^{-1/2} T^{-1}) + O_p(N^{-1/2} T^{-1/2}) + O_p(N^{-1/2} T^{-1/2} \delta_{NT}^2)$. The second term is bounded in norm by

$$N^{-1/2} T^{-1/2} \left\| N^{-1/2} T^{-1/2} \sum_{t=1}^N \varepsilon_i V_t \Gamma_t \right\| \|T^{-1/2} F^0\| \|T^{-1/2} \tilde{F}\| = O_p(N^{-1/2} T^{-1/2})$$

where $\|N^{-1/2} T^{-1/2} \sum_{t=1}^N \varepsilon_i V_t \Gamma_t \| = O_p(1)$ can be proved by following the way in the proof of (C.4). Consider the third term. Easily, we can prove $E\|T^{-1/2} \varepsilon_i\|^2 \leq C$. By Cauchy-Schwartz inequality, we have

$$E\left\| N^{-1} T^{-1} \sum_{t=1}^N \varepsilon_i^2 \|E(V_t V_t') F^i_0\|^2 \right\| = \|T^{-1} \sum_{s=1}^T \sum_{t=1}^T (N^{-1} T^{-1} \sum_{t=1}^N E(V' s_t V_t)) \varepsilon_i V_t F^i_0 \|^2 \right\| \leq \left( N^{-1} T^{-1} \sum_{t=1}^N \varepsilon_i V_t \Gamma_t \right) \left( N^{-1} T^{-1} \sum_{t=1}^N \varepsilon_i V_t \Gamma_t \right) \left( N^{-1} T^{-1} \sum_{t=1}^N \varepsilon_i V_t \Gamma_t \right) \left( N^{-1} T^{-1} \sum_{t=1}^N \varepsilon_i V_t \Gamma_t \right) \leq C(T) \left( N^{-1} T^{-1} \sum_{t=1}^T \sum_{s=1}^T \varepsilon_i V_t \Gamma_t \right) \leq C^3$$

by Assumptions A, B2, and C. With Assumption B5, we can follow the way of the proof of Lemma A.2(i) in Bai (2009) to show that $E\|N^{-1/2} T^{-1} \sum_{t=1}^N \varepsilon_i^2 [V_t V_t' - E(V_t V_t')] F^i_0\|^2 \leq C$. With the above three moment conditions, we obtain

$$\|T^{-1/2} \varepsilon_i\| = O_p(1)$$

$$\|N^{-1} T^{-1} \sum_{t=1}^N \varepsilon_i^2 \|E(V_t V_t') F^i_0\| = O_p(1)$$

$$\|N^{-1/2} T^{-1} \sum_{t=1}^N \varepsilon_i^2 [V_t V_t' - E(V_t V_t')] F^i_0\| = O_p(1)$$

Thus, the third term is bounded in norm by

$$\|T^{-1/2} \varepsilon_i\| \|N^{-1} T^{-1} \sum_{t=1}^N V_t V_t'\| \|T^{-1/2} (\tilde{F} - F^0)\| \|N^{-1} T^{-1} \sum_{t=1}^N \varepsilon_i^2 \|E(V_t V_t') F^i_0\| \|R\|$$

$$+ N^{-1/2} T^{-1} \|N^{-1/2} T^{-1} \sum_{t=1}^N \varepsilon_i^2 [V_t V_t' - E(V_t V_t')] F^i_0\| \|R\| = O_p(\delta_{NT}^2).$$

35
where
\[ \| N^{-1} T^{-1} \sum_{\ell=1}^{N} V_{\ell} V'_{\ell} \| = O_p(\delta_{NT}^{-1}). \] (C.7)
which suggest from
\[
\| N^{-1} T^{-1/2} \sum_{\ell=1}^{N} E(V_{\ell} V'_{\ell}) \|^2 = T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \left| N^{-1} \sum_{i=1}^{N} E(v'_{\ell s} v_{t i}) \right|^2
\leq C N^{-1} T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \sum_{i=1}^{N} |E(v'_{\ell s} v_{t i})| \leq C T^{-1} \sum_{s=1}^{T} \sum_{t=1}^{T} \hat{\sigma}_{st} \leq C^2,
\]
and
\[
E\left\| N^{-1/2} T^{-1} \sum_{\ell=1}^{N} [V_{\ell} V'_{\ell} - E(V_{\ell} V'_{\ell})] \right\|^2 = E\left( T^{-2} \sum_{s=1}^{T} \sum_{t=1}^{T} [N^{-1/2} \sum_{\ell=1}^{N} (v'_{\ell s} v_{t i} - E(v'_{\ell s} v_{t i}))]^2 \right)
= T^{-2} \sum_{s=1}^{T} \sum_{t=1}^{T} E\left[ N^{-1/2} \sum_{\ell=1}^{N} (v'_{\ell s} v_{t i} - E(v'_{\ell s} v_{t i})) \right]^2 \leq C,
\]
given \( |N^{-1} \sum_{s=1}^{T} |E(v'_{\ell s} v_{t i})| \leq N^{-1} \sum_{s=1}^{T} |E(v'_{\ell s} v_{t i})| \leq N^{-1} \sum_{s=1}^{T} \sqrt{E|v_{\ell s}||v_{t i}|^2} \leq C \) and Assumption B. Collecting the above three terms, the claim holds. Similarly, we can prove (b), details are omitted. This completes the proof. □

**Proof of Lemma C.2.** We first consider \( T^{-1/2} \Gamma_i^{0\prime} F^{0\prime} M_F u_i \), which is
\[
T^{-1/2} \Gamma_i^{0\prime} F^{0\prime} M_F u_i = O_p(1) \times T^{-1/2} F^{0\prime} M_F u_i.
\]
Note that \( M_F F^{0} = 0 \), we have \( M_F^2 F^{0} = (M_F - M_{F^0}) F^{0} \). We expand \( M_F^2 - M_{F^0} \) as
\[
- \frac{1}{T} (\hat{F} - F^{0\prime} R) R F^{0\prime} - \frac{1}{T} F^{0\prime} R (\hat{F} - F^{0\prime} R)' - \frac{1}{T} (\hat{F} - F^{0\prime} R) (\hat{F} - F^{0\prime} R)' - \frac{1}{T} F^{0} \left( R R' - \left( \frac{F^{0\prime} F^{0}}{T} \right)^{-1} \right) F^{0\prime},
\]
then
\[
\frac{1}{T^{1/2}} F^{0\prime} M_F^2 u_i = - \frac{1}{T^{3/2}} F^{0\prime} (\hat{F} - F^{0\prime} R) R F^{0\prime} u_i - \frac{1}{T^{3/2}} F^{0\prime} F^{0} R (\hat{F} - F^{0\prime} R)' u_i
- \frac{1}{T^{3/2}} F^{0\prime} R (\hat{F} - F^{0\prime} R)' u_i - \frac{1}{T^{3/2}} F^{0\prime} F^{0} \left( R R' - \left( \frac{F^{0\prime} F^{0}}{T} \right)^{-1} \right) F^{0\prime} u_i
= \bar{A}_1 + \bar{A}_2 + \bar{A}_3 + \bar{A}_4.
\]
Consider \( \bar{A}_1 \). Given \( u_i = H^0 \varphi^0_i + \varepsilon_i \), we have
\[
\| T^{-1/2} u_i \| \leq \| T^{-1/2} F^{0\prime} \| + \| \varphi^0_i \| \| T^{-1/2} H^0 \| + \| T^{-1/2} \varepsilon_i \|.
\]
Since \( \| \varphi^0 \| \leq C \) and the condition \( \| T^{-1/2} H^0 \| \leq C \) by Assumptions C and D, the first term is \( O_p(1) \). Similarly, we can prove that the second and the third term both are \( O_p(1) \). The above facts suggest \( \| T^{-1/2} u_i \| = O_p(1) \). Thus, \( \bar{A}_1 \) is bounded in norm by
\[
\| u_i \| \times \| T^{-1/2} F^{0\prime} (\hat{F} - F^{0\prime} R) \| \| R \| \| T^{-1/2} F^{0\prime} \| = O_p(T^{1/2} \delta_{NT}^{-2}).
\]
Similarly, we can show that \( \bar{A}_3 = O_p(T^{1/2} \delta_{NT}^{-2}) \) and \( \bar{A}_4 = O_p(T^{1/2} \delta_{NT}^{-2}) \). Consider \( \bar{A}_2 \). The term is bounded in norm by \( T^{1/2} \| T^{-1} (\hat{F} - F^{0\prime} R)' u_i \| \times \| T^{-1/2} F^{0\prime} \| \| R \| \), which is \( O_p(T^{1/2}) \times \| T^{-1} (\hat{F} - F^{0\prime} R)' u_i \| \). Furthermore, \( \| T^{-1} (\hat{F} - F^{0\prime} R)' u_i \| \) is bounded in norm by
\[
\| \varphi^0_i \| T^{-1} (\hat{F} - F^{0\prime} R)' H^0 \| + \| T^{-1} (\hat{F} R^{-1} - F^{0\prime}) \varepsilon_i \| \| R \| = O_p(\delta_{NT}^{-2})
\]
by Lemmas A.1(b) and A.2(a). Thus, \( \mathcal{A}_2 = O_p(T^{1/2}\delta_{NT}^{-2}) \). Collecting the above four terms, we have \( T^{-1/2}F_i^\prime (M_{\bar{F}} - M_{F_0})u_i = O_p(T^{1/2}\delta_{NT}^{-2}) \).

Next, we tend to prove that \( T^{-1/2}V_i(M_{\bar{F}} - M_{F_0})u_i = O_p(T^{1/2}\delta_{NT}^{-2}) \). Since \( M_{\bar{F}} - M_{F_0} = -T^{-1}((\hat{F} - F_0')R_1'F_{0'}^\prime - T^{-1}F_1'R_1'(\hat{F} - F_0') - T^{-1}F_1'(\hat{H} - F_0') \circ \gamma - T^{-1}F_0')^{-1}F_{0'}^\prime , \) we have

\[
T^{-1/2}V_i(M_{\bar{F}} - M_{F_0})u_i = -T^{-3/2}V_i((\hat{F} - F_0')R_1'F_{0'}^\prime u_i - T^{-3/2}V_iF_1'R_1'((\hat{F} - F_0')u_i - T^{-3/2}V_iF_1'R_1'((\hat{F} - F_0')u_i - T^{-3/2}V_iF_1'R_1'((\hat{F} - F_0')u_i = \mathcal{A}_5 + \mathcal{A}_6 + \mathcal{A}_7 + \mathcal{A}_8 .
\]

\( \mathcal{A}_5 \) in norm by

\[
T^{1/2}\|\varphi_i^0\| \cdot \|T^{-1}V_i((\hat{F} - F_0')u_i - T^{-1}F_0'H_0^\prime \| \|R\| + \|T^{-1/2}V_i((\hat{F} - F_0')u_i - T^{-1}F_0'H_0^\prime \| \|R\| = O_p(T^{1/2}\delta_{NT}^{-2}) .
\]

With Lemma A.2(a), \( \mathcal{A}_6 \) is bounded in norm by

\[
\|T^{-1/2}V_i((\hat{F} - F_0')u_i - T^{-1}F_0'H_0^\prime \| \|R\| + \|T^{-1/2}V_i((\hat{F} - F_0')u_i - T^{-1}F_0'H_0^\prime \| \|R\| = O_p(\delta_{NT}^{-2}) .
\]

by Lemmas A.1(b). Given Lemmas A.1(b), A.2(a) and A.2(b), \( \mathcal{A}_7 \) is bounded in norm by

\[
T^{1/2}\|\varphi_i^0\| \cdot \|T^{-1}V_i((\hat{F} - F_0')u_i - T^{-1}F_0'H_0^\prime \| \|R\| + T^{1/2}\|\varphi_i^0\| \cdot \|T^{-1}V_i((\hat{F} - F_0')u_i - T^{-1}F_0'H_0^\prime \| \|R\| = O_p(T^{1/2}\delta_{NT}^{-2}) .
\]

\( \mathcal{A}_8 \) is bounded in norm by

\[
\|T^{-1/2}V_i((\hat{F} - F_0')u_i - T^{-1}F_0'H_0^\prime \| \|R\| + \|T^{-1/2}V_i((\hat{F} - F_0')u_i - T^{-1}F_0'H_0^\prime \| \|R\| = O_p(\delta_{NT}^{-2}) .
\]

by Lemma A.1(c). With the stochastic orders of the above eight terms, we derive that

\[
T^{-1/2}X_i'M_{\bar{F}}u_i = T^{-1/2}X_i'M_{F_0}u_i + O_p(T^{1/2}\delta_{NT}^{-2}) .
\]

we complete the proof. □

**Proof of Lemma C.3.** The results are immediately obtained from Assumptions D and H. □

**Proof of Lemma C.4.** Consider (a). With the equation (A.1), we have

\[
\sup_{1 \leq i \leq N} \|T^{-1}e_i'(F_0' - \hat{F}'R_1'\| \leq \sup_{1 \leq i \leq N} N^{-1}T^{-2}\| \sum_{\ell=1}^{N} e_{i,\ell}'F_\ell^0T_{i,\ell}'V_i'\hat{F}'\| \|\Xi^{-1}R_1'\| \|T^{-1}F_0'\| + \sup_{1 \leq i \leq N} N^{-1}T^{-2}\| \sum_{\ell=1}^{N} e_{i,\ell}'V_i'F_\ell^0F_0'\| \|\Xi^{-1}R_1'\| \|T^{-1}F_0'\| + \sup_{1 \leq i \leq N} N^{-1}T^{-2}\| \sum_{\ell=1}^{N} e_{i,\ell}'V_i'V_i'\hat{F}'\| \|\Xi^{-1}R_1'\|
\]

Since \( \Xi^{-1} = O_p(1) \) and \( R_1^{-1} = O_p(1) \), we omit \( \|\Xi^{-1}R_1'\| \) in the following analysis. The first term is bounded in norm by

\[
T^{-1/2} \sup_{1 \leq i \leq N} \|T^{-1/2}e_i'F_0'\| \cdot \|N^{-1}T^{-1}\sum_{\ell=1}^{N} F_\ell^0V_i'\hat{F}'\|
\]
Since \(E|T^{-1/2}\varepsilon_t^i\mathbf{F}^0||^4 \leq C\), we have

\[
\sup_{1 \leq i \leq N} \|T^{-1/2}\varepsilon_t^i\mathbf{F}^0\| = O_p(N^{1/4}) \tag{C.8}
\]

Note that \(N^{-1}T^{-1} \sum_{t=1}^N \hat{F}^i\mathbf{V}_t\mathbf{F}^0 = O_p(N^{-1/2}T^{-1/2}) + O_p(N^{-1/4}) + O_p(N^{-1/2}\delta_{NT}^2)\) by (C.5), and with (C.8), the first term is \(O_p(N^{-1/4}T^{-1}) + O_p(N^{-3/4}T^{-1/2}) + O_p(N^{-1/2}T^{-1/2}\delta_{NT}^2)\). The second term is bounded in norm by

\[
N^{-1/2}T^{-1/2} \sup_{1 \leq i \leq N} \|N^{-1/2}T^{-1/2} \sum_{t=1}^N \varepsilon_t^i \mathbf{V}_t\mathbf{F}^0\| \cdot \|T^{-1/2}\mathbf{F}^0\| \|T^{-1/2}\hat{F}^0\| = O_p(N^{-1/4}T^{-1/2})
\]

by Lemma C.3(b). Consider the third term. We have

\[
\sup_{1 \leq i \leq N} \|T^{-1/2}\varepsilon_t^i\|^2 = \sup_{1 \leq i \leq N} T^{-1} \sum_{t=1}^T \varepsilon_{it}^2 \leq \varepsilon_{max}^2 + T^{-1/2} . \sup_{1 \leq i \leq N} T^{-1/2} \sum_{t=1}^T (\varepsilon_{it}^2 - \varepsilon_{it}^2) = O_p(1) + O_p(N^{1/3}T^{-1/2})
\]

since \(E|T^{-1/2} \sum_{t=1}^T (\varepsilon_{it}^2 - \varepsilon_{it}^2)|^3 \leq C\). With (A.8), we can show that \(E\|N^{-1/2}T^{-1} \sum_{t=1}^N [\mathbf{V}_t\mathbf{V}_t' - \mathbf{E}(\mathbf{V}_t\mathbf{V}_t')]\|^{1/2} \leq C\). Thus, the third term is bounded in norm by

\[
\sup_{1 \leq i \leq N} T^{-1/2} \|\varepsilon_t^i\| \cdot \|N^{-1/2}T^{-1} \sum_{t=1}^N \mathbf{V}_t\mathbf{V}_t'\| \cdot \|T^{-1/2}(\hat{F} - \mathbf{F}^0\mathbf{R})\| + T^{-1} \cdot \sup_{1 \leq i \leq N} T^{-1/2} \sum_{t=1}^N \varepsilon_t^i \|\mathbf{E}(\mathbf{V}_t\mathbf{V}_t')\| \cdot \|\mathbf{F}^0\| \cdot \|\mathbf{R}\|
\]

\[
= O_p(\delta_{NT}^2) + O_p(N^{1/4}T^{-1})
\]

Collecting the above three terms, the claim holds.

Consider (b). Replacing \(\hat{F} - \mathbf{F}^0\mathbf{R}\) by its expression (A.1), we have

\[
\sup_{1 \leq i \leq N} \|T^{-1/2}\mathbf{V}_t^i(\hat{F} - \mathbf{F}^0\mathbf{R})\| \leq \sup_{1 \leq i \leq N} N^{-1}T^{-2} \|\sum_{t=1}^N \mathbf{V}_t^i\mathbf{F}^0\| \|\mathbf{V}_t^i\| \|\mathbf{F}^0\| \cdot \|\mathbf{V}_t^i\| \cdot \|\mathbf{F}^0\| \cdot \|\mathbf{R}\|
\]

\[
= \sum_{t=1}^N \|\mathbf{V}_t^i\| \cdot \|\mathbf{F}^0\| \cdot \|\mathbf{R}\|
\]

Ignoring \(\|\mathbf{E}^{-1}\|\) and following the arguments of the first term and the third term in the proof of (a), the first term is \(O_p(N^{-1/4}T^{-1}) + O_p(N^{-3/4}T^{-1/2}) + O_p(N^{-1/2}T^{-1/2}\delta_{NT}^2)\) and the third term is \(O_p(\delta_{NT}^2) + O_p(N^{1/4}T^{-1})\). The second term is bounded in norm by

\[
\sup_{1 \leq i \leq N} N^{-1}T^{-1} \|\sum_{t=1}^N \mathbf{V}_t^i\| \|\mathbf{F}^0\| \cdot \|\mathbf{R}\| \leq \sup_{1 \leq i \leq N} N^{-1}T^{-1} \|\sum_{t=1}^N \mathbf{V}_t^i\| \|\mathbf{F}^0\| \cdot \|\mathbf{R}\|
\]

\[
= \sum_{t=1}^N \|\mathbf{V}_t^i\| \cdot \|\mathbf{F}^0\| \cdot \|\mathbf{R}\|
\]

Combining the above three terms, (b) holds. This completes the proof.
Proof of Lemma C.5. Consider (a). The left hand is bounded in norm by

\[ N^{-1}T^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 F^0 M_{\hat{F}} u_i \| + N^{-1}T^{-1} \sum_{i=1}^{N} \| V_i^0 M_{\hat{F}} u_i - V_i^0 M_{F^0} u_i \| \]

We first consider \( N^{-1}T^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 F^0 M_{\hat{F}} u_i \| \), which is bounded by \( N^{-1}T^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| F^0 M_{\hat{F}} u_i \| \). Note that \( M_{F^0} F^0 = 0 \), we have \( M_{\hat{F}} F^0 = (M_{\hat{F}} - M_{F^0}) F^0 \). We expand \( M_{\hat{F}} - M_{F^0} \) as following

\[ -\frac{1}{T} (\hat{F} - F^0 R) R F^0 - \frac{1}{T} F^0 R (\hat{F} - F^0 R)' - \frac{1}{T} (\hat{F} - F^0 R) (\hat{F} - F^0 R)' - \frac{1}{T} F^0 \left( RR' - \left( \frac{F^0 F^0}{T} \right)^{-1} \right) F^0, \]

then

\[ N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-1} F^0 M_{\hat{F}} u_i \| \]

\[ \leq N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-2} F^0 (\hat{F} - F^0 R) R F^0 u_i \| + N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-2} F^0 F^0 R (\hat{F} - F^0 R) u_i \| \]

\[ + N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-2} F^0 (\hat{F} - F^0 R) (\hat{F} - F^0 R)' u_i \| + N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-2} F^0 F^0 \left( RR' - \left( \frac{F^0 F^0}{T} \right)^{-1} \right) F^0 u_i \| \]

\[ = B_1 + B_2 + B_3 + B_4. \]

Consider \( B_1 \). Given \( u_i = H^0 \varphi_i^0 + \varepsilon_i \), we have

\[ N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-1} F^0 u_i \| \]

\[ \leq N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| \varphi_i^0 \| \| T^{-1} F^0 H^0 \| \| + N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-1} F^0 \varepsilon_i \| \]

\[ = O_p(1) \]

Thus, \( B_1 \) is bounded in norm by

\[ N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-1} F^0 u_i \| \times \| T^{-1} F^0 (\hat{F} - F^0 R) \| \| R \| \]

\[ = O_p(\delta_{NT}^{-2}). \]

Similarly, we can show that \( B_4 = O_p(\delta_{NT}^{-2}). \)

Consider \( B_2 \). The term is bounded in norm by \( N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-1} (\hat{F} - F^0 R)' u_i \| \| T^{-1/2} F^0 \| \| R \| \), which is \( O_p(1) \cdot N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-1} (\hat{F} - F^0 R)' u_i \| \). Furthermore, \( N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-1} (\hat{F} - F^0 R)' u_i \| \)

is bounded in norm by

\[ N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| \varphi_i^0 \| \| T^{-1} (\hat{F} - F^0 R)' H^0 \| \]

\[ + N^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 \| \| T^{-1} (\hat{F} - F^0 R)' \varepsilon_i \| \| R \| \]

\[ = O_p(\delta_{NT}^{-2}). \]

by Lemmas A.1(b) and A.2(b). Thus, \( B_2 = O_p(\delta_{NT}^{-2}). \) Analogously, we have \( B_3 = O_p(\delta_{NT}^{-4}). \)

Collecting the above four terms, we have \( N^{-1}T^{-1} \sum_{i=1}^{N} \| \Gamma_i^0 F^0 M_{\hat{F}} u_i \| = O_p(\delta_{NT}^{-2}). \)
Next, we tend to prove that $N^{-1}T^{-1} \sum_{i=1}^{N} \| V_i'(M_F - M_{F^0})u_i \| = O_p(\delta_{NT}^{-2})$. Since $M_F - M_{F^0} = -T^{-1}(\hat{F} - F^0R)R^0' - T^{-1}F^0R(\hat{F} - F^0R)' - T^{-1}(\hat{F} - F^0R)(\hat{F} - F^0R)' - T^{-1}F^0(RR' - (T^{-1}F^0F^0)^{-1})F^0,$ we have

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1}V_i'(M_F - M_{F^0})u_i \| = N^{-1} \sum_{i=1}^{N} \| T^{-2}V_i'(\hat{F} - F^0R)R^0'u_i \| + N^{-1} \sum_{i=1}^{N} \| T^{-2}V_i'F^0R(\hat{F} - F^0R)'u_i \|
\]

\[
+ N^{-1} \sum_{i=1}^{N} \| T^{-2}V_i'(\hat{F} - F^0R)(\hat{F} - F^0R)'u_i \| + N^{-1} \sum_{i=1}^{N} \| T^{-2}V_i'F^0(\hat{F} - F^0R)'u_i \| = B_5 + B_6 + B_7 + B_8
\]

We bound $B_5$ in norm by

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1}V_i'(\hat{F} - F^0R)\| T^{-1}F^0'u_i \| \| R \|
\]

\[
\leq N^{-1} \sum_{i=1}^{N} \| T^{-1}V_i'(\hat{F} - F^0R)\| \| \varphi_i' \| \| R \| \| T^{-1}F^0'H^0 \|
\]

\[
+ N^{-1} \sum_{i=1}^{N} \| T^{-1}V_i'(\hat{F} - F^0R)\| T^{-1}F^0'e_i \| \| R \|
\]

\[
= O_p(\delta_{NT}^{-2})
\]

by Lemma A.2. With Lemma A.2(a), $B_6$ is bounded in norm by

\[
N^{-1} \sum_{i=1}^{N} \| T^{-1}V_i'F^0'\| T^{-1}(\hat{F} - F^0R)'u_i \| \| R \| = O_p(T^{-1/2}\delta_{NT}^{-2})
\]

by Lemmas A.1(b). Similarly, we can show that $B_7 = O_p(\delta_{NT}^{-4})$ and $B_8 = O_p(T^{-1/2}\delta_{NT}^{-2})$. With the stochastic orders of the above eight terms, we obtain (a).

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

\[
\sup_{1 \leq i \leq N} \| T^{-1}X_i'M_FX_i - T^{-1}X_i'M_{F^0}X_i \|
\]

\[
\leq \sup_{1 \leq i \leq N} \| T^{-1}T_i^0F^0'M_FF^0T_i^0 \| + 2 \sup_{1 \leq i \leq N} \| T^{-1}V_i'M_F^0\| + \sup_{1 \leq i \leq N} \| T^{-1}V_i'(M_F - M_{F^0})V_i \|
\]

\[
=C_1 + C_2 + C_3
\]

$C_1$ is bounded in norm by

\[
\sup_{1 \leq i \leq N} \| T^{-1}T_i^0F^0'M_FF^0T_i^0 \| = \sup_{1 \leq i \leq N} \| T^{-1}T_i^0(F^0 - \hat{F}R^{-1})'M_F(F^0 - \hat{F}R^{-1})T_i^0 \|
\]

\[
\leq \sup_{1 \leq i \leq N} \| T_i^0 \| ^2 \cdot \| T^{-1/2}(F^0 - \hat{F}R^{-1}) \|^2 = O_p(N^{1/2}\delta_{NT}^{-2})
\]

Ignoring the scale 2, $C_2$ is bounded in norm by

\[
\sup_{1 \leq i \leq N} \| T^{-1}V_i'(M_F - M_{F^0})F^0T_i^0 \|
\]

\[
= \sup_{1 \leq i \leq N} \| T^{-2}V_i'(\hat{F} - F^0R)R^0'F^0T_i^0 \| + \sup_{1 \leq i \leq N} \| T^{-2}V_i'F^0R(\hat{F} - F^0R)'F^0T_i^0 \|
\]

\[
+ \sup_{1 \leq i \leq N} \| T^{-2}V_i'(\hat{F} - F^0R)(\hat{F} - F^0R)'F^0T_i^0 \| + \sup_{1 \leq i \leq N} \| T^{-2}V_i'F^0(\hat{F} - F^0R)'F^0T_i^0 \|
\]

\[
= B_5 + B_6 + B_7 + B_8
\]
We bound the first term in norm by
\[ \sup_{1 \leq i \leq N} \|T^{-1}V_i'\hat{F} - F^0_i\| \cdot \sup_{1 \leq i \leq N} \|F^0_i\| \|T^{-1}F^0F^0\| = O_p(1/2\delta_{NT}^{-2}) \]

With Lemma A.2(a), the second term is bounded in norm by
\[ \sup_{1 \leq i \leq N} \|T^{-1}V_i'F^0\| \cdot \sup_{1 \leq i \leq N} \|T^{-1}(\hat{F} - F^0_i)\| \|T^{-1}F^0\| = O_p(1/2T^{-1/2}\delta_{NT}^{-2}) \]

by Lemmas A.1(b). Given Lemmas A.1(b), A.2(a) and A.2(b), the third is bounded in norm by
\[ \sup_{1 \leq i \leq N} \|T^{-1}V_i'(\hat{F} - F^0_i)\| \cdot \sup_{1 \leq i \leq N} \|T^{-1}(\hat{F} - F^0_i)\| \|T^{-1}F^0\| = O_p(1/2\delta_{NT}^{-4}) \]

the forth term is bounded in norm by
\[ \sup_{1 \leq i \leq N} \|T^{-1}V_i'\| \cdot \sup_{1 \leq i \leq N} \|T^{-1}V_i'(R - F^0) - (T^{-1}F^0)^{-1}\| \|T^{-1}F^0\| = O_p(1/2T^{-1/2}\delta_{NT}^{-2}) \]

by Lemma A.1(c). Thus \( C_2 = O_p(N^{-1/2}\delta_{NT}^{-2}) \). \( C_3 \) is bounded in norm by
\[ \sup_{1 \leq i \leq N} \|T^{-1}V_i'(M_{\bar{F}} - M_{F^0})V_i\| = \sup_{1 \leq i \leq N} \|T^{-2}V_i'\| \cdot \sup_{1 \leq i \leq N} \|T^{-2}V_i'F^0_{\bar{R}}\| \cdot \sup_{1 \leq i \leq N} \|T^{-2}V_i'F^0_{\bar{R}}V_i\| + \sup_{1 \leq i \leq N} \|T^{-2}V_i'V_i\| \cdot \sup_{1 \leq i \leq N} \|V_i\| \]

The first term is bounded in norm by
\[ \sup_{1 \leq i \leq N} \|T^{-1}V_i'\hat{F} - F^0_i\| \cdot \sup_{1 \leq i \leq N} \|T^{-1}F^0\| \cdot \|\bar{R}\| = O_p(N^{-1/2T^{-1/2}\delta_{NT}^{-2}}) \]

Similarly, the second term is \( O_p(N^{-1/2T^{-1/2}\delta_{NT}^{-2}}) \). The third term is bounded in norm by
\[ \left( \sup_{1 \leq i \leq N} \|T^{-1}V_i'(\hat{F} - F^0_i)\| \right)^2 = O_p(N^{-1/2\delta_{NT}^{-4}}) \]

The fourth term is bounded in norm by
\[ T^{-1} \cdot \sup_{1 \leq i \leq N} \|T^{-1/2}V_i'F^0\|^2 \cdot \|\bar{R}\| \cdot \left( T^{-1}F^0\right)^{-1} = O_p(N^{-1/2T^{-1/2}\delta_{NT}^{-2}}) \]

With the above terms, we have \( C_3 = O_p(N^{-1/2T^{-1/2}\delta_{NT}^{-2}} + O_p(N^{-1/2\delta_{NT}^{-4}}) \). Then, we have (b). Thus, we complete the proof. □

D Proofs of Lemmas in Appendix D

Proof of Lemmas D.1, D.2 and D.3. The results are straightforwardly derived following the proofs in Bai (2009) and Lemmas B.1-B.6, thus, details are omitted.
### E Additional Experimental Results

Table E.1: Bias, root mean squared error (RMSE) of the estimators of $\beta_1$, and size and power of the associated t-tests when $\pi_u = \{1/2, 3/4\}$ and $N = T = 200$.

| Estimator | Homogeneous Slopes | | | | Heterogeneous Slopes | | | |
### Table E.3: Scaled bias of the estimators of $\beta_1$.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>$\hat{\beta}_1$</th>
<th>$\hat{\beta}_2$</th>
<th>$\hat{\beta}_3$</th>
<th>$\hat{\beta}_4$</th>
<th>$\hat{\beta}_5$</th>
<th>$\hat{\beta}_6$</th>
<th>$\hat{\beta}_7$</th>
<th>$\hat{\beta}_8$</th>
<th>$\hat{\beta}_9$</th>
<th>$\hat{\beta}_{10}$</th>
<th>$\hat{\beta}_{11}$</th>
<th>$\hat{\beta}_{12}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SIV</td>
<td>$-0.215$</td>
<td>$-0.074$</td>
<td>$-0.039$</td>
<td>0.005</td>
<td>0.021</td>
<td>0.067</td>
<td>0.079</td>
<td>0.083</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC-IPC</td>
<td>$-0.254$</td>
<td>0.587</td>
<td>0.923</td>
<td>0.352</td>
<td>0.010</td>
<td>0.147</td>
<td>0.159</td>
<td>0.113</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPC</td>
<td>$-0.611$</td>
<td>$-0.599$</td>
<td>$-0.222$</td>
<td>$-0.041$</td>
<td>$-0.052$</td>
<td>$-0.011$</td>
<td>0.062</td>
<td>0.079</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC-PC</td>
<td>$-0.511$</td>
<td>$-0.732$</td>
<td>$-1.632$</td>
<td>$-3.249$</td>
<td>$-0.049$</td>
<td>$-0.037$</td>
<td>$-0.075$</td>
<td>$-0.147$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC</td>
<td>$-1.841$</td>
<td>$-2.249$</td>
<td>$-3.235$</td>
<td>$-4.877$</td>
<td>$-0.314$</td>
<td>$-0.255$</td>
<td>$-0.242$</td>
<td>$-0.265$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA</td>
<td>$-0.470$</td>
<td>$-0.760$</td>
<td>$-1.056$</td>
<td>$-1.177$</td>
<td>$-0.035$</td>
<td>$-0.040$</td>
<td>$-0.029$</td>
<td>$-0.001$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGIV</td>
<td>$-0.183$</td>
<td>$-0.059$</td>
<td>$-0.054$</td>
<td>0.018</td>
<td>$-0.040$</td>
<td>$-0.010$</td>
<td>$-0.004$</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGPC</td>
<td>$-1.836$</td>
<td>$-2.245$</td>
<td>$-3.235$</td>
<td>$-4.852$</td>
<td>$-0.360$</td>
<td>$-0.323$</td>
<td>$-0.317$</td>
<td>$-0.342$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGCA</td>
<td>$-0.398$</td>
<td>$-0.692$</td>
<td>$-1.000$</td>
<td>$-1.144$</td>
<td>$-0.083$</td>
<td>$-0.102$</td>
<td>$-0.099$</td>
<td>$-0.079$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: The DGP is the same as the one for Table 3, except the differences explained below the Table E.1.