LSU | E. J. Ourso College of Business
Department of Economics

DEPARTMENT OF ECONOMICS WORKING PAPER SERIES

JIVE for Panel Dynamic Simultaneous Equations Models

Cheng Hsiao
University of Southern California

Qiankun Zhou
Louisiana State University

Working Paper 2017-10
http://faculty.bus.lsu.edu/papers/pap17_10.pdf

Department of Economics
Louisiana State University
Baton Rouge, LA 70803-6306
http://www.bus.lsu.edu/economics/

JIVE for Panel Dynamic Simultaneous Equations Models*

Cheng Hsiao[†] Qiankun Zhou[‡]

July 24, 2017

Abstract

We consider the method of moments estimation of a structural equation in a panel dynamic simultaneous equations model under different sample size combinations of cross-sectional dimension, N, and time series dimension, T. Two types of linear transformation to remove the individual-specific effects from the model, first difference and forward orthogonal demeaning, are considered. We show that the Alvarez and Arellano (2003) type GMM estimator under both transformations is consistent only if $\frac{T}{N} \to 0$ as $(N, T) \to \infty$. However, it is asymptotically biased if $\frac{T^3}{N} \to \kappa \neq 0 < \infty$. Since the validity of statistical inference depends critically on whether an estimator is asymptotically unbiased, we suggest a jackknife bias reduction method and derive its limiting distribution. Monte Carlo studies are conducted to demonstrate the importance of using an asymptotically unbiased estimator to obtain valid statistical inference.

Keywords: Panel dynamic simultaneous equations model, GMM, First difference, Forward orthogonal demeaning, Jackknife instrumental variables estimation (JIVE).
JEL classification: C01, C30, C33

*We would like to thank the co-editor, Guido Kuersteiner, and John Chao, three anonymous referees for very detailed and helpful comments and suggestions, which have greatly improved the original paper. We would also like to thank Alan Adelman and John Pedersen for editorial polishing. All remaining errors are solely ours. Partial research support of China NSFgrant #71131008 and #71631004 to the first author is gratefully acknowledged.

[†]Department of Economics, University of Southern California, University Park, Los Angeles, California 90089, chsiao@usc.edu, Department of Quantitative Finance, National Tsing Hua University, Taiwan and WISE, Xiamen University, Xiamen, China.

[‡]Department of Economics, Louisiana State University, Baton Rouge, LA, 70803, email: qzhou@lsu.edu.

1 Introduction

Inertia in human behavior, and institutional and technological rigidities have lead many to believe that "all interesting economic behavior is inherently dynamic, dynamic models are the only relevant models" (Nerlove (2000)). However, the presence of time-invariant unobservable individual-specific effects in panel dynamic models create correlations between all current, past and future jointly dependent variables. For a linear regression model, the individual-specific effects affect the outcomes linearly, and they can be removed from the specification by taking linear difference of an individual's time series observations (e.g., Anderson and Hsiao (1981, 1982), Arellano and Bond (1991), Arellano and Bover (1995)). How this linear transformation is conducted does not affect the asymptotic distribution of an estimator if the regressors are strictly exogenous with respect to the idiosyncratic time-varying equation errors (e.g., Hsiao (2014)). But for a dynamic model, different form of linear transformation creates different form of correlations between the transformed lag dependent variables and the transformed individual time-varying errors of the equation. How this correlation affects the asymptotic distribution of a panel dynamic model estimator depends on the relative size of the cross-sectional dimension N and the time series dimension T (e.g., Alvarez and Arellano (2003), Akashi and Kunitomo (2012), Hahn and Kuersteiner (2002), Hahn and Newey (2004), Phillips and Moon (1999)).

For panel dynamic simultaneous equations models, there is another source of correlations, namely, the correlations between the regressors in a behavior equation with the error of the equation due to the joint dependence (e.g., Hood and Koopmans (1953)). The asymptotic bias of the conventional method of moments estimators arising from the correlations between the contemporary regressors and contemporary errors are not easily removed by using the lagged variables as instrument variables (IVs). For example, Akashi and Kunitomo (2012) have shown that if T increases with N and $\frac{T}{N} \to c \neq 0$ as $N \to \infty$, then the GMM estimator for panel dynamic simultaneous equations model is not even consistent.

Although consistency is one of the most important and desirable properties for an estimator, whether an estimator is asymptotically unbiased also plays a critical role in obtaining valid statistical inference (e.g., Hsiao and Zhang (2015), Hsiao and Zhou (2015)). In this paper, we first consider the asymptotic properties of the GMM estimator for a structural equation in a panel dynamic simultaneous equations model. We show that for a GMM estimator to be consistent, we will need N much larger than T in the sense $\frac{T}{N} \to 0$ as $N \to \infty$. However, as long as $\frac{T^3}{N} \to \kappa \neq 0 < \infty$ as $N \to \infty$, the GMM estimator is still asymptotically biased and the bias is of order $\sqrt{\kappa}$. Since the validity of statistical inference depends critically on an estimator is asymptotically unbiased or not, we suggest a jackknife procedure (e.g., Phillips and Hale (1977), Angrist et al (1999) and Chao et al (2012)) to correct the bias of GMM. We show that under the assumption that $(N, T) \to \infty$ with $\frac{T^3}{N} \to \kappa \neq 0 < \infty$,[1] the JIVE is asymptotically normal without an asymptotic bias.

The paper is organized as follows. In Section 2 we introduce a simple panel dynamic simultaneous equations model and two transformations that are often used to eliminate the individual-specific effects in the dynamic simultaneous equations model, and discuss their valid instrumental variables (IVs). Section 3 investigates the asymptotics of the GMM estimator based on the IVs in Section 2. We characterize the many IVs bias of the GMM estimator under different sample size combinations of N and T. In Section 4 we introduce the JIVE estimator and derive its asymptotic properties. In Section 5 we investigate finite sample properties of the GMM estimator and the JIVE using Monte Carlo simulations. Section 6 concludes the paper. All the mathematical proofs and derivations are

[1] The alternative asymptotics is introduced by Lee et al (2015) where they consider $\frac{T^3}{N} \to \kappa \neq 0 < \infty$, which is alternative to the asymptotics $\frac{T}{N} \to c \neq 0 < \infty$ considered by Alvarez and Arellano (2003).

presented in the appendix.

2 Model

We consider the statistical properties of the GMM estimator of a parametrically identified equation in a panel dynamic simultaneous equations model. Since it is the joint dependence of a $G \times 1$ vector $\mathbf{y}_{it}$ and the dependence between $\mathbf{y}_{it}$ and $\mathbf{y}_{is}$ $(t \neq s)$ that impact the asymptotic distribution of an estimator, not the fixed dimension strictly exogenous explanatory variables, $\mathbf{x}_{it}$,[2] there is no loss of generality to consider the following two equations system $(G = 2)$[3]

$$\begin{aligned} y_{1,it} &= \gamma y_{1,it-1} + \beta y_{2,it} + \alpha_{1i} + u_{1,it}, \\ y_{2,it} &= \gamma_{21} y_{1,it-1} + \gamma_{22} y_{2,it-1} + \alpha_{2i} + u_{2,it}, \;\; i = 1, \ldots, N, t = 1, \ldots, T, \end{aligned} \tag{1}$$

For ease of notations, we also assume $\mathbf{y}_{i0} = (y_{1,i0}, y_{2,i0})'$ are observable. Following the limited information approach of Anderson and Rubin (1949), there is no loss of generality to consider the estimation of the first equation in system (1), β and γ.

The reduced form of (1) is

$$\mathbf{y}_{it} = \Pi \mathbf{y}_{it-1} + \boldsymbol{\xi}_i + \mathbf{v}_{it}, \tag{2}$$

where $\mathbf{y}_{it} = (y_{1,it}, y_{2,it})'$, $\mathbf{y}_{it-1} = (y_{1,it-1}, y_{2,it-1})'$ and

$$\mathbf{B} = \begin{pmatrix} 1 & -\beta \\ 0 & 1 \end{pmatrix}, \Pi = \begin{pmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{pmatrix} = \mathbf{B}^{-1} \begin{pmatrix} \gamma & 0 \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} \gamma + \beta\gamma_{21} & \beta\gamma_{22} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}, \tag{3}$$

$$\boldsymbol{\xi}_i = \mathbf{B}^{-1} \boldsymbol{\alpha}_i = \mathbf{B}^{-1} \begin{pmatrix} \alpha_{1i} \\ \alpha_{2i} \end{pmatrix}, \mathbf{v}_{it} = \mathbf{B}^{-1} \mathbf{u}_{it} = \mathbf{B}^{-1} \begin{pmatrix} u_{1,it} \\ u_{2,it} \end{pmatrix} = \begin{pmatrix} u_{1,it} + \beta u_{2,it} \\ u_{2,it} \end{pmatrix}, \tag{4}$$

with

$$E(\mathbf{v}_{it}) = 0, E(\mathbf{v}_{it}\mathbf{v}_{it}') = \Omega_v = \mathbf{B}^{-1} E(\mathbf{u}_{it}\mathbf{u}_{it}') \mathbf{B}'^{-1} = \mathbf{B}^{-1} \Omega_u \mathbf{B}'^{-1}. \tag{5}$$

For model (1), we assume

Assumption 1. $\{\mathbf{u}_{it}\}$ are i.i.d. over i and t, and are independent of $\boldsymbol{\alpha}_i$ and $\mathbf{y}_{i0}$. We also assume that $E(\mathbf{u}_{it}) = 0$, $E(\mathbf{u}_{it}\mathbf{u}_{it}') = \Omega_u = \begin{pmatrix} \sigma_{u,1}^2 & \sigma_{u,12} \\ \sigma_{u,21} & \sigma_{u,2}^2 \end{pmatrix}$ with finite eighth moment.

Assumption 2. All the eigenvalues of Π are within the unit circle.

Assumption 3. The initial values $\mathbf{y}_{i0} = (I_2 - \Pi)^{-1} \boldsymbol{\xi}_i + \mathbf{w}_{i0}$ for $i = 1, \ldots, N$, where $\mathbf{w}_{i0} = \sum_{s=0}^{\infty} \Pi^s \mathbf{v}_{i,-s}$ is independent of $\boldsymbol{\xi}_i$.

Assumption 4. $\boldsymbol{\alpha}_i$ (or $\boldsymbol{\xi}_i$) are independent of $\mathbf{u}_{jt}$ for all i, j, t and with finite fourth moment.

Assumption 1 is made to simplify the derivation. It can be replaced by heteroskedastic $\mathbf{u}_{it}$ or $\mathbf{u}_{it}$ following a finite order autoregressive process without affecting the general conclusions of the asymptotic distribution of an Alvarez and Arellano (2003) type GMM estimator to be discussed later. Assumption 2 is a stationarity assumption to ensure that the VAR model (2) is stationary. Assumption 3 actually follows from Assumption 1 and 2 through continuous substitution of (2). It is explicitly stated here for ease of exposition in later sections, as in Alvarez and Arellano (2003,

[2] Whatever transformation on the fixed dimension strictly exogenous variables $\mathbf{x}_{it}$ is conducted, the transformed $\mathbf{x}_{it}$ remains uncorrelated with the transformed idiosyncratic time-varying errors of the equation.

[3] Exclusion restriction is required for the identification of the first equation. For discussion of the identification of an equation in a general panel dynamic system, see Hsiao (1982), Hsiao and Zhou (2015). Akashi and Kunitomo (2012) consider a special case when $\gamma_{21} = 0$.

P1126) and Akashi and Kunitomo (2012, P169). Assumption 4 makes no distinction between the fixed or random effects specification because we consider estimators that remove $\boldsymbol{\alpha}_i$ (or $\boldsymbol{\xi}_i$).

For model (1), as discussed by Hsiao (2014) and Moon et al (2015), the presence of individual-specific effects $\boldsymbol{\alpha}_i$ raises the issue of incidental parameters for dynamic systems, certain linear transformation has to be used to remove the the individual-specific effects. We consider two transformations that are most frequently used in applications, (i) the forward orthogonal demeaning (FOD) in Arellano and Bover (1995), Alvarez and Arellano (2003), etc. and (ii) the first difference as in Anderson and Hsiao (1981, 1982), Arellano and Bond (1991), etc.

The FOD transformation is defined as, for $t = 1, ..., T-1$, let $y_{it}^f = c_t \left(y_{it} - \frac{1}{T-t} \sum_{s=t+1}^{T} y_{is} \right)$, $y_{it-1}^f = c_t \left(y_{it-1} - \frac{1}{T-t} \sum_{s=t+1}^{T} y_{is-1} \right)$, and $u_{it}^f = c_t \left(u_{it} - \frac{1}{T-t} \sum_{s=t+1}^{T} u_{is} \right)$, where $c_t^2 = \frac{T-t}{T-t+1}$. Then, for the first equation of (1), we have

$$y_{1,it}^f = \gamma y_{1,it-1}^f + \beta y_{2,it}^f + u_{1,it}^f, \; i = 1, \ldots, N; t = 1, \ldots, T-1. \tag{6}$$

The FOD transformation creates errors that satisfy

$$\begin{aligned} E\left(u_{1,it}^f\right) &= 0, \; E\left(u_{1,it}^{f2}\right) = \sigma_{u,1}^2, \\ E\left(u_{1,it}^f u_{1,is}^f\right) &= 0 \text{ if } t \neq s; E\left(u_{1,it}^f u_{1,js}^f\right) = 0 \text{ if } i \neq j; \end{aligned}$$

as shown by Alvarez and Bover (1995) (also see Alvarez and Arellano (2003) or Hsiao and Zhou (2017)).

Notice that although $u_{1,it}^f$ is i.i.d. over i and t, it is correlated with the transformed regressors, $y_{2,it}^f$ and $y_{1i,t-1}^f$. However, for $0 \leq s \leq t-1$, we have

$$E\left(y_{1,is} u_{1,it}^f\right) = 0, E\left(y_{2,is} u_{1,it}^f\right) = 0.$$

Let

$$\mathbf{z}_{i,t-1} = \left(y_{1,i0}, y_{2,i0}, \ldots, y_{1,it-1}, y_{2,it-1}\right)'. \tag{7}$$

Then $\mathbf{z}_{i,t-1}$ are orthogonal to the transformed error $u_{1,it}^f$ in (6). Also, under Assumption 2, $\mathbf{z}_{i,t-1}$ are correlated with the transformed regressors $y_{2,it}^f$ and $y_{1i,t-1}^f$. In this paper, we consider $\mathbf{z}_{i,t-1}$ in (7) as IVs for model (6).

An alternative transformation widely used in practice is to take the first time difference (FD) (e.g., Anderson and Hsiao (1981, 1982) and Arellano and Bond (1991)). Denote Δ to be the first difference of time series, such that $\Delta y_{it} = y_{it} - y_{it-1}$, for example. Then, the first equation of (1) becomes

$$\Delta y_{1,it} = \gamma \Delta y_{1,it-1} + \beta \Delta y_{2,it} + \Delta u_{1,it}, \; i = 1, \ldots, N, t = 2, \ldots, T. \tag{8}$$

The transformed error, $\Delta u_{1,it}$, follows a first order moving average process. However, for $0 \leq s \leq t-2$, we have

$$E\left(y_{1,is} \Delta u_{1,it}\right) = 0, E\left(y_{2,is} \Delta u_{1,it}\right) = 0.$$

From this, we choose

$$\mathbf{z}_{i,t-2} = \left(y_{1,i0}, y_{2,i0}, \ldots, y_{1,it-2}, y_{2,it-2}\right)', \tag{9}$$

as legitimate IVs for model (8).

The difference between FOD and FD is that the error in (6) is i.i.d over i and t, but the error in (8) follows a first order moving average process. Moreover, (6) uses $(y_{1,i0}, y_{2,i0}, \ldots, y_{1,it-1}, y_{2,it-1})$ as instruments and (8) uses $(y_{1,i0}, y_{2,i0}, \ldots, y_{1,it-2}, y_{2,it-2})$ as instruments.

We will show that the Alvarez and Arellano (2003) type GMM estimator using the FOD or FD transformation is inconsistent if $\frac{T}{N} \to c \neq 0$ as $N \to \infty$. It is consistent and asymptotically biased if[4]

$$\frac{T^3}{N} \to \kappa \neq 0 < \infty \text{ as } (N, T) \to \infty. \tag{10}$$

Condition (10) is also a crucial condition to establish the asymptotic unbiasedness of the JIVE to be discussed later.

3 GMM estimators and their asymptotics

Based on the instrument sets (7) and (9), we consider two types of GMM estimators for the first equation of model (1). The first one is the GMM estimator based on the FOD of model (6) (Alvarez and Arellano (2003), Akashi and Kunitomo (2012))

$$\hat{\boldsymbol{\theta}}_{GMM}^{FOD} = \left(\sum_{t=1}^{T-1} \mathbf{Y}_{t-1 \cdot t}^{f\prime} \mathbf{P}_{t-1} \mathbf{Y}_{t-1 \cdot t}^{f} \right)^{-1} \sum_{t=1}^{T-1} \mathbf{Y}_{t-1 \cdot t}^{f\prime} \mathbf{P}_{t-1} \mathbf{y}_{1,t}^{f}, \tag{11}$$

where $\boldsymbol{\theta} = (\gamma, \beta)'$, $\mathbf{Y}_{t-1 \cdot t}^{f} = \left(\mathbf{y}_{1,t-1}^{f}, \mathbf{y}_{2,t}^{f} \right)$ and $\mathbf{P}_{t-1} = \mathbf{Z}_{t-1} \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{Z}_{t-1}'$ with $\mathbf{Z}_{t-1} = (\mathbf{y}_{1,0}, \mathbf{y}_{2,0}, \ldots, \mathbf{y}_{1,t-1}, \mathbf{y}_{2,t-1})$ and

$$\begin{aligned} \mathbf{y}_{1,t-1}^{f} &= \left(y_{1,1t-1}^{f}, \ldots, y_{1,Nt-1}^{f} \right)', \mathbf{y}_{2,t}^{f} = \left(y_{2,1t}^{f}, \ldots, y_{2,Nt}^{f} \right)', \\ \mathbf{y}_{1,t-1} &= (y_{1,1t-1}, \ldots, y_{1,Nt-1})', \mathbf{y}_{2,t-1} = (y_{2,1t-1}, \ldots, y_{2,Nt-1})'. \end{aligned}$$

The asymptotic distribution of the GMM based on the FD in model (8) is difficult to derive because it involves the inverse of a $(T-1) \times (T-1)$ covariance matrix of a first order moving average process. Therefore, instead of considering the Arellano-Bond type GMM estimator for FD model (8), we consider the Alvarez and Arellano (2003) "crude GMM estimator",[5]

$$\hat{\boldsymbol{\theta}}_{GMM}^{FD} = \left(\sum_{t=2}^{T} \Delta \mathbf{Y}_{t-1 \cdot t}' \mathbf{P}_{t-2} \Delta \mathbf{Y}_{t-1 \cdot t} \right)^{-1} \sum_{t=2}^{T} \Delta \mathbf{Y}_{t-1 \cdot t}' \mathbf{P}_{t-2} \Delta \mathbf{y}_{1,t}, \tag{12}$$

where $\Delta \mathbf{Y}_{t-1 \cdot t} = (\Delta \mathbf{y}_{1,t-1}, \Delta \mathbf{y}_{2,t})$ and $\mathbf{P}_{t-2} = \mathbf{Z}_{t-2} \left(\mathbf{Z}_{t-2}' \mathbf{Z}_{t-2} \right)^{-1} \mathbf{Z}_{t-2}'$ with $\mathbf{Z}_{t-2} = (\mathbf{y}_{1,0}, \mathbf{y}_{2,0}, \ldots, \mathbf{y}_{1,t-2}, \mathbf{y}_{2,t-2})$ and

$$\Delta \mathbf{y}_{1,t-1} = (\Delta y_{1,1t-1}, \ldots, \Delta y_{1,Nt-1})', \Delta \mathbf{y}_{1,t} = (\Delta y_{1,1t}, \ldots, \Delta y_{1,Nt})', \Delta \mathbf{y}_{2,t} = (\Delta y_{2,1t}, \ldots, \Delta y_{2,Nt})'.$$

[4] Similar restriction has also been imposed by Akashi and Kunitomo (2012) and Lee et al (2015).

[5] Note that the crude GMM is not the Arellano-Bond type GMM (Arellano and Bond (1991)) which takes into account the first order dependence of transformed errors (e.g., Hsiao and Zhang (2015)).

3.1 Asymptotics for GMM estimator based on FOD

Although (11) is of the same form as the single equation GMM estimator of the Alvarez and Arellano (2003), there is a fundamental difference between the two. The single equation dynamic panel data model assumes all the regressors are either predetermined or strictly exogenous (with respect to u_{it}). On the other hand, the regressor for the single equation GMM estimator of (1) or (2) could also consist of other joint dependent variables that are correlated with the error in the equation (here, $y_{2,it}$ and $u_{1,it}$). When T is fixed, and N is large, both types of GMM estimators are consistent and asymptotically normally distributed. When T is also large, but $\frac{T}{N} \to c \neq 0 < \infty$ as $N \to \infty$, the single equation GMM estimator remains consistent but asymptotically biased of order $\sqrt{c}$ (Alvarez and Arellano (2003)). However, for GMM estimator of panel dynamic simultaneous equations model, if $\frac{T}{N} \to c \neq 0 < \infty$ as $(N,T) \to \infty$, the GMM estimator is no longer consistent. We need N to be much larger than T to obtain consistency of this GMM estimator.

For the GMM estimator based on FOD (11), Section (A.2) in the appendix shows that

Theorem 1 *Under Assumptions 1-4 and the condition (10), the GMM estimator (11) is consistent and asymptotically distributed as*

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{GMM}^{FOD} - \boldsymbol{\theta}\right) \to_d N\left(\mathbf{b}_0^f, \sigma_{u,1}^2\left(\mathbf{D}'\boldsymbol{\Gamma}_0\mathbf{D}\right)^{-1}\right), \tag{13}$$

where

$$\mathbf{b}_0^f = -\left(\mathbf{D}'\boldsymbol{\Gamma}_0\mathbf{D}\right)^{-1}\begin{pmatrix} 0 \\ \sigma_{u,12} \end{pmatrix}\sqrt{\kappa},$$

denotes the asymptotic bias and

$$\mathbf{D} = \begin{pmatrix} 1 & \beta\gamma_{22} \\ 0 & \gamma_{22} \end{pmatrix}, \Gamma_0 = \sum_{s=0}^{\infty}\Pi^s\mathbf{B}^{-1}\Omega_u\mathbf{B}'^{-1}\Pi^{s\prime},$$

with Π *and* $\mathbf{B}$ *are given in (3) and* κ *is given in (10).*

Remark 1 *It follows from Theorem 1 that if* $\kappa = 0$*, the FOD GMM estimator* $\hat{\boldsymbol{\theta}}_{GMM}^{FOD}$ *is asymptotically unbiased. However, if* $\kappa \neq 0$*, the* $\hat{\boldsymbol{\theta}}_{GMM}^{FOD}$ *is asymptotically biased of order* $\sqrt{\frac{T^3}{N}}$*, which is the same order as in Akashi and Kunitomo (2012). If* $\frac{T}{N} \to c \neq 0$ *as* $(N,T) \to \infty$ *considered in Alvarez and Arellano (2003), the GMM estimator (11) is inconsistent, i.e.,*

$$\hat{\boldsymbol{\theta}}_{GMM}^{FOD} - \boldsymbol{\theta} = O_p\left(\frac{T}{N}\right).$$

3.2 Asymptotics for crude GMM estimator based on FD

For the crude GMM estimator (12), it is shown in Section (A.3) of the appendix that

Theorem 2 *Under Assumptions 1-4 and the restriction (10), the crude GMM estimator (12) is asymptotically distributed as*

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{GMM}^{FD} - \boldsymbol{\theta}\right) \to_d N\left(\mathbf{b}_0^{\Delta}, \sigma_{u,1}^2\mathbf{D}'^{-1}\left(I_2 - \Pi'\right)^{-1}\left[\Gamma_0^{-1}\left(I_2 - \Pi\right) + \left(I_2 - \Pi'\right)\Gamma_0^{-1}\right]\left(I_2 - \Pi\right)^{-1}\mathbf{D}^{-1}\right). \tag{14}$$

where $\mathbf{b}_0^{\Delta} = \left[\mathbf{D}'\left(I_2 - \Pi\right)\Gamma_0\left(I_2 - \Pi'\right)\mathbf{D}\right]^{-1}\tilde{\mathbf{b}}_0$,

$$\tilde{\mathbf{b}}_0 = -\left[\mathbf{D}'\begin{pmatrix}\sigma_{u,1}^2 + \beta\sigma_{u,12} \\ \sigma_{u,21}\end{pmatrix} - \begin{pmatrix}0 \\ 2\sigma_{u,21}\end{pmatrix}\right]\sqrt{\kappa}. \tag{15}$$

and $\mathbf{D}$, Π *and* Γ_0 *are defined in Theorem 1.*

Remark 2 *As in the FOD case, the crude GMM estimator (12) is inconsistent when* $\frac{T}{N} \to c \neq 0$. *Even if* $\frac{T}{N} \to 0$ *as* $(N, T) \to \infty$ *with* $\frac{T^3}{N} \to \kappa \neq 0$, *it is asymptotically biased of order* $\sqrt{\frac{T^3}{N}}$.

Remark 3 *Due to the complicated variance-covariance matrix of the crude GMM based on FD, it is difficult to directly compare the asymptotic efficiency for these two GMM estimators based on FOD (estimator (11)) and FD (estimator (12)) for the panel dynamic simultaneous equations models. We note that the crude GMM based on FD (estimator (12)) is not a GMM estimator, and it is expected to have a larger covariance matrix than the GMM based on FOD. Since GMM based FOD or FD uses identical number of moment conditions with almost identical IVs, we expect that (12) is not as efficient as (11). In a similar context, Lee et al (2015) find that GMM based on FOD is asymptotically more efficient than the crude GMM based on FD. Moreover, from the simulation below, we observe that the iqr (inter-quantile range) of GMM based on FOD is much smaller than that of the crude GMM based FD. Thus, for models of the form (1), we conjecture that the GMM based on FOD is more efficient than the crude GMM based on FD.*

Remark 4 *As noted by a referee, the GMM estimator of the reduced form parameters,* Π, *are consistent if* $\frac{T}{N} \to c \neq 0 < \frac{1}{2}$ *as* $N \to \infty$ *by a straightforward generalization of Alvarez and Arellano (2003). However, the structural form parameters could be ratios of the reduced form parameters (e.g., here* $\beta = \frac{\pi_{12}}{\pi_{22}}$*). Moreover, if a structural equation is over-identified, there is an issue of how to account for the complicated nonlinear restrictions in obtaining efficient estimators of structural form parameters (e.g., Intrilgator et al (1996)). These issues are complicated and deserve an independent study.*

4 JIVE and its asymptotics

We note that the GMM estimator (11) or (12) can also be viewed as finding optimal instruments $\mathbf{W}_{it}$ or $\tilde{\mathbf{W}}_{it}$ that satisfy $E\left(\mathbf{W}_{it}u_{1,it}^f\right) = 0$ and $E\left(\tilde{\mathbf{W}}_{it}\Delta u_{1,it}\right) = 0$. In sample analogue, we have

$$\frac{1}{N(T-1)}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\mathbf{W}_{it}u_{1,it}^f = 0, \tag{16}$$

or

$$\frac{1}{N(T-1)}\sum_{t=2}^{T}\sum_{i=1}^{N}\tilde{\mathbf{W}}_{it}\Delta u_{1,it} = 0, \tag{17}$$

where

$$\mathbf{W}_{it} = \mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\left[\sum_{j=1}^{N}\mathbf{z}_{jt-1}\mathbf{y}_{j,t-1}^{f\prime}\right], \tilde{\mathbf{W}}_{it} = \mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\left[\sum_{j=1}^{N}\mathbf{z}_{jt-2}\Delta\mathbf{y}_{j,t-1}'\right], \tag{18}$$

with $\mathbf{y}^{f\prime}_{j,t-1} = \left(y^f_{1,it-1}, y^f_{2,it}\right)$ and $\Delta\mathbf{y}'_{j,t-1} = \left(\Delta y_{1,it-1}, \Delta y_{2,it}\right).$

Under the assumption that $\mathbf{u}_{it}$ are i.i.d over i, and t, so are $\mathbf{u}^f_{it}$, therefore,

$$E\left[\left(\sum_{j=1}^{N}\mathbf{z}_{jt-1}\mathbf{y}^{f\prime}_{j,t-1}\right)u^f_{1,it}\right] = E\left(\mathbf{z}_{it-1}\mathbf{y}^{f\prime}_{i,t-1}u^f_{1,it}\right) \neq 0, \tag{19}$$

or

$$E\left[\left(\sum_{j=1}^{N}\mathbf{z}_{jt-2}\Delta\mathbf{y}'_{j,t-1}\right)\Delta u_{1,it}\right] = E\left(\mathbf{z}_{it-2}\Delta\mathbf{y}'_{i,t-1}\Delta u_{1,it}\right) \neq 0. \tag{20}$$

Equation (19) or (20) is the source of asymptotic bias for the GMM estimator (11) or (12). However, if we remove the ith individual's observation in the construction of $\mathbf{W}_{it}$ or $\tilde{\mathbf{W}}_{it}$, so

$$\mathbf{W}^*_{it} = \mathbf{z}'_{it-1}\left(\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\right)^{-1}\left[\sum_{j\neq i}^{N}\mathbf{z}_{jt-1}\mathbf{y}^{f\prime}_{j,t-1}\right],\ \tilde{\mathbf{W}}^*_{it} = \mathbf{z}'_{it-2}\left(\mathbf{Z}'_{t-2}\mathbf{Z}_{t-2}\right)^{-1}\left[\sum_{j\neq i}^{N}\mathbf{z}_{jt-2}\Delta\mathbf{y}'_{j,t-1}\right], \tag{21}$$

then

$$E\left(\mathbf{W}^*_{it}u^f_{1,it}\right) = 0 \text{ or } E\left(\tilde{\mathbf{W}}^*_{it}\Delta u_{1,it}\right) = 0. \tag{22}$$

Thus, using $\mathbf{W}^*_{it}$ or $\tilde{\mathbf{W}}^*_{it}$ as IV removes the source of asymptotic bias. That's how JIVE corrects the asymptotic bias of GMM estimator.

4.1 JIVE based on FOD

The JIVE for (11) is defined as

$$\begin{aligned}\hat{\boldsymbol{\theta}}^{FOD}_{JIVE} &= \left(\sum_{t=1}^{T-1}\left[\sum_{j=1}^{N}\sum_{i\neq j}\left(\mathbf{y}^{f\prime}_{i,t-1\cdot t}\mathbf{z}'_{it-1}\right)\left(\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{jt-1}\mathbf{y}^f_{j,t-1\cdot t}\right]\right)^{-1} \\ &\quad\times\sum_{t=1}^{T-1}\left[\sum_{j=1}^{N}\sum_{i\neq j}\left(\mathbf{y}^{f\prime}_{i,t-1\cdot t}\mathbf{z}'_{it-1}\right)\left(\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{jt-1}y^f_{1,jt}\right],\end{aligned} \tag{23}$$

where $\mathbf{y}^f_{i,t-1\cdot t} = \left(y^f_{1,it-1}, y^f_{2,it}\right)$ and $\mathbf{y}^f_{j,t-1\cdot t} = \left(y^f_{1,jt-1}, y^f_{2,jt}\right).$

Remark 5 *When $T-1=1$, the above JIVE (23) is identical to the JIVE2 proposed in Angrist et al (1999, P61) for cross-sectional models.*

For this JIVE (23), it is shown in the Section (A.4) of the appendix that

Theorem 3 *Under Assumptions 1-4 and the restriction (10), the JIVE estimator (23) is asymptotically distributed as*

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}^{FOD}_{JIVE} - \boldsymbol{\theta}\right) \to_d N\left(0, \sigma^2_{u,1}\mathbf{D}'^{-1}\boldsymbol{\Gamma}_0^{-1}\mathbf{D}^{-1}\right), \tag{24}$$

where $\mathbf{D}$, Π and Γ_0 are defined in Theorem 1.

Remark 6 *As can be seen from (24), the JIVE of (1) using FOD is asymptotically unbiased. Additionally, the JIVE is as efficient as the original GMM estimator (11), i.e., there is no efficiency loss for the JIVE.*[6]

Remark 7 *As pointed out by a referee, the purpose of this JIVE is to construct an instrument for observation (i,t) which does not involve any observation dependent with (i,t). If there is spatial-temporal dependence in the observation, then the JIVE bias reduction property would vanish or diminish.*[7]

Remark 8 *As noted by a referee that one of the advantage of JIVE is its ability to handle heteroskedastic disturbance. This is indeed the case because the construction of JIVE for the (i,t) observation excludes the use of ith individual observations. However, the asymptotic covariance matrix of the JIVE estimator no longer possesses the neat form as in (24).*[8]

Standard Error Computation

Under Assumption 1 of homoskedasticity, given that the JIVE estimators are consistent and asymptotically unbiased, a consistent estimator for the variance of $\hat{\boldsymbol{\theta}}^{FOD}_{JIVE} = \left(\hat{\beta}^{FOD}_{JIVE}, \hat{\gamma}^{FOD}_{JIVE}\right)'$ can be obtained by replacing $\sigma^2_{u,1}$, $\mathbf{D}$ and $\boldsymbol{\Gamma}_0$ in (24) by their estimates $\hat{\sigma}^2_{u,1}$, $\hat{\mathbf{D}}$ and $\hat{\boldsymbol{\Gamma}}_0$, respectively. Let

$$\hat{\sigma}^2_{u,1} = \frac{1}{NT}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left(\hat{u}^f_{1,it}\right)^2, \hat{\sigma}^2_{u,2} = \frac{1}{NT}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left(\hat{u}^f_{2,it}\right)^2 \text{ and } \hat{\sigma}_{u,12} = \frac{1}{NT}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\hat{u}^f_{1,it}\hat{u}^f_{2,it},$$

with $\hat{u}^f_{1,it} = y^f_{1,it} - \hat{\beta}^{FOD}_{JIVE}y^f_{2,it} - \hat{\gamma}^{FOD}_{JIVE}y^f_{1,it-1}$ and $\hat{u}^f_{2,it} = y^f_{2,it} - \hat{\gamma}_{21}y_{2,it-1} - \hat{\gamma}_{22}y_{2,it-1}$ given the estimators of γ_{21} and γ_{22}, we have

$$\begin{aligned}
\hat{\mathbf{D}} &= \begin{pmatrix} 1 & \hat{\beta}\hat{\gamma}_{22} \\ 0 & \hat{\gamma}_{22} \end{pmatrix}, \hat{\Pi} = \begin{pmatrix} \hat{\pi}_{11} & \hat{\pi}_{12} \\ \hat{\pi}_{21} & \hat{\pi}_{22} \end{pmatrix} = \begin{pmatrix} \hat{\gamma} + \hat{\beta}\hat{\gamma}_{21} & \hat{\beta}\hat{\gamma}_{22} \\ \hat{\gamma}_{21} & \hat{\gamma}_{22} \end{pmatrix}, \\
\hat{\mathbf{B}} &= \begin{pmatrix} 1 & \hat{\beta} \\ 0 & 1 \end{pmatrix}, \hat{\boldsymbol{\Gamma}}_0 = \sum_{s=0}^{\infty}\hat{\Pi}^s\hat{\mathbf{B}}^{-1}\hat{\Omega}_u\hat{\mathbf{B}}'^{-1}\hat{\Pi}^{s\prime},
\end{aligned}$$

by using the above estimators $\hat{\sigma}^2_{u,1}, \hat{\sigma}^2_{u,2}$ and $\hat{\sigma}_{u,12}$.

4.2 JIVE based on FD

For model (8) based on FD, we can define the JIVE as

$$\begin{aligned}
\hat{\boldsymbol{\theta}}^{FD}_{JIVE} &= \left(\sum_{t=2}^{T}\sum_{j=1}^{N}\sum_{i\neq j}\Delta\mathbf{y}'_{i,t-1\cdot t}\mathbf{z}'_{it-2}\left(\mathbf{Z}'_{t-2}\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{jt-2}\Delta\mathbf{y}_{j,t-1\cdot t}\right)^{-1} \qquad (25) \\
&\quad \times \sum_{t=2}^{T}\sum_{j=1}^{N}\sum_{i\neq j}\Delta\mathbf{y}'_{i,t-1\cdot t}\mathbf{z}'_{it-2}\left(\mathbf{Z}'_{t-2}\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{jt-2}\Delta y_{1,jt},
\end{aligned}$$

[6] In a different context, Hahn and Newey (2004) have also established that jackknife correction does not affect the asymptotic variance.

[7] When $u_{1,it}$ is serially correlated, FOD loses its attraction. It does not yield i.i.d errors. Also, unless the pattern of serial correlation is known, lagged variables may not be legitimate IVs. Neither is (11) a GMM estimator. For the application of JIVE in a simple univariate dynamic panel model with errors following a first order moving average process, see Lee at al (2015).

[8] Under certain assumptions, one can show that the GMM is asymptotically biased of order $\sqrt{\frac{T^3}{N}}$. However, the exact bias is complicated and depends on $\sigma_{u,12i}$ where $\sigma_{u,12i} = Cov\left(u_{1,it}, u_{2,it}\right)$ $(i = 1, \ldots, N)$.

where $\Delta \mathbf{y}_{i,t-1\cdot t} = (\Delta y_{1,it-1}, \Delta y_{2,it})$ and $\Delta \mathbf{y}_{j,t-1\cdot t} = (\Delta y_{1,jt-1}, \Delta y_{2,jt})$.

For this JIVE (25), it is shown in the Section (A.5) of the appendix that

Theorem 4 *Under Assumptions 1-4 and the restriction (10), the JIVE estimator (25) is asymptotically distributed as*

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}^{FD}_{JIVE} - \boldsymbol{\theta}\right) \to_d N\left(0, \sigma^2_{u,1}\mathbf{D}'^{-1}\left(I_2 - \Pi'\right)^{-1}\left[\Gamma_0^{-1}\left(I_2 - \Pi\right) + \left(I_2 - \Pi'\right)\Gamma_0^{-1}\right]\left(I_2 - \Pi\right)^{-1}\mathbf{D}^{-1}\right), \tag{26}$$

where $\mathbf{D}$, Π *and* Γ_0 *are defined in Theorem 1.*

Remark 9 *As can be seen from (26), the JIVE indeed corrects the asymptotic bias of the crude GMM estimator of (1) using FD, as long as* $T^3/N < \infty$ *as* $N \to \infty$ *despite the number of instruments increasing with* T.

Standard Error Computation

The standard error computation for JIVE based on FD is similar to the case where FOD is used. The only difference is in the estimation of $\sigma^2_{u,1}$, $\sigma^2_{u,2}$ and $\sigma_{u,12}$, which are based on the residuals from the JIVE using FD. For instance, we can estimate $\sigma^2_{u,1}$ by $\hat{\sigma}^2_{u,1} = \frac{1}{2}\frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}\left(\Delta\hat{u}_{1,it}\right)^2$ where $\Delta\hat{u}_{1,it} = \Delta y_{1,it} - \hat{\beta}^{FD}_{JIVE}\Delta y_{2,it} - \hat{\gamma}^{FD}_{JIVE}\Delta y_{1,it-1}$. Similarly for the estimation of $\sigma^2_{u,2}$ and $\sigma_{u,12}$, and estimation of the remaining terms in (26).

5 Monte Carlo Simulations

In this section, we study the finite sample properties of the GMM and JIVE estimators considered above. The data generating process is

$$\begin{aligned} y_{1,it} &= \gamma y_{1,it-1} + \beta y_{2,it} + \alpha_{1,i} + u_{1,it}, \\ y_{2,it} &= \gamma_{21} y_{1,it-1} + \gamma_{22} y_{1,it-1} + \alpha_{2,i} + u_{2,it}, \end{aligned}$$

with $\gamma = 0.5$, $\beta = 0.5$, $\gamma_{21} = 0.2$, $\gamma_{22} = 0.6$.[9] For the individual-specific effects, we assume $\alpha_{1,i} \sim IIDN(0,1)$ and $\alpha_{2,i} \sim IIDN(0,2)$ for $i = 1, 2, \ldots, N$. For the error terms, we assume that

$$\begin{pmatrix} u_{1,it} \\ u_{2,it} \end{pmatrix} \sim_{iid} N\left(0, \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}\right),$$

for $i = 1, \ldots, N, t = 1, \ldots, T$. Also, $(\alpha_{1,i}, \alpha_{2,i})'$ and $(u_{1,it}, u_{2,it})'$ are independent over i and t. We consider different combinations of N, T by letting $N = 1000$, 2000 and 5000, and $T = 10, 25$ and 50. We generate $T + 100$ observations, and the first 100 observations are discarded. The number of replication is set at 1000 times.

We consider the GMM estimation as well as the JIVE proposed in the paper to estimate γ and β in the above DGP. For comparison, we also consider the LIML estimation considered by Akashi and Kunitomo (2012)[10], the regularized JIVE (RJIVE) proposed by Hansen and Kozbur (2014)[11].

[9] Under this specification, we have $\Pi = \mathbf{B}^{-1}\begin{pmatrix} \gamma & 0 \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} 0.6 & 0.3 \\ 0.2 & 0.6 \end{pmatrix}$, whose eigenvalues are given by 0.8449 and 0.3551, so Assumption 2 is satisfied.

[10] As shown by Akashi and Kunitomo (2012, P170), the LIML estimation of model (1) is asymptotically biased of order $\sqrt{\frac{T}{N}}$.

[11] We follow Hansen and Kozbur (2014, P297) to choose the penalty matrix in the RJIVE, i.e., we set the penalty to $\sqrt{K}$ where K is the number of instruments used in the estimation.

Because our simulated model is exactly identified, unique estimates of β and γ can be obtained from $\beta = \frac{\pi_{12}}{\pi_{22}}$ and $\hat{\gamma} = \pi_{11} - \beta\pi_{21}$. We also consider a referee's suggestion by first estimating Π by the GMM estimation based on the reduced form using all available instruments, and then solve for β and γ[12]. To eliminate α_{1i}, we use both FOD and FD transformations. We calculate the mean and bias of the estimates, iqr (inter-quantile range), and size for these estimators. Our estimation results are summarized in the tables 1-4.

Several interesting findings can be observed from the simulation results. First, we note that there is significant bias for GMM estimators of both γ and β using either FOD or FD transformation to eliminate the individual effects. Moreover, the size is severely distorted for GMM estimators and the coverage ratio is quite poor. On the other hand, if one considers GMM estimator based on the reduced form of the simultaneous equations models, the GMM estimator based on FOD transformation appears to perform better in the exactly identified case and has smaller size distortion than the GMM based on FOD or FD (Table 1-2). However, the GMM estimator based on the reduced form using FD transformation still shows significant bias and distorted size (Table 3-4). Second, for the JIVE for both γ and β, the bias is almost negligible in both FOD and FD cases, which suggests that the JIVE indeed corrects the bias of GMM estimators as desired. Moreover, the actual size for JIVE of γ and β is very close to the nominal value of 5% significance level. Alternatively, for the JIVE, one can consider the regularized JIVE by Hansen and Kozbur (2014), which is also found to be asymptotically unbiased and has correct size.[13] One can also observe that the iqr (inter quantile range) of JIVE estimators are quite close to the GMM estimators for large N. For example, when $N = 5000$, the iqr of JIVE for both γ and β are quite close to the iqr of GMM estimators in the simulation, which is evident of the fact that JIVE doesn't inflate the variance as shown in the paper. Finally, if one considers LIML of Akashi and Kunitomo (2012), it is observed that LIML estimation is indeed asymptotically unbiased for both γ and β using either FOD or FD transformation, and has correct size. This is because LIML is asymptotically biased of order $\sqrt{\frac{T}{N}}$, while in our simulation, we have $\frac{T}{N} \to 0$, which in turn leads LIML to be asymptotically unbiased.[14]

Finally, for comparison, we draw the empirical densities of $\sqrt{NT}\left(\hat{\gamma} - \gamma\right)$ and $\sqrt{NT}\left(\hat{\beta} - \beta\right)$ in Fig 1 and 2, respectively, of the GMM estimators and the JIVE using FOD and FD transformation for the DGP when $N = 2000$ and $T = 25$. It is clear that the empirical densities of JIVE and LIML estimators are centered at zero and are normally distributed, while the empirical densities of GMM estimators are not centered at zero. In all, the findings from simulation confirms our theoretical findings in the paper.

[12] By following Alvarez and Arellano (2003), it can be shown that the GMM using FOD is consistent and asymptotically normal as long as $\frac{T}{N} \to c$ as $(N, T) \to \infty$, but the GMM using FD is inconsistent if $\frac{T}{N} \to c$ and asymptotically biased of order $\sqrt{\frac{T^3}{N}}$. See Remark 4 for more discussion on GMM based on reduced form.

[13] It should be noted that even if regularized JIVE behaves similarly to the JIVE proposed in this paper, it is quite computational extensive, and it becomes more computationally demanding if T is large. For example, for the simulation when $N = 1000$ and $T = 25$ with 1000 replications, the cpu time for the JIVE and regularized JIVE are 17491s and 28293s, respectively. When $T = 50$, it takes days for regularized JIVE to finish.

[14] Even if the LIML behaves similarly to JIVE in our designs, the JIVE is relatively easy to implement, while LIML requires extra work to obtain the estimators, such as solving the characteristic function to get the smallest root (Akashi and Kunitomo (2012, P169)).

Table 1 Simulation results of estimation of γ based FOD

	N	1000					2000					5000				
T		GMM	JIVE	LIML	RJIVE	GMMR	GMM	JIVE	LIML	RJIVE	GMMR	GMM	JIVE	LIML	RJIVE	GMMR
	estim	0.4569	0.5001	0.4968	0.5001	0.4897	0.4775	0.5006	0.4988	0.5006	0.4951	0.4911	0.5007	0.5000	0.5007	0.4984
10	bias	-0.0431	0.0001	-0.0032	0.0001	-0.0103	-0.0225	0.0006	-0.0012	0.0006	-0.0049	-0.0089	0.0007	0.0000	0.0007	-0.0016
	iqr	0.0329	0.0389	0.0368	0.0390	0.0358	0.0226	0.0253	0.0240	0.0253	0.0242	0.0166	0.0173	0.0167	0.0172	0.0165
	size	40.6%	5%	4.9%	5.3%	7%	23.9%	5.5%	5%	5.4%	5.2%	11.5%	4.4%	4.3%	4.4%	4.5%
	estim	0.4436	0.5004	0.4979	0.5004	0.4948	0.4690	0.5002	0.4990	0.5002	0.4973	0.4868	0.5000	0.4995	0.5000	0.4989
25	bias	-0.0564	0.0004	-0.0021	0.0004	-0.0052	-0.0310	0.0002	-0.0010	0.0002	-0.0027	-0.0132	0.0000	-0.0005	0.0000	-0.0011
	iqr	0.0147	0.0188	0.0182	0.0188	0.0169	0.0116	0.0130	0.0130	0.0130	0.0124	0.0070	0.0075	0.0071	0.0075	0.0072
	size	99.9%	5.6%	5.5%	5.4%	5.6%	95.3%	5.8%	5.6%	5.7%	5.9%	68.8%	5.3%	4.9%	5.3%	5.3%
	estim	0.4195	0.5001	0.4980	0.5001	0.4958	0.4538	0.5001	0.4991	0.5001	0.4980	0.4797	0.5000	0.4996	0.5000	0.4992
50	bias	-0.0805	0.0001	-0.0020	0.0001	-0.0042	-0.0462	0.0001	-0.0009	0.0001	-0.0020	-0.0203	0.0000	-0.0004	0.0000	-0.0008
	iqr	0.0082	0.0123	0.0114	0.0123	0.0102	0.0067	0.0080	0.0079	0.0080	0.0076	0.0043	0.0047	0.0047	0.0047	0.0045
	size	100%	5.3%	5.4%	5.4%	8.5%	100%	4.7%	5.4%	4.7%	6.8%	100%	4.5%	4.9%	4.5%	6%

Notes: 1. "GMM" refers to GMM estimation, "JIVE" refers to Jackknife instrumental variable estimation, "LIML" refers to the limited information likelihood estimation of Akashi and Kunitomo (2012), "RJIVE" refers to the regularized JIVE of Hansen and Kozbur (2014), and "GMMR" refers to GMM estimation based on the reduced form (2) using all available instruments.

2. "iqr" refers to inter quantile range (75% quantile-25% quantile).

3. size is calculated for $H_0 : \gamma = 0.5$.

Table 2 Simulation results of estimation of β based FOD

	N	1000					2000					5000				
T		GMM	JIVE	LIML	RJIVE	GMMR	GMM	JIVE	LIML	RJIVE	GMMR	GMM	JIVE	LIML	RJIVE	GMMR
	estim	0.5823	0.4955	0.4961	0.4955	0.4759	0.5450	0.4982	0.4989	0.4982	0.4881	0.5189	0.4993	0.4995	0.4993	0.4950
10	bias	0.0823	-0.0045	-0.0039	-0.0045	-0.0241	0.0450	-0.0018	-0.0011	-0.0018	-0.0119	0.0189	-0.0007	-0.0005	-0.0007	-0.0050
	iqr	0.0645	0.0847	0.0869	0.0848	0.0793	0.0500	0.0567	0.0557	0.0568	0.0533	0.0296	0.0318	0.0318	0.0318	0.0296
	size	39%	5.3%	5.1%	4.7%	6.7%	27%	4.4%	4.6%	4.4%	5.8%	13.1%	5.3%	4.7%	5.3%	5.8%
	estim	0.5975	0.4983	0.4984	0.4983	0.4934	0.5543	0.4995	0.4995	0.4995	0.4969	0.5231	0.4996	0.4997	0.4996	0.4985
25	bias	0.0975	-0.0017	-0.0016	-0.0017	-0.0066	0.0543	-0.0005	-0.0005	-0.0005	-0.0031	0.0231	-0.0004	-0.0003	-0.0004	-0.0015
	iqr	0.0240	0.0320	0.0320	0.0319	0.0286	0.0190	0.0224	0.0216	0.0223	0.0210	0.0124	0.0130	0.0130	0.0131	0.0130
	size	100%	4.8%	4.4%	4.7%	6.5%	97.7%	4.5%	4.8%	4.5%	5.1%	73.4%	4.1%	3.8%	4.1%	4.6%
	estim	0.6336	0.4999	0.5000	0.4999	0.4987	0.5771	0.4998	0.4997	0.4998	0.4988	0.5342	0.5001	0.5001	0.5001	0.4997
50	bias	0.1336	-0.0001	0.0000	-0.0001	-0.0013	0.0771	-0.0002	-0.0003	-0.0002	-0.0012	0.0342	0.0001	0.0001	0.0001	-0.0003
	iqr	0.0134	0.0198	0.0179	0.0199	0.0164	0.0103	0.0128	0.0124	0.0129	0.0116	0.0071	0.0078	0.0078	0.0078	0.0073
	size	100%	4.7%	6.1%	4.8%	4.8%	100%	4.8%	4.9%	4.8%	5.9%	100%	4.8%	5%	4.8%	5%

Notes: size is calculated for $H_0 : \beta = 0.5$. Also see notes of Table 1.

Table 3 Simulation results of estimation of γ based FD

	N	1000					2000					5000				
T		GMM	JIVE	LIML	RJIVE	GMMR	GMM	JIVE	LIML	RJIVE	GMMR	GMM	JIVE	LIML	RJIVE	GMMR
	estim	0.3863	0.5035	0.5006	0.5035	0.4151	0.4368	0.5018	0.5002	0.5018	0.4546	0.4730	0.5007	0.5004	0.5007	0.4812
10	bias	-0.1137	0.0035	0.0006	0.0035	-0.0849	-0.0632	0.0018	0.0002	0.0018	-0.0454	-0.0270	0.0007	0.0004	0.0007	-0.0188
	iqr	0.0409	0.0625	0.0515	0.0626	0.0480	0.0342	0.0425	0.0361	0.0424	0.0368	0.0216	0.0239	0.0219	0.0239	0.0217
	size	95%	5.1%	4.8%	5.4%	63.9%	73%	5.3%	5%	5.4%	43.1%	38.1%	5%	5.3%	5%	21.7%
	estim	0.3087	0.5022	0.5001	0.5022	0.3642	0.3855	0.5010	0.5003	0.5010	0.4246	0.4475	0.5003	0.5003	0.5003	0.4677
25	bias	-0.1913	0.0022	0.0001	0.0022	-0.1358	-0.1145	0.0010	0.0003	0.0010	-0.0754	-0.0525	0.0003	0.0003	0.0003	-0.0323
	iqr	0.0222	0.0433	0.0308	0.0437	0.0254	0.0168	0.0249	0.0201	0.0249	0.0193	0.0109	0.0128	0.0118	0.0128	0.0116
	size	100%	4.6%	5.8%	4.5%	100%	100%	4.2%	4.3%	4.2%	100%	100%	5.2%	5.4%	5.2%	95.8%
	estim	0.2239	0.5008	0.4999	0.5008	0.2979	0.3208	0.5001	0.5002	0.5001	0.3822	0.4110	0.4999	0.4999	0.4999	0.4476
50	bias	-0.2761	0.0008	-0.0001	0.0008	-0.2021	-0.1792	0.0001	0.0002	0.0001	-0.1178	-0.0890	-0.0001	-0.0001	-0.0001	-0.0524
	iqr	0.0127	0.0341	0.0220	0.0342	0.0169	0.0104	0.0188	0.0147	0.0001	0.0127	0.0070	0.0083	0.0083	0.0083	0.0075
	size	100%	5.6%	5.2%	5.5%	100%	100%	4.9%	4.4%	4.9%	100%	100%	5.5%	5.6%	5.5%	100%

Notes: Size is calculated for $H_0 : \gamma = 0.5$. Also see notes of Table 1.

Table 4 Simulation results of estimation of β based FD

	N	1000					2000					5000				
T		GMM	JIVE	LIML	RJIVE	GMMR	GMM	JIVE	LIML	RJIVE	GMMR	GMM	JIVE	LIML	RJIVE	GMMR
	estim	0.6839	0.4856	0.4919	0.4855	0.2719	0.6135	0.4931	0.4965	0.4931	0.3803	0.5551	0.4999	0.5005	0.4999	0.4523
10	bias	0.1839	-0.0144	-0.0081	-0.0145	-0.2281	0.1135	-0.0069	-0.0035	-0.0069	-0.1197	0.0551	-0.0001	0.0005	-0.0001	-0.0477
	iqr	0.0924	0.1731	0.1615	0.1732	0.1614	0.0700	0.1033	0.0907	0.1040	0.0924	0.0425	0.0503	0.0493	0.0503	0.0521
	size	74%	5.4%	5.6%	5.2%	41.9%	57.9%	4.5%	4.8%	4.5%	38.5%	40.2%	4.5%	5.3%	4.5%	24.2%
	estim	0.7597	0.4948	0.4987	0.4948	0.2174	0.6784	0.4978	0.4992	0.4978	0.3453	0.5918	0.4997	0.4998	0.4997	0.4337
25	bias	0.2597	-0.0052	-0.0013	-0.0052	-0.2826	0.1784	-0.0022	-0.0008	-0.0022	-0.1547	0.0918	-0.0003	-0.0002	-0.0003	-0.0663
	iqr	0.0389	0.1017	0.0876	0.1023	0.0918	0.0305	0.0539	0.0470	0.0541	0.0505	0.0212	0.0280	0.0260	0.0279	0.0247
	size	100%	5.3%	5.9%	5.3%	99.7%	100%	5%	5.8%	5.1%	99.8%	100%	5.1%	4.9%	5%	96.5%
	estim	0.8254	0.4985	0.5000	0.4986	0.1482	0.7484	0.4992	0.4990	0.4992	0.2941	0.6428	0.5003	0.5002	0.5003	0.4083
50	bias	0.3254	-0.0015	0.0000	-0.0014	-0.3518	0.2484	-0.0008	-0.0010	-0.0008	-0.2059	0.1428	0.0003	0.0002	0.0003	-0.0917
	iqr	0.0236	0.0877	0.0661	0.0879	0.0710	0.0179	0.0423	0.0378	0.0424	0.0360	0.0127	0.0207	0.0182	0.0208	0.0171
	size	100%	5.1%	4.7%	5.1%	100%	100%	5.5%	5.3%	5.5%	100%	100%	5.3%	4.5%	5.3%	100%

Notes: size is calculated for $H_0 : \beta = 0.5$. Also see notes of Table 1.

Fig 1. Empirical densities of GMM and JIVE estimators for $\sqrt{NT}\,(\hat{\gamma}-\gamma)$ when $N=2000$ and $T=25$

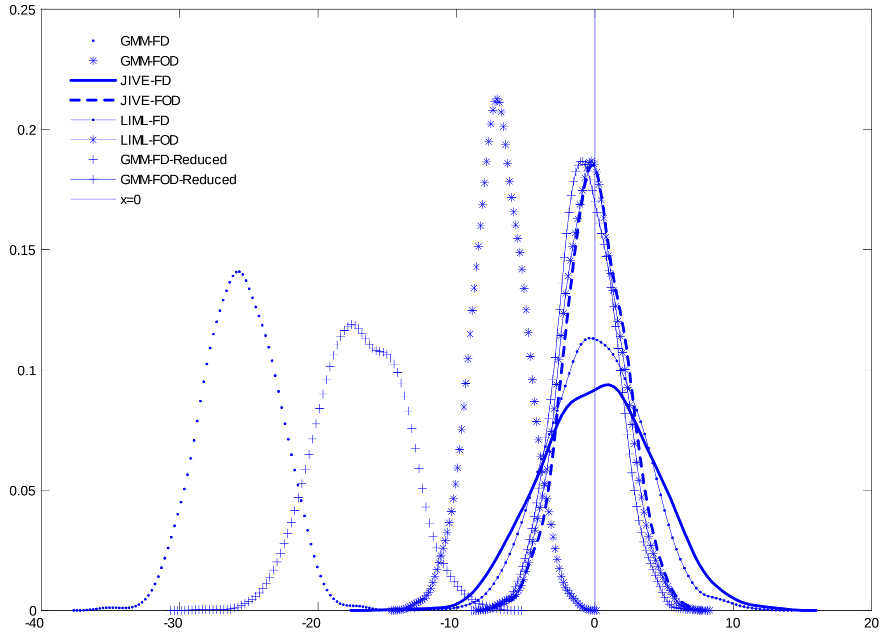
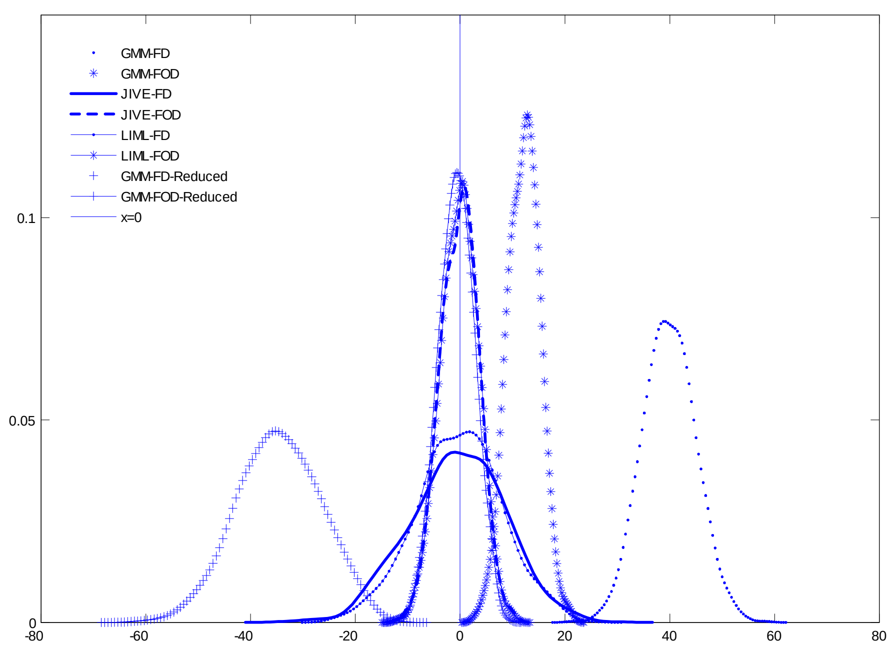


Fig 2. Empirical densities of GMM and JIVE estimators for $\sqrt{NT}\left(\hat{\beta}-\beta\right)$ when $N=2000$ and $T=25$



6 Conclusion

In this paper we investigate the statistical properties of the GMM estimators for linear panel dynamic simultaneous equations models. Using the alternative asymptotics $(N, T) \to \infty$ with $\frac{T^3}{N} \to \kappa \neq 0 < \infty$, we characterize the many IVs bias of the GMM estimators. To reduce the bias of the GMM estimators, we consider the JIVE and establish its asymptotics. Monte Carlo simulations show that the JIVE estimator can eliminate the asymptotic bias, hence allowing us to obtain valid statistical inference. It would be very interesting to extend the above JIVE procedure to models with heteroskedastic errors, as in Chao et al (2012); and to models with spatio-temporal dependence, as in Lee and Yu (2014). We leave these topics for future research.

References

[1] Akashi, K., and N. Kunitomo, 2012, Some properties of the LIML estimator in a dynamic panel structural equation, *Journal of Econometrics* 166, 167-183.

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

[3] Anderson, T.W. and C. Hsiao, 1981 Estimation of dynamic models with error components, *Journal of the American Statistical Association*, 589-606.

[4] Anderson, T.W., and C. Hsiao, 1982, Formulation and estimation of dynamic models using panel data, *Journal of Econometrics* 18, 47-82.

[5] Anderson, T.W., and H. Rubin, 1949, Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations, The Annals of Mathematical Statistics 20, 46-63.

[6] Angrist, J.D., G.W. Imbens, and A.B. Krueger, 1999. Jackknife instrumental variables estimation, *Journal of Applied Econometrics* 14, 57-67.

[7] Arellano, M., and S. Bond, 1991. Some tests of specification for panel data: monte carlo evidence and an application to employment equations, *Review of Economic Studies* 58, 277-297.

[8] Arellano, M., and O. Bover, 1995, Another look at the instrumental variable estimation of error-components models, *Journal of Econometrics* 68, 29-51.

[9] Chao, J. N. Swanson, J. Hausman, W. Newey, and T. Woutersen (2012), Asymptotic distribution of JIVE in a heteroskedastic IV regression with many instruments, *Econometric Theory* 28, 42-86.

[10] Hahn, J., and W. Newey, 2004, Jackknife and Analytical Bias Reduction for Nonlinear Panel Models, *Econometrica*, 72, 1295-1319.

[11] Hahn, J. and G. M. Kuersteiner (2002): Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large, *Econometrica*, 70, 1639-1657.

[12] Hansen, C., and D. Kozbur, 2014, Instrumental variables estimation with many weak instruments using regularized JIVE, Journal of Econometrics 182, 290-308.

[13] Hood, W. C., and T. C. Koopmans, 1953, *Studies in Econometric Method*, Cowles Foundation Monograph 14, Yale University Press, New Heaven, USA.

[14] Hsiao, C., 1982, Identification, in *Handbook of Econometrics* vol 1, edited by Z. Griliches and M. Intriligator, pp. 223-283, Amsterdam, Netherlands.

[15] Hsiao, C., 2014, *Analysis of Panel Data*, 3rd edition, Cambridge University Press, Cambridge, UK.

[16] Hsiao, C., and J. Zhang, 2015, IV, GMM or MLE to estimate dynamic panel data models. *Journal of Econometrics* 187, 312-322.

[17] Hsiao, C., and Q. Zhou, 2015, Statistical inference for panel dynamic simultaneous equations models, *Journal of Econometrics* 189, 383–396.

[18] Hsiao, C., and Q. Zhou, 2017, First difference or forward demeaning: Implications for the method of moments estimators, Econometric Reviews, forthcoming.

[19] Intriligator, D., R. Bodkin and C. Hsiao, 1996, *Econometric Models, Techniques, and Applications*, 2nd edition, Pearson.

[20] Lee, L-F, and J. Yu, 2014, Efficient GMM estimation of spatial dynamic panel data models with fixed effects, *Journal of Econometrics* 180, 174–197.

[21] Lee, N., R. Moon, and Q. Zhou, 2015, Many IVs estimation of dynamic panel regression models with measurement error, Journal of Econometrics, forthcoming.

[22] Moon, H.R., B. Perron, and P.C.B. Phillips, 2015, Incidental parameters and dynamic panel modeling, in *Oxford Handbook of Panel Data,* Edited by Badi H. Baltagi.

[23] Nerlove, M., 2000, An essay on the history of panel data econometrics, working paper;

[24] Phillips, G.D.A. and C. Hale, 1977, The bias of instrumental variable estimators of simultaneous equation systems, *International Economic Review* 18, 219-228.

[25] Phillips, P.C.B., and H.R. Moon, 1999, Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica* 67, 1057-1112.

Appendix: Mathematical Proofs

This appendix includes the mathematical proofs that are omitted in the main paper. In what follows, we shall let $\|A\| = \sqrt{tr\,(AA')}$ denote the Frobenius norm, $\|A\|_0 = \lambda_{\max}(A)$, where $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ denote the maximum and minimum eigenvalues of A, respectively. We also let C denote a generic finite constant, whose value may vary case by case.

A.1 Useful Lemmas

Before we introduce the lemmas, we notice that for the reduced form (2), we have

$$\mathbf{y}_{it} = \Pi \mathbf{y}_{it-1} + \boldsymbol{\xi}_i + \mathbf{v}_{it},$$

then

$$\mathbf{y}_{it} = (I_2 - \Pi)^{-1} \boldsymbol{\xi}_i + (I_2 - \Pi L)^{-1} \mathbf{v}_{it},$$

where L denotes the lag operator. Consequently, we can decompose $\mathbf{y}_{it}$ as

$$\mathbf{y}_{it} = (I_2 - \Pi)^{-1} \boldsymbol{\xi}_i + \mathbf{w}_{it}, \tag{A.1}$$

where

$$\mathbf{w}_{it} = (w_{1,it}, w_{2,it})' = \Pi \mathbf{w}_{it-1} + \mathbf{v}_{it} \tag{A.2}$$

and $\mathbf{w}_{it}$ is a stationary VAR(1) process under Assumption 2.

As a result, the forward demeaning transformation for (A.1) is given by

$$\mathbf{y}_{it}^f = \mathbf{w}_{it}^f = \Pi \mathbf{w}_{it-1}^f + \mathbf{v}_{it}^f, \tag{A.3}$$

and

$$\mathbf{w}_{it}^f = c_t \left(\mathbf{w}_{it} - \frac{1}{T-t} \sum_{s=t+1}^{T} \mathbf{w}_{is} \right) = c_t \mathbf{w}_{it} - c_t \bar{\mathbf{w}}_{i,t+1T},$$

where

$$\bar{\mathbf{w}}_{i,t+1T} = \frac{1}{T-t} \sum_{s=t+1}^{T} \mathbf{w}_{is}.$$

Similarly, the first difference transformation for (A.1) is given by

$$\Delta \mathbf{y}_{it} = \Delta \mathbf{w}_{it} = \Pi \Delta \mathbf{w}_{it-1} + \Delta \mathbf{v}_{it}. \tag{A.4}$$

Now let's turn to the lemmas. Lemma (A.1) to lemma (A.4) are used to derive the results of lemma (A.5) to lemma (A.8), and the latter are used to establish the theorems in the paper.

Lemma A.1 *Let $\mathbf{d}_t$ and $\mathbf{d}_s$ be $N \times 1$ vectors containing the diagonal elements of $\mathbf{P}_t$ and $\mathbf{P}_s$, respectively, so that $tr(\mathbf{P}_t) = \mathbf{d}_t' 1_N = 2t$ and $tr(\mathbf{P}_s) = \mathbf{d}_s' 1_N = 2s$, and $\mathbf{d}_t'\mathbf{d}_s \leq 2\min(t,s)$, then under assumptions 1-4, for $l \geq r > t$, $p \geq q > s$ and $t \geq s$*

$$Cov\left(\mathbf{u}_{(a),l}' \mathbf{P}_t \mathbf{u}_{(b),r}, \mathbf{u}_{(a),p}' \mathbf{P}_s \mathbf{u}_{(b),q}\right) = \begin{cases} \left(m^{(3)} + m^{(2)}\right) 2s + m^{(0)} E\left(\mathbf{d}_t'\mathbf{d}_s\right) & \text{if } l = r = p = q \\ E\left(u_{(a),it}^2 u_{(b),it}\right) E\left(\mathbf{d}_t' \mathbf{P}_s \mathbf{u}_{(b),q}\right) & \text{if } l = r = p \neq q < t \\ m^{(3)} 2s & \text{if } l = p \neq r = q \\ 0 & \text{otherwise} \end{cases}$$

and $\left|E\left(\mathbf{d}_t'\mathbf{P}_s\mathbf{u}_{(b),q}\right)\right| \leq 2\left[stE\left(u_{(b),it}^2\right)\right]^{1/2}$, *and*

$$\begin{aligned} m^{(1)} &= E\left[u_{(a),it}^2 u_{(b),it}^2\right], m^{(2)} = \left(E\left[u_{(a),it}u_{(b),it}\right]\right)^2, \\ m^{(3)} &= E\left[u_{(a),it}^2\right]E\left[u_{(b),it}^2\right], m^{(0)} = m^{(1)} - 2m^{(2)} - m^{(3)}, \end{aligned}$$

and $\left(\mathbf{u}_{(a),t}, \mathbf{u}_{(b),t}\right)$ *takes any pair of* $N \times 1$ *vectors from random variables* $u_{g,it}$ $(g = 1, 2)$.

Proof can be found in Akashi and Kunitomo (2012).

Lemma A.2 *Under Assumptions 1-2, as well as the condition (10), then the following hold for all* $t = 1, \ldots, T-1$.

(a) $\left\|\tilde{B}_{Nt} - B_{Nt}\right\|^2 = O_p\left(\frac{t^2}{N}\right)$;

(b) $\lambda_{\min}\left(\tilde{B}_{Nt}\right) \geq C > 0$;

(c) $\lambda_{\min}\left(B_{Nt}\right) \geq C > 0$;

where $B_{Nt} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{z}_{it}\mathbf{z}_{it}'$, $\tilde{B}_{Nt} = \frac{1}{N}\sum_{i=1}^{N}E\left(\mathbf{z}_{it}\mathbf{z}_{it}'\right)$ *with* $\mathbf{z}_{it} = \left(y_{1,i0}, y_{2,i0}, \ldots, y_{1,it}, y_{2,it}\right)'$.

Proof can be found in Lemma A.4 of Lee et al (2015).

Lemma A.3 *Under Assumptions 1-4, as* $(N,T) \to \infty$, *we have*

$$\begin{aligned} \frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{W}_{t-1}'\mathbf{P}_{t-1}\mathbf{W}_{t-1} &= \frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-1}'\mathbf{W}_{t-1} + o_p(1) \\ &\to {}_p\Gamma_0, \end{aligned}$$

where $\mathbf{W}_{t-1} = \left(\mathbf{w}_{1t}, \ldots, \mathbf{w}_{Nt}\right)'$ *and* $\Gamma_0 = E\left(\mathbf{w}_{it}\mathbf{w}_{it}'\right) = \sum_{i=0}^{\infty}\Pi^i\Omega_v\Pi^{i\prime} = \sum_{s=0}^{\infty}\Pi^s\mathbf{B}^{-1}\Omega_u\mathbf{B}'^{-1}\Pi^{s\prime}$.

Proof can be found in Akashi and Kunitomo (2012).

Lemma A.4 *Under Assumptions 1-4 as well as the restriction (10), as* $(N,T) \to \infty$, *we have*

(a):

$$\frac{1}{NT}\sum_{t=2}^{T}\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t} = o_p(1),$$

the above results still hold if we replace $\mathbf{u}_{1,t}$ *by* $\mathbf{u}_{2,t}$, $\mathbf{v}_{1,t}$ *or* $\mathbf{v}_{2,t}$.

(b):

$$\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t} \to_p \sigma_{u,1}^2\sqrt{\kappa},$$

similar results can be derived if we replace $\mathbf{u}_{1,t}$ *by* $\mathbf{u}_{2,t}$, $\mathbf{v}_{1,t}$ *or* $\mathbf{v}_{2,t}$

Proof. (a) In order to show this, we first notice that

$$\begin{aligned}\frac{1}{NT}\sum_{t=2}^{T}E\left(\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\right) &= \frac{1}{NT}\sum_{t=2}^{T}tr\left(E\left(\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\right)\right)=\frac{1}{NT}\sum_{t=2}^{T}E\left(tr\left(\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\right)\right)\\ &= \frac{1}{NT}\sum_{t=2}^{T}E\left(tr\left(\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,t}'\right)\right)=\frac{1}{NT}\sum_{t=2}^{T}tr\left(E\left(\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,t}'\right)\right)\\ &= \frac{1}{NT}\sum_{t=2}^{T}tr\left(E\left(\mathbf{P}_{t-2}E_{t-1}\left(\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,t}'\right)\right)\right)=\frac{2\sigma_{u,1}^2}{NT}\sum_{t=2}^{T}tr\left(E\left(\mathbf{P}_{t-2}\right)\right)\\ &= \frac{2\sigma_{u,1}^2}{NT}\sum_{t=2}^{T}tr\left(E\left(\mathbf{P}_{t-2}\right)\right)=\frac{2\sigma_{u,1}^2}{NT}\sum_{t=2}^{T}E\left(tr\left(\mathbf{P}_{t-2}\right)\right)\\ &= \frac{2\sigma_{u,1}^2}{NT}\sum_{t=2}^{T}2\left(t-1\right)=O\left(\frac{T}{N}\right)\\ &= o\left(1\right),\end{aligned}$$

under restriction (10). Also, we have

$$\begin{aligned}&Var\left(\frac{1}{NT}\sum_{t=2}^{T}\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\right)\\ =\;& \frac{1}{N^2T^2}\sum_{s,t=2}^{T}E\left(\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,s}'\mathbf{P}_{s-2}\Delta\mathbf{u}_{1,s}\right)\\ =\;& \frac{1}{N^2T^2}\sum_{t=2}^{T}E\left(\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\right)+\frac{2}{N^2T^2}\sum_{s<t}E\left(\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,s}'\mathbf{P}_{s-2}\Delta\mathbf{u}_{1,s}\right),\end{aligned} \tag{A.5}$$

where the first term can be shown that

$$\begin{aligned}\frac{1}{N^2T^2}\sum_{t=2}^{T}E\left(\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\right) &= \frac{C}{N^2T^2}\sum_{t=2}^{T}t^2+o\left(1\right)\\ &= o\left(1\right),\end{aligned}$$

by using the results from lemma (A.1). Similarly, for the second term of (A.5), we have

$$\begin{aligned}\frac{2}{N^2T^2}\sum_{s<t}E\left(\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,s}'\mathbf{P}_{s-2}\Delta\mathbf{u}_{1,s}\right) &= \frac{C}{N^2T^2}\sum_{t=2}^{T}t\\ &= o\left(1\right).\end{aligned}$$

Consequently, we have

$$Var\left(\frac{1}{NT}\sum_{t=2}^{T}\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\right)=o\left(1\right),$$

as $(N,T)\to\infty$, which in turn gives

$$\frac{1}{NT}\sum_{t=2}^{T}\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}=o_p\left(1\right),$$

as required.

(b) For this result, it is obvious that

$$\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}E\left(\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)=\sigma_{u,1}^{2}\sqrt{\kappa}+o\left(1\right),$$

by following the above derivation. For its variance, we notice that

$$\begin{aligned}
&Var\left(\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)\\
=\;&E\left(\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)^{2}-\left[\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}E\left(\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)\right]^{2}\\
=\;&\frac{1}{NT}\sum_{t=1}^{T}E\left(\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)+\frac{2}{NT}\sum_{s<t}E\left(\mathbf{u}_{1,s}'\mathbf{P}_{s-1}\mathbf{u}_{1,s}\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)-\sigma_{u,1}^{4}\kappa+o\left(1\right)\\
=\;&o\left(1\right),
\end{aligned}$$

since

$$\begin{aligned}
&\frac{1}{NT}\sum_{t=1}^{T}E\left(\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)\\
=\;&\frac{1}{NT}\sum_{t=1}^{T}Cov\left(\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t},\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)+\frac{1}{NT}\sum_{t=1}^{T}E\left(\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)E\left(\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)\\
=\;&\frac{C}{NT}\sum_{t=2}^{T}t+\frac{1}{NT}\sum_{t=2}^{T}4\sigma_{u,1}^{4}\left(t-1\right)^{2}\\
=\;&o\left(1\right),
\end{aligned}$$

by using the results of lemma (A.1). Similarly,

$$\begin{aligned}
\frac{2}{NT}\sum_{s<t}E\left(\mathbf{u}_{1,s}'\mathbf{P}_{s-1}\mathbf{u}_{1,s}\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)&=\frac{2}{NT}\sum_{s<t}E\left(\mathbf{u}_{1,s}'\mathbf{P}_{s-1}\mathbf{u}_{1,s}E_{s}\left(\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)\right)\\
&=\frac{8\sigma_{u,1}^{4}}{NT}\sum_{s<t}\left(t-1\right)\left(s-1\right)\\
&=\frac{4\sigma_{u,1}^{4}}{NT}\sum_{t=3}^{T}\left(t-1\right)^{3}+o\left(1\right)\\
&=\sigma_{u,1}^{4}\kappa+o\left(1\right),
\end{aligned}$$

consequently, we have

$$Var\left(\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\right)\to 0,$$

as $(N,T)\to\infty$, which gives

$$\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\mathbf{u}_{1,t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}\to_{p}\sigma_{u,1}^{2}\sqrt{\kappa},$$

as required. ■

The following two lemmas provide the theoretical results needed for the GMM estimation based on FOD and FD.

Lemma A.5 *Under Assumptions 1-4, as well as the condition (10), for the FOD transformed model (6), as* $(N,T)\to\infty$*, we have*

(a)

$$\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{f\prime}\mathbf{P}_{t-1}\mathbf{Y}_{t-1\cdot t}^{(f)}\to_p \mathbf{D}'\Gamma_0\mathbf{D},$$

where $\Gamma_0=\sum_{i=0}^{\infty}\Pi^i\mathbf{B}^{-1}\Omega_u\mathbf{B}'^{-1}\Pi^{i\prime}$ *and*

$$\mathbf{D}=\begin{pmatrix}1 & \pi_{21}\\ 0 & \pi_{22}\end{pmatrix}=\begin{pmatrix}1 & \beta\gamma_{22}\\ 0 & \gamma_{22}\end{pmatrix}.$$

(b)

$$\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{(f)\prime}\mathbf{P}_{t-1}\mathbf{u}_{1,t}^{f}\to_d N\left(0,\sigma_{u,1}^2\mathbf{D}'\Gamma_0\mathbf{D}\right)-\begin{pmatrix}0\\ \sigma_{u,12}\end{pmatrix}\sqrt{\kappa}.$$

Proof. (a) We note that

$$\begin{aligned}\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{f\prime}\mathbf{P}_{t-1}\mathbf{Y}_{t-1\cdot t}^{(f)} &= \frac{1}{NT}\sum_{t=1}^{T-1}\pi_t^2\left(\mathbf{W}_{t-1\cdot t}-\bar{\mathbf{W}}_{t-1\cdot t}\right)'\mathbf{P}_{t-1}\left(\mathbf{W}_{t-1\cdot t}-\bar{\mathbf{W}}_{t-1\cdot t}\right)\\ &= \frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{W}_{t-1\cdot t}'\mathbf{P}_{t-1}\mathbf{W}_{t-1\cdot t}-\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{W}_{t-1\cdot t}'\mathbf{P}_{t-1}\bar{\mathbf{W}}_{t-1\cdot t}\\ &\quad -\frac{1}{NT}\sum_{t=1}^{T-1}\bar{\mathbf{W}}_{t-1\cdot t}'\mathbf{P}_{t-1}\mathbf{W}_{t-1\cdot t}+\frac{1}{NT}\sum_{t=1}^{T-1}\bar{\mathbf{W}}_{t-1\cdot t}'\mathbf{P}_{t-1}\bar{\mathbf{W}}_{t-1\cdot t}+o_p(1),\end{aligned}$$

where $\mathbf{W}_{t-1\cdot t}=(\mathbf{w}_{1,t-1},\mathbf{w}_{2,t})$ with $\mathbf{w}_{j,t}=(w_{j,1t},w_{j,2t},\ldots,w_{j,Nt})'$ for $j=1,2$, $\bar{\mathbf{W}}_{t-1\cdot t}=\left(\frac{1}{T-t+1}\sum_{s=t}^{T}\mathbf{w}_{1,t},\frac{1}{T-t}\sum_{s=t+1}^{T}\mathbf{w}_{2,t}\right)=(\bar{\mathbf{w}}_{1,t-1T},\bar{\mathbf{w}}_{2,tT})$ and $\mathbf{w}_{it}$ is defined in (A.2).

It is obvious that

$$\begin{aligned}\mathbf{W}_{t-1\cdot t} &= (\mathbf{w}_{1,t-1},\mathbf{w}_{2,t})=(\mathbf{w}_{1,t-1},\pi_{21}\mathbf{w}_{1,t-1}+\pi_{22}\mathbf{w}_{2,t-1}+\mathbf{v}_{2,t})\\ &= \mathbf{W}_{t-1}\mathbf{D}+(\mathbf{0},\mathbf{v}_{2,t}),\end{aligned}\tag{A.6}$$

where

$$\mathbf{D}=\begin{pmatrix}1 & \pi_{21}\\ 0 & \pi_{22}\end{pmatrix}=\begin{pmatrix}1 & \beta\gamma_{22}\\ 0 & \gamma_{22}\end{pmatrix},$$

and

$$\begin{aligned}\bar{\mathbf{W}}_{t-1\cdot t} &= (\bar{\mathbf{w}}_{1,t-1T},\bar{\mathbf{w}}_{2,tT})=(\bar{\mathbf{w}}_{1,t-1T},\pi_{21}\bar{\mathbf{w}}_{1,t-1T}+\pi_{22}\bar{\mathbf{w}}_{2,t-1T}+\bar{\mathbf{v}}_{2,tT})\\ &= \bar{\mathbf{W}}_{t-1T}\mathbf{D}+(\mathbf{0},\bar{\mathbf{v}}_{2,tT}),\end{aligned}\tag{A.7}$$

Then substituting (A.6) and (A.7) yields

$$\begin{aligned}
\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{W}'_{t-1\cdot t}\mathbf{P}_{t-1}\bar{\mathbf{W}}_{t-1\cdot t} &= \frac{1}{NT}\sum_{t=1}^{T-1}\left[\mathbf{D}'\mathbf{W}'_{t-1}+(\mathbf{0},\mathbf{v}_{2,t})'\right]\mathbf{P}_{t-1}\left(\bar{\mathbf{W}}_{t-1T}\mathbf{D}+(\mathbf{0},\bar{\mathbf{v}}_{2,tT})\right)\\
&= \frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{D}'\mathbf{W}'_{t-1}\mathbf{P}_{t-1}\bar{\mathbf{W}}_{t-1T}\mathbf{D}+\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{D}'\mathbf{W}'_{t-1}\mathbf{P}_{t-1}\left(\mathbf{0},\bar{\mathbf{v}}_{2,tT}\right)\\
&\quad+\frac{1}{NT}\sum_{t=1}^{T-1}\left(\mathbf{0},\mathbf{v}_{2,t}\right)'\mathbf{P}_{t-1}\left(\mathbf{0},\bar{\mathbf{v}}_{2,tT}\right)+\frac{1}{NT}\sum_{t=1}^{T-1}\left(\mathbf{0},\mathbf{v}_{2,t}\right)'\mathbf{P}_{t-1}\left(\mathbf{0},\bar{\mathbf{v}}_{2,tT}\right),
\end{aligned}$$

and each term can be shown to be $o_p(1)$ by using the results of part (a) of lemma (A.4), for instance, the first term is given by

$$\begin{aligned}
\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{D}'\mathbf{W}'_{t-1}\mathbf{P}_{t-1}\bar{\mathbf{W}}_{t-1T}\mathbf{D} &= \mathbf{D}'\left(\frac{1}{NT}\sum_{t=1}^{T-1}\frac{1}{T-t+1}\sum_{s=t}^{T}\mathbf{W}'_{t-1}\mathbf{P}_{t-1}\mathbf{W}_s\right)\mathbf{D}\\
&= \mathbf{D}'\left(\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{W}'_{t-1}\mathbf{P}_{t-1}\mathbf{W}_t\frac{1}{T-t+1}\sum_{s=t}^{T}\Pi'^{s-t}\right)\mathbf{D}+o_p(1)\\
&= O_p\left(\frac{\log T}{T}\right)=o_p(1),
\end{aligned}$$

as $T\to\infty$. Similarly, all other terms can be shown to be $o_p(1)$, which yields

$$\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{W}'_{t-1\cdot t}\mathbf{P}_{t-1}\bar{\mathbf{W}}_{t-1\cdot t}=o_p(1).$$

By using the same argument, we can show that

$$\begin{aligned}
\frac{1}{NT}\sum_{t=1}^{T-1}\bar{\mathbf{W}}'_{t-1\cdot t}\mathbf{P}_{t-1}\mathbf{W}_{t-1\cdot t} &= o_p(1),\\
\frac{1}{NT}\sum_{t=1}^{T-1}\bar{\mathbf{W}}'_{t-1\cdot t}\mathbf{P}_{t-1}\bar{\mathbf{W}}_{t-1\cdot t} &= o_p(1).
\end{aligned}$$

Combining the above yields

$$
\begin{aligned}
\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{f\prime}\mathbf{P}_{t-1}\mathbf{Y}_{t-1\cdot t}^{(f)} &= \frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{W}_{t-1\cdot t}^{\prime}\mathbf{P}_{t-1}\mathbf{W}_{t-1\cdot t}+o_p(1) \\
&= \frac{1}{NT}\sum_{t=1}^{T-1}\left[\mathbf{D}^{\prime}\mathbf{W}_{t-1}^{\prime}+(\mathbf{0},\mathbf{v}_{2,t})^{\prime}\right]\mathbf{P}_{t-1}\left[\mathbf{W}_{t-1}\mathbf{D}+(\mathbf{0},\mathbf{v}_{2,t})\right]+o_p(1) \\
&= \frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{D}^{\prime}\mathbf{W}_{t-1}^{\prime}\mathbf{P}_{t-1}\mathbf{W}_{t-1}\mathbf{D}+\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{D}^{\prime}\mathbf{W}_{t-1}^{\prime}\mathbf{P}_{t-1}(\mathbf{0},\mathbf{v}_{2,t}) \\
&\quad +\frac{1}{NT}\sum_{t=1}^{T-1}(\mathbf{0},\mathbf{v}_{2,t})^{\prime}\mathbf{P}_{t-1}\mathbf{W}_{t-1}\mathbf{D}+\frac{1}{NT}\sum_{t=1}^{T-1}(\mathbf{0},\mathbf{v}_{2,t})^{\prime}\mathbf{P}_{t-1}(\mathbf{0},\mathbf{v}_{2,t})+o_p(1) \\
&= \frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{D}^{\prime}\mathbf{W}_{t-1}^{\prime}\mathbf{P}_{t-1}\mathbf{W}_{t-1}\mathbf{D}+o_p(1) \\
&\to {}_p\mathbf{D}^{\prime}\Gamma_0\mathbf{D},
\end{aligned}
$$

by using the result of lemma (A.3) and the fact that

$$
\begin{aligned}
\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{D}^{\prime}\mathbf{W}_{t-1}^{\prime}\mathbf{P}_{t-1}(\mathbf{0},\mathbf{v}_{2,t}) &= o_p(1),\ \frac{1}{NT}\sum_{t=1}^{T-1}(\mathbf{0},\mathbf{v}_{2,t})^{\prime}\mathbf{P}_{t-1}\mathbf{W}_{t-1}\mathbf{D}=o_p(1), \\
\frac{1}{NT}\sum_{t=1}^{T-1}(\mathbf{0},\mathbf{v}_{2,t})^{\prime}\mathbf{P}_{t-1}(\mathbf{0},\mathbf{v}_{2,t}) &= o_p(1),
\end{aligned}
$$

which holds since, for example, $E\left[\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{D}^{\prime}\mathbf{W}_{t-1}^{\prime}\mathbf{P}_{t-1}(\mathbf{0},\mathbf{v}_{2,t})\right]=0$, and

$$
\begin{aligned}
& E\left[\left(\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{D}^{\prime}\mathbf{W}_{t-1}^{\prime}\mathbf{P}_{t-1}(\mathbf{0},\mathbf{v}_{2,t})\right)\left(\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{D}^{\prime}\mathbf{W}_{t-1}^{\prime}\mathbf{P}_{t-1}(\mathbf{0},\mathbf{v}_{2,t})\right)^{\prime}\right] \\
=\ & \frac{1}{N^2T^2}\sum_{s,t=1}^{T-1}\mathbf{D}^{\prime}E\left(\mathbf{W}_{t-1}^{\prime}\mathbf{P}_{t-1}(\mathbf{0},\mathbf{v}_{2,t})(\mathbf{0},\mathbf{v}_{2,s})^{\prime}\mathbf{P}_{s-1}\mathbf{W}_{s-1}\right)\mathbf{D} \\
=\ & \frac{1}{N^2T^2}\sum_{t=1}^{T-1}\mathbf{D}^{\prime}E\left(\mathbf{W}_{t-1}^{\prime}\mathbf{P}_{t-1}(\mathbf{0},\mathbf{v}_{2,t})(\mathbf{0},\mathbf{v}_{2,t})^{\prime}\mathbf{P}_{t-1}\mathbf{W}_{t-1}\right)\mathbf{D} \\
\leq\ & \frac{C}{N^2T^2}\sum_{t=1}^{T-1}\mathbf{D}^{\prime}E\left(\mathbf{W}_{t-1}^{\prime}\mathbf{P}_{t-1}\mathbf{W}_{t-1}\right)\mathbf{D}=O_p\left(\frac{1}{NT}\right).
\end{aligned}
$$

A similar argument can be applied to all other remaining terms.

To summarize, as $(N,T)\to\infty$, we obtain

$$
\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{f\prime}\mathbf{P}_{t-1}\mathbf{Y}_{t-1\cdot t}^{(f)}\to_p\mathbf{D}^{\prime}\Gamma_0\mathbf{D}, \tag{A.8}
$$

as required.

(b) By using the result (A.6) and (A.7), we obtain

$$
\begin{aligned}
& \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \mathbf{Y}_{t-1\cdot t}^{(f)\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t}^{f} \\
= & \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \mathbf{W}_{t-1\cdot t}^{(f)\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t}^{f} \\
= & \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \left[\mathbf{W}_{t-1} \mathbf{D} + (\mathbf{0}, \mathbf{v}_{2,t}) - \bar{\mathbf{W}}_{t-1T} \mathbf{D} - (\mathbf{0}, \bar{\mathbf{v}}_{2,tT}) \right]^{\prime} \mathbf{P}_{t-1} \left(\mathbf{u}_{1,t} - \bar{\mathbf{u}}_{1,tT} \right) + o_p(1) \\
= & \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \mathbf{D}^{\prime} \mathbf{W}_{t-1}^{\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t} - \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} (\mathbf{0}, \mathbf{v}_{2,t})^{\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t} + o_p(1), \qquad \text{(A.9)}
\end{aligned}
$$

where the last identity holds since all remaining terms can be shown to be $o_p(1)$. For instance, for the term $\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \mathbf{D}^{\prime} \bar{\mathbf{W}}_{t-1T}^{\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t}$, we have

$$
\begin{aligned}
\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \mathbf{D}^{\prime} \bar{\mathbf{W}}_{t-1T}^{\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t} &= \mathbf{D}^{\prime} \left[\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \frac{1}{T-t+1} \sum_{s=t}^{T} \mathbf{W}_{s}^{\prime} \mathbf{P}_{t-1} \mathbf{V}_{t} \right] \mathbf{B} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&= \mathbf{D}^{\prime} \left[\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \frac{1}{T-t+1} \sum_{s=t}^{T} \Pi^{\prime s-t} \mathbf{V}_{t}^{\prime} \mathbf{P}_{t-1} \mathbf{V}_{t} \right] \mathbf{B} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + o_p(1) \\
&= \mathbf{D}^{\prime} \left[\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \frac{1}{T-t+1} \sum_{s=t}^{T} \Pi^{\prime s-t} \mathbf{V}_{t}^{\prime} \mathbf{P}_{t-1} \mathbf{V}_{t} \right] \mathbf{B} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + o_p(1) \\
&= O_p \left(\sqrt{\frac{T (\log T)^2}{N}} \right),
\end{aligned}
$$

which will be $o_p(1)$ under alternative restriction (10). Similarly, we can show all other remaining terms are $o_p(1)$.

For (A.9), it is obvious that the first term will contribute to the limiting distribution as $(N,T) \to \infty$

$$
\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \mathbf{D}^{\prime} \mathbf{W}_{t-1}^{\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t} \to_d N\left(0, \sigma_{u,1}^2 \mathbf{D}^{\prime} \Gamma_0 \mathbf{D}\right),
$$

and the second term will contribute to the asymptotic bias under (10) with

$$
\begin{aligned}
\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} (\mathbf{0}, \mathbf{v}_{2,t})^{\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t} &= \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \begin{pmatrix} 0 \\ \mathbf{v}_{2,t}^{\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t} \end{pmatrix} = \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \begin{pmatrix} 0 \\ \mathbf{u}_{2,t}^{\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t} \end{pmatrix} \\
&\to_p \begin{pmatrix} 0 \\ \sigma_{u,12} \end{pmatrix} \sqrt{\kappa},
\end{aligned}
$$

by using the results from part (b) of lemma (A.4). Combining these results yields

$$
\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \mathbf{Y}_{t-1\cdot t}^{(f)\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t}^{f} \to_d N\left(0, \sigma_{u,1}^2 \mathbf{D}^{\prime} \Gamma_0 \mathbf{D}\right) - \begin{pmatrix} 0 \\ \sigma_{u,12} \end{pmatrix} \sqrt{\kappa},
$$

as $(N,T) \to \infty$ under alternative restriction (10). ■

Lemma A.6 *Under Assumptions 1-4, as well as the condition (10), for the FD transformed model (8), as* $(N,T)\to\infty$, *we have*

(a)

$$\frac{1}{NT}\sum_{t=2}^{T}\Delta\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-2}\Delta\mathbf{Y}_{t-1\cdot t}\to_{p}\mathbf{D}'\left(I_2-\Pi\right)\Gamma_0\left(I_2-\Pi'\right)\mathbf{D}$$

where $\mathbf{D}$,Π *and* Γ_0 *are defined in lemma (A.5).*

(b)

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\Delta\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\to N\left(\tilde{\mathbf{b}}_0,\sigma_{u,1}^2\mathbf{D}'\left(I_2-\Pi\right)\left[\left(I_2-\Pi\right)\Gamma_0+\Gamma_0\left(I_2-\Pi'\right)\right]\left(I_2-\Pi'\right)\mathbf{D}\right),$$

where

$$\tilde{\mathbf{b}}_0=-\left[\mathbf{D}'\begin{pmatrix}\sigma_{u,1}^2+\beta\sigma_{u,12}\\ \sigma_{u,21}\end{pmatrix}-\begin{pmatrix}0\\ 2\sigma_{u,21}\end{pmatrix}\right]\sqrt{\kappa}.$$

Proof. (a) To show this, we first notice that

$$\begin{aligned}\Delta\mathbf{Y}_{t-1\cdot t} &= \Delta\mathbf{W}_{t-1\cdot t}=\left(\Delta\mathbf{w}_{1,t-1},\Delta\mathbf{w}_{2,t}\right)=\left(\Delta\mathbf{w}_{1,t-1},\pi_{21}\Delta\mathbf{w}_{1,t-1}+\pi_{22}\Delta\mathbf{w}_{2,t-1}+\Delta\mathbf{v}_{2,t}\right)\\ &= \left(\Delta\mathbf{w}_{1,t-1},\Delta\mathbf{w}_{2,t-1}\right)\mathbf{D}+\left(0,\Delta\mathbf{v}_{2,t}\right)\\ &= \Delta\mathbf{W}_{t-1}\mathbf{D}+\left(0,\Delta\mathbf{v}_{2,t}\right),\end{aligned}$$

where $\mathbf{D}$ is defined in lemma (A.5). Then

$$\begin{aligned}&\frac{1}{NT}\sum_{t=2}^{T}\Delta\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-2}\Delta\mathbf{Y}_{t-1\cdot t}\\ =\ &\frac{1}{NT}\sum_{t=2}^{T}\mathbf{D}'\left(\mathbf{W}_{t-1}-\mathbf{W}_{t-2}\right)'\mathbf{P}_{t-2}\left(\mathbf{W}_{t-1}-\mathbf{W}_{t-2}\right)\mathbf{D}\\ &+\frac{1}{NT}\sum_{t=2}^{T}\mathbf{D}'\left(\mathbf{W}_{t-1}-\mathbf{W}_{t-2}\right)'\mathbf{P}_{t-2}\left(0,\Delta\mathbf{v}_{2,t}\right)+\frac{1}{NT}\sum_{t=2}^{T}\left(0,\Delta\mathbf{v}_{2,t}\right)'\mathbf{P}_{t-2}\left(\mathbf{W}_{t-1}-\mathbf{W}_{t-2}\right)\mathbf{D}\\ &+\frac{1}{NT}\sum_{t=2}^{T}\left(0,\Delta\mathbf{v}_{2,t}\right)'\mathbf{P}_{t-2}\left(0,\Delta\mathbf{v}_{2,t}\right)\\ =\ &\frac{1}{NT}\sum_{t=2}^{T}\mathbf{D}'\left(\mathbf{W}_{t-1}-\mathbf{W}_{t-2}\right)'\mathbf{P}_{t-2}\left(\mathbf{W}_{t-1}-\mathbf{W}_{t-2}\right)\mathbf{D}+o_p\left(1\right),\end{aligned}\tag{A.10}$$

with

$$\frac{1}{NT}\sum_{t=2}^{T}\left(0,\Delta\mathbf{v}_{2,t}\right)'\mathbf{P}_{t-2}\left(0,\Delta\mathbf{v}_{2,t}\right)=\begin{pmatrix}0\\ \frac{1}{NT}\sum_{t=2}^{T}\Delta\mathbf{v}_{2,t}'\mathbf{P}_{t-2}\Delta\mathbf{v}_{2,t}\end{pmatrix},$$

where it can be shown that

$$\frac{1}{NT}\sum_{t=2}^{T}\left(0,\Delta\mathbf{v}_{2,t}\right)'\mathbf{P}_{t-2}\left(0,\Delta\mathbf{v}_{2,t}\right)=o_p\left(1\right),$$

by using the results from lemma (A.4), and

$$
\begin{aligned}
& \frac{1}{NT}\sum_{t=2}^{T}\left(\mathbf{W}_{t-1}-\mathbf{W}_{t-2}\right)'\mathbf{P}_{t-2}\left(\mathbf{W}_{t-1}-\mathbf{W}_{t-2}\right) \\
= \;& \frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-1}'\mathbf{P}_{t-2}\mathbf{W}_{t-1}-\frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-1}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}-\frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-1} \\
& +\frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2} \\
= \;& \frac{1}{NT}\sum_{t=2}^{T}\left(\Pi\mathbf{W}_{t-2}'+\mathbf{V}_{t-1}'\right)\mathbf{P}_{t-2}\left(\mathbf{W}_{t-2}\Pi'+\mathbf{V}_{t-1}\right)-\frac{1}{NT}\sum_{t=2}^{T}\left(\Pi\mathbf{W}_{t-2}'+\mathbf{V}_{t-1}'\right)\mathbf{P}_{t-2}\mathbf{W}_{t-2} \\
& -\frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\left(\mathbf{W}_{t-2}\Pi'+\mathbf{V}_{t-1}\right)+\frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2} \\
= \;& \frac{1}{NT}\sum_{t=2}^{T}\Pi\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}\Pi'+\frac{1}{NT}\sum_{t=2}^{T}\Pi\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{V}_{t-1}+\frac{1}{NT}\sum_{t=2}^{T}\mathbf{V}_{t-1}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}\Pi' \\
& +\frac{1}{NT}\sum_{t=2}^{T}\mathbf{V}_{t-1}'\mathbf{P}_{t-2}\mathbf{V}_{t-1}-\frac{1}{NT}\sum_{t=2}^{T}\Pi\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}-\frac{1}{NT}\sum_{t=2}^{T}\mathbf{V}_{t-1}'\mathbf{P}_{t-2}\mathbf{W}_{t-2} \\
& -\frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}\Pi'-\frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{V}_{t-1}+\frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2} \\
= \;& \frac{1}{NT}\sum_{t=2}^{T}\Pi\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}\Pi'-\frac{1}{NT}\sum_{t=2}^{T}\Pi\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}-\frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}\Pi' \\
& +\frac{1}{NT}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}+o_p(1) \\
= \;& \Pi\Gamma_0\Pi'-\Pi\Gamma_0-\Gamma_0\Pi'+\Gamma_0+o_p(1) \\
= \;& \left(I_2-\Pi\right)\Gamma_0\left(I_2-\Pi'\right)+o_p(1),
\end{aligned}
$$

since from (A.2)

$$\mathbf{W}_{t-1}=\mathbf{W}_{t-2}\Pi'+\mathbf{V}_{t-1},$$

and by using the results from the lemma (A.4).

Substituting the above back to (A.10), as $(N,T)\to\infty$, we have

$$\frac{1}{NT}\sum_{t=2}^{T}\Delta\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-2}\Delta\mathbf{Y}_{t-1\cdot t}\to_p\mathbf{D}'\left(I_2-\Pi\right)\Gamma_0\left(I_2-\Pi'\right)\mathbf{D},$$

as required.

(b) To show this, we note that

$$\begin{aligned}
& \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\Delta\mathbf{Y}'_{t-1\cdot t}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t} \\
= & \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{D}'\Delta\mathbf{W}'_{t-1}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}+\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left(0,\Delta\mathbf{v}_{2,t}\right)'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t} \\
= & \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{D}'\mathbf{W}'_{t-1}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}-\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{D}'\mathbf{W}'_{t-2}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}+\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left(0,\Delta\mathbf{v}_{2,t}\right)'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t} \\
= & \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{D}'\left(\Pi-I_2\right)\mathbf{W}'_{t-2}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}+\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{D}'\mathbf{V}'_{t-1}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t} \\
& +\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left(0,\Delta\mathbf{v}_{2,t}\right)'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}, \qquad \text{(A.11)}
\end{aligned}$$

where the first term will contribute to the limiting distribution and the last two terms will contribute to the bias. For the second term, we have

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{D}'\mathbf{V}'_{t-1}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}=\mathbf{D}'\begin{pmatrix}\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{v}'_{1,t-1}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\\ \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{v}'_{2,t-1}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\end{pmatrix},$$

where

$$\begin{aligned}
\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{v}'_{1,t-1}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t} &= -\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{v}'_{1,t-1}\mathbf{P}_{t-2}\mathbf{u}_{1,t-1}+o_p(1) \\
&= -\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left(\mathbf{u}'_{1,t-1}+\beta\mathbf{u}'_{2,t-1}\right)\mathbf{P}_{t-2}\mathbf{u}_{1,t-1}+o_p(1), \\
&\to_p -\left(\sigma_{u,1}^2+\beta\sigma_{u,12}\right)\sqrt{\kappa},
\end{aligned}$$

and

$$\begin{aligned}
\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{v}'_{2,t-1}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t} &= -\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{v}'_{2,t-1}\mathbf{P}_{t-2}\mathbf{u}_{1,t-1}+o_p(1) \\
&= -\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{u}'_{2,t-1}\mathbf{P}_{t-2}\mathbf{u}_{1,t-1}+o_p(1), \\
&\to_p -\sigma_{u,21}\sqrt{\kappa},
\end{aligned}$$

by using the results of part (b) of lemma (A.4) then

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{D}'\mathbf{V}'_{t-1}\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\to_p -\mathbf{D}'\begin{pmatrix}\sigma_{u,1}^2+\beta\sigma_{u,12}\\ \sigma_{u,21}\end{pmatrix}\sqrt{\kappa}.$$

Similarly, we can show that

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left(0,\Delta\mathbf{v}_{2,t}\right)'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}=\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left(0,\Delta\mathbf{u}_{2,t}\right)'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\to_p\begin{pmatrix}0\\ 2\sigma_{u,21}\end{pmatrix}\sqrt{\kappa}.$$

As a result, we have

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{D}\mathbf{V}_{t-1}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}+\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left(0,\Delta\mathbf{u}_{2,t}\right)'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\rightarrow_{p}\tilde{\mathbf{b}}_{0},$$

where $\tilde{\mathbf{b}}_0$ denotes the asymptotic bias term and is given by

$$\tilde{\mathbf{b}}_{0}=-\left[\mathbf{D}'\begin{pmatrix}\sigma_{u,1}^{2}+\beta\sigma_{u,12}\\ \sigma_{u,21}\end{pmatrix}-\begin{pmatrix}0\\ 2\sigma_{u,21}\end{pmatrix}\right]\sqrt{\kappa}. \tag{A.12}$$

which in turn yields

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\Delta\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}=\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{D}'\left(\Pi-I_{2}\right)\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}+\tilde{\mathbf{b}}_{0}+o_{p}\left(1\right), \tag{A.13}$$

and for the first term, it is obvious that

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}E\left[\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\right]=0,$$

and

$$\begin{aligned}
&Var\left(\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\right)=\frac{1}{NT}\sum_{s,t}E\left[\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,s}'\mathbf{P}_{s-2}\mathbf{W}_{s-2}\right]\\
=\;&\frac{1}{NT}\sum_{t=2}^{T}E\left[\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}\right]+\frac{1}{NT}\sum_{t=3}^{T}E\left[\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\Delta\mathbf{u}_{1,t-1}'\mathbf{P}_{t-3}\mathbf{W}_{t-3}\right]\\
&+\frac{1}{NT}\sum_{t=3}^{T}E\left[\mathbf{W}_{t-3}'\mathbf{P}_{t-3}\Delta\mathbf{u}_{1,t-1}\Delta\mathbf{u}_{1,t}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}\right]\\
=\;&2\sigma_{u,1}^{2}\frac{1}{NT}\sum_{t=2}^{T}E\left[\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}\right]-\sigma_{u,1}^{2}\frac{1}{NT}\sum_{t=3}^{T}E\left[\left(\Pi\mathbf{W}_{t-3}'+\mathbf{V}_{t-2}'\right)\mathbf{P}_{t-3}\mathbf{W}_{t-3}\right]\\
&-\sigma_{u,1}^{2}\frac{1}{NT}\sum_{t=3}^{T}E\left[\mathbf{W}_{t-3}'\mathbf{P}_{t-3}\left(\mathbf{W}_{t-3}\Pi'+\mathbf{V}_{t-2}\right)\right]\\
=\;&2\sigma_{u,1}^{2}\frac{1}{NT}\sum_{t=2}^{T}E\left[\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\mathbf{W}_{t-2}\right]-\sigma_{u,1}^{2}\frac{1}{NT}\sum_{t=3}^{T}E\left[\Pi\mathbf{W}_{t-3}'\mathbf{P}_{t-3}\mathbf{W}_{t-3}\right]\\
&-\sigma_{u,1}^{2}\frac{1}{NT}\sum_{t=3}^{T}E\left[\mathbf{W}_{t-3}'\mathbf{P}_{t-3}\mathbf{W}_{t-3}\Pi'\right]\\
=\;&2\sigma_{u,1}^{2}\Gamma_{0}-\sigma_{u,1}^{2}\Pi\Gamma_{0}-\sigma_{u,1}^{2}\Gamma_{0}\Pi'+o\left(1\right)\\
=\;&\sigma_{u,1}^{2}\left(\left(I_{2}-\Pi\right)\Gamma_{0}+\Gamma_{0}\left(I_{2}-\Pi'\right)\right).
\end{aligned}$$

As a result, we have

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\mathbf{W}_{t-2}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\rightarrow_{d}N\left(0,\sigma_{u,1}^{2}\left(\left(I_{2}-\Pi\right)\Gamma_{0}+\Gamma_{0}\left(I_{2}-\Pi'\right)\right)\right),$$

by following Akashi and Kunitomo (2012), which in turn gives

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\Delta\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\to N\left(\tilde{\mathbf{b}}_0,\sigma_{u,1}^2\mathbf{D}'\left(I_2-\Pi\right)\left[\left(I_2-\Pi\right)\Gamma_0+\Gamma_0\left(I_2-\Pi'\right)\right]\left(I_2-\Pi'\right)\mathbf{D}\right), \tag{A.14}$$

as required. ■

The last two lemmas provide the theoretical results needed for the JIVE estimation based on FOD and FD.

Lemma A.7 *Under Assumptions 1-4, as well as the condition (10), then for the JIVE based FOD, we have*

(a).

$$\frac{1}{NT}\sum_{t=1}^{T-1}\left[\sum_{i=1}^{N}\mathbf{y}_{i,t-1\cdot t}^{f\prime}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\mathbf{y}_{i,t-1\cdot t}^{f}\right]=o_p\left(1\right),$$

(b).

$$\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\mathbf{y}_{i,t-1\cdot t}^{f\prime}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}u_{1,it}^{f}\to_p\begin{pmatrix}0\\ \sigma_{u,12}\end{pmatrix}\sqrt{\kappa}.$$

Proof. (a) We notice that

$$\frac{1}{NT}\sum_{t=1}^{T-1}\left[\sum_{i=1}^{N}\mathbf{y}_{i,t-1\cdot t}^{f\prime}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\mathbf{y}_{i,t-1\cdot t}^{f}\right]=\frac{1}{NT}\sum_{t=1}^{T-1}$$

$$\begin{pmatrix}\sum_{i=1}^{N}y_{1,it-1}^{f}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}y_{1,it-1}^{f} & \sum_{i=1}^{N}y_{1,it-1}^{f}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}y_{2,it}^{f}\\ \sum_{i=1}^{N}y_{2,it}^{f}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}y_{1,it-1}^{f} & \sum_{i=1}^{N}y_{2,it}^{f}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}y_{2,it}^{f}\end{pmatrix}, \tag{A.15}$$

then we need to show that each element of (A.15) has zero limit. To this end, we first notice that for the (1,1)-th element of (A.15), we have

$$\begin{aligned}&\frac{1}{NT}\sum_{t=1}^{T-1}\sum_{i=1}^{N}y_{1,it-1}^{f}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}y_{1,it-1}^{f}\\ =&\frac{1}{NT}\sum_{t=1}^{T-1}\sum_{i=1}^{N}w_{1,it-1}^{f}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}w_{1,it-1}^{f}\\ =&\frac{1}{N^2T}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left(w_{1,it-1}-\bar{w}_{1,itT-1}\right)^2\mathbf{z}_{it-1}'\left(\left(B_{Nt-1}-\tilde{B}_{Nt-1}\right)+\tilde{B}_{Nt-1}\right)^{-1}\mathbf{z}_{it-1},\end{aligned}$$

where $B_{Nt-1}=\frac{1}{N}\sum_{j=1}^{N}\mathbf{z}_{jt-1}\mathbf{z}_{jt-1}'$ and $\tilde{B}_{Nt-1}=\frac{1}{N}\sum_{j=1}^{N}E\left(\mathbf{z}_{jt-1}\mathbf{z}_{jt-1}'\right)$. Similar strategy has also been used by Lee et al (2015).

Since $\left(\left(B_{Nt-2}-\tilde{B}_{Nt-2}\right)+\tilde{B}_{Nt-2}\right)^{-1}=\tilde{B}_{Nt-2}^{-1}+\frac{1}{\sqrt{N}}\tilde{B}_{Nt-2}^{-1}\left(\sqrt{N}\left(B_{Nt-2}-\tilde{B}_{Nt-2}\right)\right)\tilde{B}_{Nt-2}^{-1}+$

$O_p\left(\frac{t^2}{N}\right)$, we have

$$
\begin{aligned}
& \frac{1}{N^2T}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left(w_{1,it-1}-\bar{w}_{1,itT-1}\right)^2\mathbf{z}_{it-1}'\left(\left(B_{Nt-1}-\tilde{B}_{Nt-1}\right)+\tilde{B}_{Nt-1}\right)^{-1}\mathbf{z}_{it-1} \\
= & \frac{1}{N^2T}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left(w_{1,it-1}-\bar{w}_{1,itT-1}\right)^2\mathbf{z}_{it-1}'\tilde{B}_{Nt-2}^{-1}\mathbf{z}_{it-1} \\
& +\frac{1}{N^2T}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left(w_{1,it-1}-\bar{w}_{1,itT-1}\right)^2\mathbf{z}_{it-1}'\frac{1}{\sqrt{N}}\tilde{B}_{Nt-2}^{-1}\left(\sqrt{N}\left(B_{Nt-2}-\tilde{B}_{Nt-2}\right)\right)\tilde{B}_{Nt-2}^{-1}\mathbf{z}_{it-1} \\
& +O_p\left(\frac{t^2}{N}\right)\frac{1}{N^2T}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left(w_{1,it-1}-\bar{w}_{1,itT-1}\right)^2\mathbf{z}_{it-1}'\mathbf{z}_{it-1} \\
= & I_1+I_2+I_3,
\end{aligned}
$$

For term I_1,

$$
\begin{aligned}
|I_1| \leq & \frac{C}{N^2T}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left(w_{1,it-1}-\bar{w}_{1,itT-1}\right)^2\mathbf{z}_{it-1}'\mathbf{z}_{it-1} \leq \frac{C}{N^2T}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left\|\left(w_{1,it-1}-\bar{w}_{1,itT-1}\right)\mathbf{z}_{it-1}\right\|^2 \\
= & O_p\left(\frac{1}{NT}\sum_{t=1}^{T-1}t^2\right)=o_p(1), \qquad \text{(A.16)}
\end{aligned}
$$

and for term I_2,

$$
\begin{aligned}
|I_2| \leq & \frac{1}{N^2T}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left|\left(w_{1,it-1}-\bar{w}_{1,itT-1}\right)^2\mathbf{z}_{it-1}'\frac{1}{\sqrt{N}}\tilde{B}_{Nt-1}^{-1}\left(\sqrt{N}\left(B_{Nt-1}-\tilde{B}_{Nt-1}\right)\right)\tilde{B}_{Nt-1}^{-1}\mathbf{z}_{it-1}\right| \\
\leq & \frac{1}{N^{3/2}T}\sum_{t=1}^{T-1}\left\|\tilde{B}_{Nt-2}^{-1}\right\|_0^2\left\|\sqrt{N}\left(B_{Nt-2}-\tilde{B}_{Nt-2}\right)\right\|\frac{1}{N}\sum_{i=1}^{N}\left\|\left(w_{1,it-1}-\bar{w}_{1,itT-1}\right)\mathbf{z}_{it-1}\right\|^2 \\
\leq & \frac{C}{N^{3/2}}\sqrt{\frac{1}{T}\sum_{t=1}^{T-1}\left(\frac{1}{N}\sum_{i=1}^{N}\left\|\left(w_{1,it-1}-\bar{w}_{1,itT-1}\right)\mathbf{z}_{it-1}\right\|^2\right)^2}\sqrt{\frac{1}{T}\sum_{t=1}^{T-1}\left\|\sqrt{N}\left(B_{Nt-2}-\tilde{B}_{Nt-2}\right)\right\|^2} \\
= & \frac{C}{N^{3/2}}\sqrt{O_p\left(\frac{1}{T}\sum_{t=1}^{T-1}t^4\right)}\sqrt{O_p\left(\frac{1}{T}\sum_{t=1}^{T-1}t^2\right)} \\
= & O_p\left(\frac{T^3}{N^{3/2}}\right)=O_p\left(\frac{\kappa}{N^{1/2}}\right) \\
= & o_p(1). \qquad \text{(A.17)}
\end{aligned}
$$

by using the results of lemma (A.2). Similarly, we can show that

$$I_3=o_p(1).$$

Combing these results gives us

$$\frac{1}{NT}\sum_{t=1}^{T-1}\sum_{i=1}^{N}y_{1,it-1}^{f}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}y_{1,it-1}^{f}=o_p(1).$$

as required.

Similarly, we can show that

$$\frac{1}{NT}\sum_{i=1}^{N} y^{f}_{1,it-1}\mathbf{z}'_{it-1}\left(\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}y^{f}_{2,it} = o_p(1),$$

$$\frac{1}{NT}\sum_{i=1}^{N} y^{f}_{2,it}\mathbf{z}'_{it-1}\left(\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}y^{f}_{2,it} = o_p(1),$$

By combining the above results, we obtain

$$\frac{1}{NT}\sum_{t=1}^{T-1}\left[\sum_{i=1}^{N}\mathbf{y}^{f\prime}_{i,t-1\cdot t}\mathbf{z}'_{it-1}\left(\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\mathbf{y}^{f\prime}_{i,t-1\cdot t}\right] = o_p(1),$$

as required.

(b). We first note that $\mathbf{y}^{f}_{i,t-1\cdot t} = \left(y^{f}_{1,it-1}, y^{f}_{2,it}\right)'$, then

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\left[\sum_{i=1}^{N}\mathbf{y}^{f\prime}_{i,t-1\cdot t}\mathbf{z}'_{it-1}\left(\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}u^{f}_{1,it}\right] \\
= &\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\begin{pmatrix}\sum_{i=1}^{N} y^{f}_{1,it-1}\mathbf{z}'_{it-1}\left(\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}u^{f}_{1,it} \\ \sum_{i=1}^{N} y^{f}_{2,it}\mathbf{z}'_{it-1}\left(\mathbf{Z}'_{t-1}\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}u^{f}_{1,it}\end{pmatrix},
\end{aligned} \tag{A.18}$$

For the first element of (A.18), we notice that, from the reduced form (2) and (A.3),

$$\mathbf{y}^{f}_{it} = \mathbf{w}^{f}_{it} = \Pi\mathbf{w}^{f}_{it-1} + \mathbf{v}^{f}_{it}, \quad \mathbf{w}_{it} = \Pi\mathbf{w}_{it-1} + \mathbf{v}_{it}, \tag{A.19}$$

then we have

$$\begin{aligned}
y^{f}_{1,it-1} &= w^{f}_{1,it-1} = \pi_{11}w^{f}_{1,it-2} + \pi_{12}w^{f}_{2,it-2} + v^{f}_{1,it-1} \\
&= \pi_{11}c_t\left(w_{1,it-1} - \bar{w}_{1,itT}\right) + \pi_{12}c_t\left(w_{2,it-1} - \bar{w}_{2,itT}\right) + c_t\left(v_{1,it-1} - \bar{v}_{1,itT}\right),
\end{aligned} \tag{A.20}$$

where $\bar{w}_{1,itT} = \frac{1}{T-t}\sum_{s=t}^{T-1} w_{1,is}$ and $\bar{w}_{2,itT} = \frac{1}{T-t}\sum_{s=t}^{T-1} w_{2,is}$, and since

$$\Pi = \begin{pmatrix}\pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22}\end{pmatrix} = \begin{pmatrix}\gamma + \beta\gamma_{21} & \beta\gamma_{22} \\ \gamma_{21} & \gamma_{22}\end{pmatrix},$$

and

$$v_{1,it} = u_{1,it} + \beta u_{2,it} = (1,\beta)\,\mathbf{u}_{it}, \tag{A.21}$$

because

$$\mathbf{v}_{it} = \mathbf{B}^{-1}\begin{pmatrix}u_{1,it} \\ u_{2,it}\end{pmatrix} = \begin{pmatrix}1 & \beta \\ 0 & 1\end{pmatrix}\begin{pmatrix}u_{1,it} \\ u_{2,it}\end{pmatrix} = \begin{pmatrix}u_{1,it} + \beta u_{2,it} \\ u_{2,it}\end{pmatrix}. \tag{A.22}$$

Then, we have

$$\begin{aligned}
& \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} y_{1,it-1}^{f} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} u_{1,it}^{f} \\
= & \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} \pi_{11} c_t^2 \left(w_{1,it-1} - \bar{w}_{1,itT} \right) \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} \left(u_{1,it} - \bar{u}_{1,it+1T} \right) \\
& + \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} \pi_{12} c_t^2 \left(w_{1,it-1} - \bar{w}_{1,itT} \right) \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} \left(u_{1,it} - \bar{u}_{1,it+1T} \right) \\
& + \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} c_t^2 \left(v_{1,it-1} - \bar{v}_{1,itT} \right) \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} \left(u_{1,it} - \bar{u}_{1,it+1T} \right) \\
= & I_1 + I_2 + I_3, \text{ say,}
\end{aligned}$$

For the first term I_1, we have

$$\begin{aligned}
E\left(I_1\right) &= E\left(\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} \pi_{11} c_t^2 \left(w_{1,it-1} - \bar{w}_{1,itT} \right) \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} \left(u_{1,it} - \bar{u}_{1,it+1T} \right) \right) \\
&= -\frac{\pi_{11}}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} E\left(\bar{w}_{1,it-1T} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} \left(u_{1,it} - \bar{u}_{1,it+1T} \right) \right) + o\left(1\right) \\
&= -\frac{\pi_{11}}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} \frac{1}{T-t} \sum_{s=t}^{T-1} E\left(w_{1,is} u_{1,it} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} \right) \\
&\quad + \frac{\pi_{11}}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} \frac{1}{\left(T-t\right)^2} \sum_{s_1 \geq s_2} E\left(w_{1,is_1} u_{1,is_2} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} \right) + o\left(1\right) \\
&= I_{11} + I_{21} + o\left(1\right),
\end{aligned}$$

where

$$\begin{aligned}
I_{11} &= -\frac{\pi_{11}}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} \frac{1}{T-t} \sum_{s=t-1}^{T-1} E\left(w_{1,is} u_{1,it} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} \right) \\
&= -\frac{\pi_{11}}{\sqrt{NT}} \sum_{t=1}^{T-1} \frac{1}{T-t} E\left(\sum_{i=1}^{N} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} \right) \sum_{s=t}^{T-1} \left(1,0\right) \Pi^{s-t} E\left(\mathbf{v}_{it} \mathbf{v}_{it}' \right) \left(1,0\right)' \\
&= -\frac{C}{\sqrt{NT}} \sum_{t=1}^{T-1} \frac{2t}{T-t} \\
&= o\left(1\right),
\end{aligned}$$

under restriction (10). The above holds since for $s > t$, from (A.2) and (A.19),

$$\mathbf{w}_{is} = \Pi^{s-t-1} \mathbf{w}_{it-1} + \mathbf{v}_{is} + \Pi \mathbf{v}_{is-1} + \cdots + \Pi^{s-t} \mathbf{v}_{it},$$

then

$$\begin{aligned}
E\left(w_{1,is} u_{1,it} \right) &= \left(1,0\right) E\left(\mathbf{w}_{is} \mathbf{v}_{it}' \right) \left(1,-\beta\right)' \\
&= \left(1,0\right) \Pi^{s-t} E\left(\mathbf{v}_{it} \mathbf{v}_{it}' \right) \left(1,-\beta\right)',
\end{aligned}$$

and since the process is stationary by assumption 2. Also, we have

$$
\begin{aligned}
E\left(\sum_{i=1}^{N} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \mathbf{z}_{it-1}\right) &= tr\left(E\left(\sum_{i=1}^{N} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \mathbf{z}_{it-1}\right)\right) \\
&= \sum_{j=1}^{N} E\left(tr\left(\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \mathbf{z}_{it-1}\mathbf{z}_{it-1}'\right)\right) \\
&= tr\left(E\left(\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)\right)\right) \\
&= tr\left(I_{2t}\right) \\
&= 2t.
\end{aligned}
$$

For I_{12}, we have

$$
\begin{aligned}
I_{12} &= \frac{\pi_{11}}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} \frac{1}{(T-t)^2} \sum_{s_1 \geq s_2} E\left(w_{1,is_1} u_{1,is_2} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \mathbf{z}_{it-1}\right) \\
&= \frac{\pi_{11}}{\sqrt{NT}} \sum_{t=1}^{T-1} \frac{1}{(T-t)^2} E\left(\sum_{i=1}^{N} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \mathbf{z}_{it-1}\right) \sum_{s_1 \geq s_2} (1,0)\, \Pi^{s_1-s_2} E\left(\mathbf{v}_{it}\mathbf{v}_{it}'\right)(1,0)' \\
&= \frac{\pi_{11}}{\sqrt{NT}} \sum_{t=1}^{T-1} \frac{2t}{(T-t)^2} \sum_{s_1 \geq s_2 > t} (1,0)\, \Pi^{s_1-s_2} E\left(\mathbf{v}_{it}\mathbf{v}_{it}'\right)(1,0)' \\
&= \frac{C}{\sqrt{NT}} \sum_{t=1}^{T-1} \frac{2t}{T-t} \\
&= o(1).
\end{aligned}
$$

Consequently, we have

$$
E\left(I_1\right) = o(1).
$$

For the variance of I_1, we notice that

$$\begin{aligned}
& Var\left(I_1\right) \\
= & \frac{\pi_{11}^2}{NT} Var\left(\sum_{t=1}^{T-1}\sum_{i=1}^{N}\left(w_{1,it-1}-\bar{w}_{1,itT}\right)\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\left(u_{1,it}-\bar{u}_{1,it+1T}\right)\right) \\
= & \frac{\pi_{11}^2}{NT}\sum_{t_1,t_2}\sum_{i_1,i_2} E\left(\begin{array}{c}\left(w_{1,i_1t_1-1}-\bar{w}_{1,i_1t_1T}\right)\mathbf{z}_{i_1t_1-1}'\left(\mathbf{Z}_{t_1-1}'\mathbf{Z}_{t_1-1}\right)^{-1}\mathbf{z}_{i_1t_1-1}\left(u_{1,i_1t_1}-\bar{u}_{1,i_1t_1+1T}\right) \\ \times\left(w_{1,i_2t_2-1}-\bar{w}_{1,i_2t_2T}\right)\mathbf{z}_{i_2t_2-1}'\left(\mathbf{Z}_{t_2-1}'\mathbf{Z}_{t_2-1}\right)^{-1}\mathbf{z}_{i_2t_2-1}\left(u_{1,i_2t_2}-\bar{u}_{1,i_2t_2+1T}\right)\end{array}\right) \\
= & \frac{\pi_{11}^2}{NT}\sum_{t_1,t_2}\sum_{i=1}^{N} E\left(\begin{array}{c}\left(w_{1,it_1-1}-\bar{w}_{1,it_1T}\right)\mathbf{z}_{it_1-1}'\left(\mathbf{Z}_{t_1-1}'\mathbf{Z}_{t_1-1}\right)^{-1}\mathbf{z}_{it_1-1}\left(u_{1,it_1}-\bar{u}_{1,it_1+1T}\right) \\ \times\left(w_{1,it_2-1}-\bar{w}_{1,it_2T}\right)\mathbf{z}_{it_2-1}'\left(\mathbf{Z}_{t_2-1}'\mathbf{Z}_{t_2-1}\right)^{-1}\mathbf{z}_{it_2-1}\left(u_{1,it_2}-\bar{u}_{1,it_2+1T}\right)\end{array}\right) \\
& +\frac{\pi_{11}^2}{NT}\sum_{t_1,t_2}\sum_{i_1\neq i_2}\begin{array}{c} E\left[\left(w_{1,i_1t_1-1}-\bar{w}_{1,i_1t_1T}\right)\mathbf{z}_{i_1t_1-1}'\left(\mathbf{Z}_{t_1-1}'\mathbf{Z}_{t_1-1}\right)^{-1}\mathbf{z}_{i_1t_1-1}\left(u_{1,i_1t_1}-\bar{u}_{1,i_1t_1+1T}\right)\right] \\ \times E\left[\left(w_{1,i_2t_2-1}-\bar{w}_{1,i_2t_2T}\right)\mathbf{z}_{i_2t_2-1}'\left(\mathbf{Z}_{t_2-1}'\mathbf{Z}_{t_2-1}\right)^{-1}\mathbf{z}_{i_2t_2-1}\left(u_{1,i_2t_2}-\bar{u}_{1,i_2t_2+1T}\right)\right]\end{array} \\
\leq & \frac{C}{N^3T}\sum_{i=1}^{N}\left[\sum_{t}\left[E\left(\begin{array}{c}\left(w_{1,it-1}-\bar{w}_{1,it,T}\right)^2\mathbf{z}_{it-1}'\left(\frac{1}{N}\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\right)^{-1}\mathbf{z}_{it-1} \\ \left(u_{1,it}-\bar{u}_{1,it+1T}\right)^2\mathbf{z}_{it-1}'\left(\frac{1}{N}\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\right)^{-1}\mathbf{z}_{it-1}\end{array}\right)\right]^{1/2}\right]^2+o\left(1\right) \\
\leq & \frac{C}{N^3T}\sum_{i=1}^{N}\left[\sum_{t}\left[E\left(\left(w_{1,it-1}-\bar{w}_{1,it,T}\right)^2\mathbf{z}_{it-1}'\mathbf{z}_{it-1}\left(u_{1,it}-\bar{u}_{1,it+1T}\right)^2\mathbf{z}_{it-1}'\mathbf{z}_{it-1}\right)\right]^{1/2}\right]^2+o\left(1\right) \\
\leq & \frac{C}{N^3T}\sum_{i=1}^{N}\left[\sum_{t}\left(\left[E\left[\left(w_{1,it-1}-\bar{w}_{1,it,T}\right)^8\right]E\left[\left(u_{1,it}-\bar{u}_{1,it+1T}\right)^8\right]\left(E\left(\mathbf{z}_{it-1}'\mathbf{z}_{it-1}\right)^4\right)^2\right]\right)^{1/8}\right]^2+o\left(1\right) \\
\leq & \frac{C}{N^2T}\sum_{t}t^2+o\left(1\right) \\
= & o\left(1\right),
\end{aligned}$$

since $E\left[\left(w_{1,it-1}-\bar{w}_{1,it,T}\right)^8\right]$ and $E\left[\left(u_{1,it}-\bar{u}_{1,it+1T}\right)^8\right]$ are finite under assumption 1.

As a result, we can conclude that

$$I_1 = o_p\left(1\right). \tag{A.23}$$

Similarly, we can show that

$$I_2 = o_p\left(1\right). \tag{A.24}$$

For I_3, we notice that

$$\begin{aligned}
E\left(I_3\right) &= \frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}E\left[\left(v_{1,it-1}-\bar{v}_{1,itT}\right)\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\left(u_{1,it}-\bar{u}_{1,it+1T}\right)\right]+o\left(1\right)\\
&= -\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\frac{1}{T-t}\sum_{s_1=t}^{T-1}E\left[v_{1,is_1}u_{1,it}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\right]\\
&\quad+\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\frac{1}{\left(T-t\right)^2}\sum_{s_1,s_2}E\left[v_{1,is_1}u_{1,is_2}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\right]+o\left(1\right)\\
&= -\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\frac{1}{T-t}E\left[v_{1,it}u_{1,it}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\right]\\
&\quad+\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\frac{1}{\left(T-t\right)}E\left[v_{1,it}u_{1,it}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\right]+o\left(1\right)\\
&= o\left(1\right),
\end{aligned}$$

and we can also show that

$$Var\left(I_3\right)=o\left(1\right),$$

by following the derivation above.

Combining the above results, we obtain

$$\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}y_{1,it-1}^{f}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}u_{1,it}^{f}\rightarrow_p 0. \tag{A.25}$$

For the second element of (A.18), we first notice that

$$\begin{aligned}
y_{2,it}^{f} &= w_{2,it}^{f}=\gamma_{21}w_{1,it-1}^{f}+\gamma_{22}w_{2,it-1}^{f}+v_{2,it}^{f}\\
&= \gamma_{21}c_t\left(w_{1,it-1}-\bar{w}_{1,itT}\right)+\gamma_{22}c_t\left(w_{2,it-1}-\bar{w}_{2,itT}\right)+c_t\left(v_{2,it}-\bar{v}_{2,it+1T}\right),
\end{aligned}$$

then we have

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}y_{2,it}^{f}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}u_{1,it}^{f}\\
=&\ \frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\gamma_{21}c_t^2\left(w_{1,it-1}-\bar{w}_{1,itT}\right)\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\left(u_{1,it}-\bar{u}_{1,it+1T}\right)\\
&+\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\gamma_{22}c_t^2\left(w_{2,it-1}-\bar{w}_{2,itT}\right)\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\left(u_{1,it}-\bar{u}_{1,it+1T}\right)\\
&+\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\sum_{i=1}^{N}c_t^2\left(v_{2,it}-\bar{v}_{2,it+1T}\right)\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\left(u_{1,it}-\bar{u}_{1,it+1T}\right)\\
=&\ I_4+I_5+I_6,\ \text{say,}
\end{aligned}$$

it can be shown by using the derivation above that

$$I_4\rightarrow_p 0,\ \text{and}\ I_5\rightarrow_p 0. \tag{A.26}$$

Using a similar argument for I_6, we obtain

$$
\begin{aligned}
I_6 &= \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} c_t^2 \left(v_{2,it} - \bar{v}_{2,it+1T}\right) \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \mathbf{z}_{it-1} \left(u_{1,it} - \bar{u}_{1,it+1T}\right) \\
&= \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} v_{2,it} u_{1,it} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \mathbf{z}_{it-1} + o_p(1) \\
&= \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} u_{2,it} u_{1,it} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \mathbf{z}_{it-1} + o_p(1) \\
&= \frac{\sigma_{u,12}}{\sqrt{NT}} \sum_{t=1}^{T-1} 2t + o_p(1) \\
&\rightarrow {}_p \sigma_{u,12} \sqrt{\kappa},
\end{aligned}
$$

since $v_{2,it} = u_{2,it}$ from (A.22) and by applying lemma (A.4).

Consequently, we can get

$$
\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} y_{2,it}^{f} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \mathbf{z}_{it-1} u_{1,it}^{f} \rightarrow_p \sigma_{u,12} \sqrt{\kappa}, \tag{A.27}
$$

which in turn yields

$$
\frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \left[\sum_{i=1}^{N} \mathbf{y}_{i,t-1 \cdot t}^{f\prime} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1} \mathbf{z}_{it-1} u_{1,it}^{f} \right] \rightarrow_p \begin{pmatrix} 0 \\ \sigma_{u,12} \end{pmatrix} \sqrt{\kappa},
$$

as required. ■

Lemma A.8 *Under Assumptions 1-4, as well as the condition (10), we have*

(a) For FD case, we have

$$
\frac{1}{NT} \sum_{t=2}^{T} \left[\sum_{i=1}^{N} \Delta \mathbf{y}_{i,t-1 \cdot t}' \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1} \mathbf{z}_{it-2} \Delta \mathbf{y}_{i,t-1 \cdot t} \right] = o_p(1).
$$

(b) For the FD, we have

$$
\frac{1}{\sqrt{NT}} \sum_{t=2}^{T} \sum_{i=1}^{N} \Delta \mathbf{y}_{i,t-1 \cdot t}' \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1} \mathbf{z}_{it-2} \Delta u_{1,it} \rightarrow_p - \left[\mathbf{D}' \begin{pmatrix} \sigma_{u,1}^2 + \beta \sigma_{u,12} \\ \sigma_{u,21} \end{pmatrix} - \begin{pmatrix} 0 \\ 2\sigma_{u,21} \end{pmatrix} \right] \sqrt{\kappa}.
$$

Proof. (a) To show this result, we first notice that

$$
\frac{1}{NT} \sum_{t=2}^{T} \left[\sum_{i=1}^{N} \Delta \mathbf{y}_{i,t-1 \cdot t}' \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1} \mathbf{z}_{it-2} \Delta \mathbf{y}_{i,t-1 \cdot t}' \right] = \frac{1}{NT} \sum_{t=2}^{T}
$$

$$
\begin{pmatrix} \sum_{i=1}^{N} \Delta y_{1,it-1} \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1} \mathbf{z}_{it-2} \Delta y_{1,it-1} & \sum_{i=1}^{N} \Delta y_{1,it-1} \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1} \mathbf{z}_{it-2} \Delta y_{2,it} \\ \sum_{i=1}^{N} \Delta y_{2,it} \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1} \mathbf{z}_{it-2} \Delta y_{1,it-1} & \sum_{i=1}^{N} \Delta y_{2,it} \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1} \mathbf{z}_{it-2} \Delta y_{2,it} \end{pmatrix}, \tag{A.28}
$$

as before, we need to show that each element of (A.28) has zero limit. For the (1,1)-th element of (A.28), we have

$$\begin{aligned}
&\frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}\Delta y_{1,it-1}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\Delta y_{1,it-1} \\
=\;& \frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}\left(w_{1,it-1}-w_{1,it-2}\right)^{2}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2} \\
=\;& \frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}w_{1,it-1}^{2}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}+\frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}w_{1,it-2}^{2}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2} \\
&-\frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}2w_{1,it-1}w_{1,it-2}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2},
\end{aligned}$$

by following the derivation of part (b) of lemma (A.7), we can show that

$$\begin{aligned}
\frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}w_{1,it-1}^{2}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2} &= o_p(1), \\
\frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}w_{1,it-2}^{2}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2} &= o_p(1), \\
\frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}2w_{1,it-1}w_{1,it-2}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2} &= o_p(1),
\end{aligned}$$

which gives

$$\frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}\Delta y_{1,it-1}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\Delta y_{1,it-1}=o_p(1).$$

Similarly, we can show all other elements of (A.28) have zero probability limit, i.e.,

$$\frac{1}{NT}\sum_{t=2}^{T}\sum_{i=1}^{N}\Delta\mathbf{y}_{i,t-1\cdot t}'\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\Delta\mathbf{y}_{i,t-1\cdot t}'=o_p(1),$$

as required.

(b) In order to show this, we notice that

$$\begin{aligned}
&\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left[\sum_{i=1}^{N}\Delta\mathbf{y}_{i,t-1\cdot t}'\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\Delta u_{1,it}\right] \\
=\;& \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left[\sum_{i=1}^{N}\begin{pmatrix}\Delta w_{1,it-1}\\ \Delta w_{2,it}\end{pmatrix}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\Delta u_{1,it}\right] \\
=\;& \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left[\sum_{i=1}^{N}\mathbf{D}'\Delta\mathbf{w}_{i,t-1}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\Delta u_{1,it}\right] \\
&+\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left[\sum_{i=1}^{N}\left(0,\Delta v_{2,it}\right)'\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\Delta u_{1,it}\right], \qquad \text{(A.29)}
\end{aligned}$$

which holds since

$$\left(\begin{array}{c}\Delta w_{1,it-1}\\ \Delta w_{2,it}\end{array}\right) = \left(\begin{array}{c}\Delta w_{1,it-1}\\ \pi_{21}\Delta w_{1,it-1}+\pi_{22}\Delta w_{2,it-1}\end{array}\right)+\left(\begin{array}{c}0\\ \Delta v_{2,it}\end{array}\right)$$
$$= \mathbf{D}'\Delta\mathbf{w}_{i,t-1}+\left(\begin{array}{c}0\\ \Delta v_{2,it}\end{array}\right),$$

by using (A.4).

For the first element of (A.29), we have

$$\begin{aligned}
& \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left[\sum_{i=1}^{N}\mathbf{D}'\Delta\mathbf{w}_{i,t-1}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\Delta u_{1,it}\right] \\
= & -\mathbf{D}'\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left[\sum_{i=1}^{N}\mathbf{w}_{i,t-1}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}u_{1,it-1}\right] \\
= & -\mathbf{D}'\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left[\sum_{i=1}^{N}\mathbf{w}_{i,t-1}u_{1,it-1}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\right] \\
= & -\mathbf{D}'\mathbf{B}^{-1}\Omega_u\left(1,0\right)'\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}2\left(t-1\right)+o_p\left(1\right) \\
\rightarrow_p & -\mathbf{D}'\mathbf{B}^{-1}\Omega_u\left(1,0\right)'\sqrt{\kappa},
\end{aligned} \tag{A.30}$$

as $(N,T)\rightarrow\infty$ and under (10), since

$$\begin{aligned}
E\left(\mathbf{w}_{i,t-1}u_{1,it-1}\right) &= E\left(\mathbf{v}_{i,t-1}\mathbf{u}_{it-1}'\right)\left(1,0\right)' \\
&= \mathbf{B}^{-1}E\left(\mathbf{u}_{it}\mathbf{u}_{it}'\right)\left(1,0\right)' \\
&= \mathbf{B}^{-1}\Omega_u\left(1,0\right)' = \left(\begin{array}{c}\sigma_{u,1}^2+\beta\sigma_{u,12}\\ \sigma_{u,12}\end{array}\right),
\end{aligned}$$

and by following the derivation of (A.4).

Similarly, for the second element of (A.29), as $(N,T)\rightarrow\infty$ and under (10), we have

$$\begin{aligned}
& \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left[\sum_{i=1}^{N}\left(\begin{array}{c}0\\ \Delta v_{2,it}\end{array}\right)\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\Delta u_{1,it}\right] \\
= & \left(\begin{array}{c}0\\ \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\sum_{i=1}^{N}\Delta v_{2,it}\Delta u_{1,it}\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\end{array}\right) \\
= & \left(0,1\right)\mathbf{B}^{-1}\Omega_u\left(1,0\right)'\left(\begin{array}{c}0\\ \frac{1}{\sqrt{NT}}\sum_{t=2}^{T}2\left(t-1\right)\end{array}\right)+o_p\left(1\right) \\
\rightarrow_p & \left(\begin{array}{c}0\\ 2\sigma_{u,21}\end{array}\right)\sqrt{\kappa},
\end{aligned} \tag{A.31}$$

since

$$\begin{aligned} E\left(\Delta v_{2,it}\Delta u_{1,it}\right) &= 2E\left(v_{2,it}u_{1,it}\right) \\ &= 2\left(0,1\right)E\left(\mathbf{v}_{it}\mathbf{u}_{it}'\right)\left(1,0\right)' \\ &= 2\left(0,1\right)\mathbf{B}^{-1}E\left(\mathbf{u}_{it}\mathbf{u}_{it}'\right)\left(1,0\right)' \\ &= 2\left(0,1\right)\mathbf{B}^{-1}\Omega_u\left(1,0\right)' \\ &= 2\sigma_{u,21}. \end{aligned}$$

Combining (A.30) and (A.31) yields

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\left[\sum_{i=1}^{N}\Delta\mathbf{y}_{i,t-1\cdot t}'\mathbf{z}_{it-2}'\left(\mathbf{Z}_{t-2}'\mathbf{Z}_{t-2}\right)^{-1}\mathbf{z}_{it-2}\Delta u_{1,it}\right]\to_p -\left[\mathbf{D}'\begin{pmatrix}\sigma_{u,1}^2+\beta\sigma_{u,12}\\ \sigma_{u,21}\end{pmatrix}-\begin{pmatrix}0\\ 2\sigma_{u,21}\end{pmatrix}\right]\sqrt{\kappa},$$

as required. ■

A.2 Derivation of GMM based on FOD

In order to show the results of Theorem 1, we notice that

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{GMM}^{FOD}-\boldsymbol{\theta}\right)=\left(\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{f\prime}\mathbf{P}_{t-1}\mathbf{Y}_{t-1\cdot t}^{f}\right)^{-1}\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{f\prime}\mathbf{P}_{t-1}\mathbf{u}_{1,t}^{f}, \tag{A.32}$$

where for the denominator, by using the results in part (a) of lemma (A.5), we obtain

$$\begin{aligned} \frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{f\prime}\mathbf{P}_{t-1}\mathbf{Y}_{t-1\cdot t}^{f} &= \frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-1}\mathbf{Y}_{t-1\cdot t}+o_p\left(1\right) \\ &\to {}_p\mathbf{D}'\boldsymbol{\Gamma}_0\mathbf{D}, \end{aligned} \tag{A.33}$$

where $\boldsymbol{\Gamma}_0=\sum_{i=0}^{\infty}\Pi^i\Omega_v\Pi^{i\prime}=\sum_{i=0}^{\infty}\Pi^i\mathbf{B}^{-1}\Omega_u\mathbf{B}'^{-1}\Pi^{i\prime}$ and

$$\mathbf{D}=\begin{pmatrix}1 & \pi_{21}\\ 0 & \pi_{22}\end{pmatrix}=\begin{pmatrix}1 & \beta\gamma_{22}\\ 0 & \gamma_{22}\end{pmatrix}. \tag{A.34}$$

For the numerator of (A.32), by using the results in part (b) of lemma (A.5), we have

$$\begin{aligned} \frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{f\prime}\mathbf{P}_{t-1}\mathbf{u}_{1,t}^{f} &= \frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-1}\mathbf{u}_{1,t}+o_p\left(1\right) \\ &\to {}_dN\left(0,\sigma_{u,1}^2\mathbf{D}'\Gamma_0\mathbf{D}\right)-\begin{pmatrix}0\\ \sigma_{u,12}\end{pmatrix}\sqrt{\kappa}, \end{aligned} \tag{A.35}$$

where $\kappa=\frac{T^3}{N}<\infty$ as $N\to\infty$.

Substituting (A.33) and (A.35) into (A.32) yields the required result.

A.3 Derivation of GMM based on FD

For the results of Theorem 2, we notice that

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{GMM}^{FD}-\boldsymbol{\theta}\right)=\left(\frac{1}{NT}\sum_{t=2}^{T}\Delta\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-2}\Delta\mathbf{Y}_{t-1\cdot t}\right)^{-1}\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\Delta\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t},\quad\text{(A.36)}$$

where $\Delta\mathbf{u}_{1,t}=(\Delta u_{1,1t},\ldots,\Delta u_{1,Nt})'$.

Using the results of lemma (A.6), as $(N,T)\to\infty$, we have

$$\frac{1}{NT}\sum_{t=2}^{T}\Delta\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-2}\Delta\mathbf{Y}_{t-1\cdot t}\to_p\mathbf{D}'\left(I_2-\Pi\right)\Gamma_0\left(I_2-\Pi'\right)\mathbf{D},\quad\text{(A.37)}$$

where $\mathbf{D}$ and Π are defined in (A.34) and (2), respectively, and

$$\frac{1}{\sqrt{NT}}\sum_{t=2}^{T}\Delta\mathbf{Y}_{t-1\cdot t}'\mathbf{P}_{t-2}\Delta\mathbf{u}_{1,t}\to_d N\left(\tilde{\mathbf{b}}_0,\sigma_{u,1}^2\mathbf{D}'\left(I_2-\Pi\right)\left[\left(I_2-\Pi\right)\Gamma_0+\Gamma_0\left(I_2-\Pi'\right)\right]\left(I_2-\Pi'\right)\mathbf{D}\right),\quad\text{(A.38)}$$

where

$$\tilde{\mathbf{b}}_0=-\left[\mathbf{D}'\begin{pmatrix}\sigma_{u,1}^2+\beta\sigma_{u,12}\\ \sigma_{u,21}\end{pmatrix}-\begin{pmatrix}0\\ 2\sigma_{u,21}\end{pmatrix}\right]\sqrt{\kappa}.\quad\text{(A.39)}$$

Substituting (A.37) and (A.38) into (A.36) yields the required result.

A.4 Derivation of JIVE based on FOD

In order to show the results of Theorem 3, we first notice that

$$\begin{aligned}\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{JIVE}^{FOD}-\boldsymbol{\theta}\right)&=\left(\frac{1}{NT}\sum_{t=1}^{T-1}\left[\sum_{j=1}^{N}\sum_{i\neq j}\mathbf{y}_{i,t-1\cdot t}^{f\prime}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{jt-1}\mathbf{y}_{j,t-1\cdot t}^{f}\right]\right)^{-1}\\&\quad\times\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}\left[\sum_{j=1}^{N}\sum_{i\neq j}\mathbf{y}_{i,t-1\cdot t}^{f\prime}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{jt-1}u_{1,jt}^{f}\right].\quad\text{(A.40)}\end{aligned}$$

For the denominator of (A.40), we have

$$\begin{aligned}&\frac{1}{NT}\sum_{t=1}^{T-1}\sum_{j=1}^{N}\sum_{i\neq j}\mathbf{y}_{i,t-1\cdot t}^{f\prime}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{jt-1}\mathbf{y}_{j,t-1\cdot t}^{f}\\=&\ \frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{f\prime}\mathbf{P}_{t-1}\mathbf{Y}_{t-1\cdot t}^{(f)}-\frac{1}{NT}\sum_{t=1}^{T-1}\left[\sum_{i=1}^{N}\mathbf{y}_{i,t-1\cdot t}^{f\prime}\mathbf{z}_{it-1}'\left(\mathbf{Z}_{t-1}'\mathbf{Z}_{t-1}\right)^{-1}\mathbf{z}_{it-1}\mathbf{y}_{i,t-1\cdot t}^{f}\right]\\=&\ \frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t-1\cdot t}^{f\prime}\mathbf{P}_{t-1}\mathbf{Y}_{t-1\cdot t}^{(f)}+o_p(1)\\\to&\ _p\mathbf{D}'\boldsymbol{\Gamma}_0\mathbf{D},\quad\text{(A.41)}\end{aligned}$$

by using the result of part (a) from both lemma (A.5) and lemma (A.7).

For the numerator,

$$\begin{aligned}
& \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} \sum_{j \neq i} \mathbf{y}_{i,t-1 \cdot t}^{f\prime} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{jt-1} u_{1,jt}^{f} \\
= \;& \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \mathbf{Y}_{t-1 \cdot t}^{(f)\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t}^{f} - \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \sum_{i=1}^{N} \mathbf{y}_{i,t-1 \cdot t}^{f\prime} \mathbf{z}_{it-1}' \left(\mathbf{Z}_{t-1}' \mathbf{Z}_{t-1} \right)^{-1} \mathbf{z}_{it-1} u_{1,it}^{f} \\
= \;& \frac{1}{\sqrt{NT}} \sum_{t=1}^{T-1} \mathbf{Y}_{t-1 \cdot t}^{(f)\prime} \mathbf{P}_{t-1} \mathbf{u}_{1,t}^{f} - \begin{pmatrix} 0 \\ \sigma_{u,12} \end{pmatrix} \sqrt{\kappa} + o_p(1) \\
\to \;& {}_d N\left(0, \sigma_{u,1}^2 \mathbf{D}' \mathbf{\Gamma}_0 \mathbf{D}\right), \qquad \text{(A.42)}
\end{aligned}$$

by using the results of part (b) from both lemma (A.5) and lemma (A.7).

As a result, combining equations (A.41) and (A.42), we obtain the result of Theorem 3 as required.

A.5 Derivation of JIVE based on FD

Finally, for the results of Theorem 4, we have

$$\begin{aligned}
\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{JIVE}^{FD} - \boldsymbol{\theta}\right) = \;& \left(\frac{1}{NT} \sum_{t=2}^{T} \sum_{j=1}^{N} \sum_{i \neq j} \Delta \mathbf{y}_{i,t-1 \cdot t}' \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}' \mathbf{Z}_{t-2} \right)^{-1} \mathbf{z}_{jt-2} \Delta \mathbf{y}_{j,t-1 \cdot t} \right)^{-1} \\
& \times \frac{1}{\sqrt{NT}} \sum_{t=2}^{T} \sum_{j=1}^{N} \sum_{i \neq j} \Delta \mathbf{y}_{i,t-1 \cdot t}' \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}' \mathbf{Z}_{t-2} \right)^{-1} \mathbf{z}_{jt-2} \Delta u_{1,jt}, \qquad \text{(A.43)}
\end{aligned}$$

where for the denominator, we have

$$\begin{aligned}
& \frac{1}{NT} \sum_{t=2}^{T} \sum_{j=1}^{N} \sum_{i \neq j} \Delta \mathbf{y}_{i,t-1 \cdot t}' \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}' \mathbf{Z}_{t-2} \right)^{-1} \mathbf{z}_{jt-2} \Delta \mathbf{y}_{j,t-1 \cdot t} \\
= \;& \frac{1}{NT} \sum_{t=2}^{T} \Delta \mathbf{Y}_{t-1 \cdot t}' \mathbf{P}_{t-2} \Delta \mathbf{Y}_{t-1 \cdot t} - \frac{1}{NT} \sum_{t=2}^{T} \left[\sum_{i=1}^{N} \Delta \mathbf{y}_{i,t-1 \cdot t}' \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}' \mathbf{Z}_{t-2} \right)^{-1} \mathbf{z}_{it-2} \Delta \mathbf{y}_{i,t-1 \cdot t} \right] \\
= \;& \frac{1}{NT} \sum_{t=2}^{T} \Delta \mathbf{Y}_{t-1 \cdot t}' \mathbf{P}_{t-2} \Delta \mathbf{Y}_{t-1 \cdot t} + o_p(1) \\
\to \;& {}_p \mathbf{D}' \left(I_2 - \Pi \right) \Gamma_0 \left(I_2 - \Pi' \right) \mathbf{D}, \qquad \text{(A.44)}
\end{aligned}$$

by using the result of part (a) from both lemma (A.6) and lemma (A.8). For the numerator,

$$\begin{aligned}
& \frac{1}{\sqrt{NT}} \sum_{t=2}^{T} \sum_{j=1}^{N} \sum_{i \neq j} \Delta \mathbf{y}_{i,t-1 \cdot t}' \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}' \mathbf{Z}_{t-2} \right)^{-1} \mathbf{z}_{jt-2} \Delta u_{1,jt} \\
= \;& \frac{1}{\sqrt{NT}} \sum_{t=2}^{T} \Delta \mathbf{Y}_{t-1 \cdot t}' \mathbf{P}_{t-2} \Delta \mathbf{u}_{1,t} - \frac{1}{\sqrt{NT}} \sum_{t=2}^{T} \sum_{i=1}^{N} \Delta \mathbf{y}_{i,t-1 \cdot t}' \mathbf{z}_{it-2}' \left(\mathbf{Z}_{t-2}' \mathbf{Z}_{t-2} \right)^{-1} \mathbf{z}_{it-2} \Delta u_{1,it} \\
= \;& \frac{1}{\sqrt{NT}} \sum_{t=2}^{T} \Delta \mathbf{Y}_{t-1 \cdot t}' \mathbf{P}_{t-2} \Delta \mathbf{u}_{1,t} - \tilde{\mathbf{b}}_0 + o_p(1) \\
\to \;& {}_d N\left(0, \sigma_{u,1}^2 \mathbf{D}' \left(I_2 - \Pi \right) \left[\left(I_2 - \Pi \right) \Gamma_0 + \Gamma_0 \left(I_2 - \Pi' \right) \right] \left(I_2 - \Pi' \right) \mathbf{D} \right), \qquad \text{(A.45)}
\end{aligned}$$

by using the results of part (b) from both lemma (A.6) and lemma (A.8).

Consequently, combining equations (A.44) and (A.45), we obtain the result of Theorem 4 as required.