Some Properties of the LIML Estimator in a Dynamic Panel Structural Equation

Kentaro Akashi
Institute of Statistical Mathematics

Naoto Kunitomo
University of Tokyo

January 2010

Some Properties of the LIML Estimator in a Dynamic Panel Structural Equation [1]

Kentaro Akashi [2]

and
Naoto Kunitomo [3]

January 25, 2010

abstract

We investigate the finite sample and asymptotic properties of several estimation methods (Within-Groups, GMM and LIML) for a panel autoregressive structural equation model with random effects when both T and N are large. When we use the forward filtering to a structural model as Alvarez and Arellano (2003), both the WG and GMM estimators are significantly biased when both T and N go to infinity while T/N is different from zero. The LIML (limited information maximum likelihood) estimator has consistency and the asymptotic normality when T/N converges to a constant as both T and N go to infinity. Its asymptotic distribution has some bias and covariance which depend on the limiting behavior of T/N.

Keywords : Dynamic Panel Model, Simultaneous Equation, Within-Groups Estimator, GMM, LIML (limited information maximum likelihood), Many Orthogonal Conditions.

1   Introduction

Recently there has been a growing interest on panel econometric models in micro-econometrics and they are indispensable tools for econometric analysis. (See Hsiao (2003), Arellano (2003) and Baltagi (2005), for instance.) However, there are still non-trivial statistical problems of estimating dynamic panel econometric models to be investigated. When we use the lagged explained variables as well as other explanatory variables with individual effects in panel regression models, there could be a natural question among econometricians on what would happen if one of variables was actually endogenous in the economic system. When we have an endogenous variable in the dynamic panel models with individual effects, it would not be obvious how to estimate such a particular structural equation because some complicated interactions would be occurred by the lagged endogenous variables and the individual effects in the econometric model at the same time. Earlier

---

[1] AK10-1-25-2.

[2] Institute of Statistical Mathematics

[3] Graduate School of Economics, University of Tokyo

1

investigations on some aspect of the dynamic panel models were Anderson and Hsiao (1981, 1982).

In a pioneering work Alvarez and Allerano (2003) have investigated the asymptotic behaviors of alternative estimation methods, namely the WG (Within-Groups), the GMM and the LIML (limited information maximum likelihood) estimators, for a coefficient in a dynamic panel regression model when both N (the number of individuals) and T (the number of observation periods) go to infinity. They have investigated the asymptotic properties of estimators when both N and T go to infinity and derived the asymptotic distributions of these estimators. Although they have obtained interesting findings, however, one remaining major issue in econometrics is to investigate the effects of the endogeneity of possible explanatory variables in the dynamic panel structural equations. One important aspect in this problem is the fact that when there are many orthogonal conditions in dynamic panel models except some cases when T is really small, the use of GMM would be problematic due to *incidental parameters* in the recent light on the estimation of structural equations in econometric studies as Anderson, Kunitomo and Matsushita (2008a, b).

The main purpose of this paper is to investigate the finite sample and the asymptotic properties of three estimators (the WG, the GMM and the LIML estimators) for a coefficient in a specific panel dynamic structural equation. The model we shall consider is intentionally very simple because it is possible to obtain the precise information of the finite sample as well as the asymptotic properties of alternative estimators, which would be useful for practical problems eventually. In a companion paper (Akashi and Kunitomo (2010)), we shall develop the general formulation of the estimation methods of dynamic panel structural equation with endogeneity, individual effects and many orthogonal conditions. They have tried to draw rather general results including the asymptotic distributions of alternative estimation methods and the asymptotic optimality of estimation, of which the results would be rather complicated at the first glance. In order to make our expositions of our general results useful in a meaningful and persuasive way, this paper utilizes a particular dynamic panel structural equation with an endogenous variable and individual effects as the typical case, which was originally used by Blundell and Bond (2000).

In Section 2, we present the panel structural model and define its alternative estimation methods. In Section 3 we shall establish the asymptotic properties of three estimators considered and discuss their asymptotic behaviors. Then in Section 4 we also discuss the finite sample properties of estimators based on their empirical distribution functions in the Monte Carlo simulations. In Section 5, some concluding remarks will be presented. All mathematical derivations of our

theoretical results are in Section 6 and some figures on the distribution functions of normalized alternative estimators are in Appendix.

2 The Panel Model and Estimation Methods

We consider a dynamic panel model with an endogenous variable (Blundell and Bond (2000))

$$y_{it}^{(1)} = \beta_2 y_{it}^{(2)} + \gamma_1 y_{it-1}^{(1)} + \eta_i + u_{it}^{(1)}, \tag{2.1}$$

$$y_{it}^{(2)} = \gamma_2 y_{it-1}^{(2)} + \delta\eta_i + u_{it}^{(2)}, \tag{2.2}$$

for $i = 1, ..., N$; $t = 1, ..., T$, where $(u_{it}^{(1)}, u_{it}^{(2)})$ are the disturbance terms. In the first structural equation, there is an endogenous variable $y_{it}^{(2)}$ with a lagged endogenous regressor $y_{it-1}^{(1)}$ and individual effects η_i. Define the reduced form

$$\mathbf{y}_{it} = \mathbf{\Pi}\mathbf{y}_{it-1} + \boldsymbol{\pi}_i + \mathbf{v}_{it} , \tag{2.3}$$

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

$$\mathbf{\Pi} = \mathbf{B}^{-1}\begin{pmatrix} \gamma_1 & 0 \\ 0 & \gamma_2 \end{pmatrix}, \ \mathbf{B} = \begin{pmatrix} 1 & -\beta_2 \\ 0 & 1 \end{pmatrix}, \ \boldsymbol{\pi}_i = \mathbf{B}^{-1}\begin{pmatrix} 1 \\ \delta \end{pmatrix}\eta_i, \ \mathbf{v}_{it} = \mathbf{B}^{-1}\begin{pmatrix} u_{it}^{(1)} \\ u_{it}^{(2)} \end{pmatrix}.$$

By using the forward-filters [4] to both sides of (2.3), we obtain the forward-filtered reduced form

$$\mathbf{y}_{it}^{(f)} = \mathbf{\Pi}\mathbf{y}_{it-1}^{(f)} + \mathbf{v}_{it}^{(f)}, \tag{2.4}$$

where the superscript (f) denotes the forward-filtered variables, which are free from the individual effects.

Although we call the equation (2.3) or (2.4) the *reduced form*, $\mathbf{y}_{it-1}$ is correlated with unobserved $\boldsymbol{\pi}_i$ in (2.3), and $\mathbf{y}_{it-1}^{(f)}$ is also correlated with $\mathbf{v}_{it}^{(f)}$ in (2.4) by the consequence of applying the forward filtering. It is an important difference from the standard simultaneous equation problems. Alvarez and Arellano (2003) considered the single equation (2.1) when $\beta_2 = 0$ and investigate the estimator for a coefficient γ_1 without (2.2). They have shown that three estimators the WG, the GMM and the LIML estimators are consistent and have the same asymptotic

---

[4]The forward-filters are defined by the set of transformations such that the elements $z_{it}^{(f)}$ from z_{is} has the form

$$z_{it}^{(f)} = c_t[z_{it} - (\frac{1}{T-t})(z_{it+1} + ... + z_{iT})] \quad (t = 1, ..., T-1),$$

with $c_t^2 = (T-t)/(T-t+1)$.

variance when both N and T go to infinity while their ratio converges to a constant. An interpretation of this aspect is the lack of the endogeneity in the reduced form (2.4) with the particular equation as we shall discuss later. As for a dynamic structural equation problem, the parameters of interests are both β_2 and γ_1, and we shall forcus on three estimators for these parameters in this paper.

2.1 The Within-Groups Estimator

Define the forward-filtered variables $\mathbf{y}_t^{(g,f)} = (y_{1t}^{(g,f)}, ..., y_{Nt}^{(g,f)})'$ for the two endogenous variables $y_{it}^{(g)}$ ($g = 1, 2$), and set $N \times 2$ vector $\mathbf{Y}_{t \cdot t-1}^{(f)} = (\mathbf{y}_t^{(2,f)}, \mathbf{y}_{t-1}^{(1,f)})$. Then the WG (within-groups) estimator for $\boldsymbol{\theta} = (\beta_2, \gamma_1)'$ is given by

$$\hat{\boldsymbol{\theta}}_{WG} = (\sum_{t=1}^{T-1} \mathbf{Y}_{t \cdot t-1}^{(f)'} \mathbf{Y}_{t \cdot t-1}^{(f)})^{-1} (\sum_{t=1}^{T-1} \mathbf{Y}_{t \cdot t-1}^{(f)'} \mathbf{y}_t^{(1,f)}). \tag{2.5}$$

where the forward-filtered variables are operated by the forward orthogonal deviations $(T-1) \times T$ matrix $\mathbf{A}_f$, such that $\mathbf{A}_f' \mathbf{A}_f = \mathbf{Q}_T, \mathbf{A}_f \mathbf{A}_f = \mathbf{I}_{T-1}$, $\mathbf{Q}_T = \mathbf{I}_T - \boldsymbol{\iota}_T \boldsymbol{\iota}_T'/T$ (the WG operator) and $\boldsymbol{\iota}_T$ is a $T \times 1$ vector whose elements are ones. The form (2.5) is written as the OLS estimator in terms of the orthogonal deviations and notice that since $\mathbf{A}_f \boldsymbol{\iota}_T = \mathbf{0}$, the individual effects are differenced out in the orthogonal deviations for the associated variables from the original observations.

2.2 A GMM estimator

A GMM (generalized method of moments) estimator is given by

$$\hat{\boldsymbol{\theta}}_{GMM} = (\sum_{t=1}^{T-1} \mathbf{Y}_{t \cdot t-1}^{(f)'} \mathbf{M}_t \mathbf{Y}_{t \cdot t-1}^{(f)})^{-1} (\sum_{t=1}^{T-1} \mathbf{Y}_{t \cdot t-1}^{(f)'} \mathbf{M}_t \mathbf{y}_t^{(1,f)}), \tag{2.6}$$

where $\mathbf{M}_t = \mathbf{Z}_t (\mathbf{Z}_t' \mathbf{Z}_t)^{-1} \mathbf{Z}_t'$, and $\mathbf{Z}_t = (\mathbf{y}_0^{(1)}, \mathbf{y}_0^{(2)}, ..., \mathbf{y}_{t-1}^{(1)}, \mathbf{y}_{t-1}^{(2)})$ are $N \times 2t$ instrumental variables matrix. The GMM estimator is identical to the one given by Arellano and Bond (1991) [5], which have the form written in the orthogonal deviations. An interesting feature of this estimation method is to use all available instrumental variables at each t, therefore the orthogonal conditions can be given by

$$\mathcal{E}[y_{is}^{(g)} u_{it}^{(1,f)}] = 0 \qquad (0 \le s < t = 1, ..., T-1; \ g = 1, 2),$$

---

[5] See Chapter 4 of Hsiao (2003), for instance.

where $u_{it}^{(1,f)}$ stands for the forward-filtered structural error term. In this formulation the number of the total orthogonal conditions r_n can be often substantial, i.e., $r_n = 2 \times T(T-1)/2$. In this paper we use the notation that the total number of observations $n = NT$ and r_n can be dependent on n.

2.3 The LIML estimator

The LIML (limited information maximum likelihood) estimator was originally developed by Anderson and Rubin (1949, 1950) for the classical simultaneous equation problem, and we shall apply this estimation method to the filtered variables in a dynamic panel structural equation model.

Define two 3×3 matrices by

$$\mathbf{G}^{(f)} = \sum_{t=1}^{T-1} \begin{pmatrix} \mathbf{y}_t^{(1,f)'} \\ \mathbf{Y}_{t \cdot t-1}^{(f)'} \end{pmatrix} \mathbf{M}_t(\mathbf{y}_t^{(1,f)}, \mathbf{Y}_{t \cdot t-1}^{(f)}), \tag{2.7}$$

and

$$\mathbf{H}^{(f)} = \sum_{t=1}^{T-1} \begin{pmatrix} \mathbf{y}_t^{(1,f)'} \\ \mathbf{Y}_{t \cdot t-1}^{(f)'} \end{pmatrix} [\mathbf{I}_N - \mathbf{M}_t](\mathbf{y}_t^{(1,f)}, \mathbf{Y}_{t \cdot t-1}^{(f)}). \tag{2.8}$$

Then the LIML estimator $(1, -\hat{\boldsymbol{\theta}}_{LI}') = (1, -\hat{\beta}_{2.LI}, -\hat{\gamma}_{1.LI})'$ is defined by

$$(\frac{1}{n}\mathbf{G}^{(f)} - \lambda_n \frac{1}{q_n}\mathbf{H}^{(f)}) \begin{pmatrix} 1 \\ -\hat{\boldsymbol{\theta}}_{LI} \end{pmatrix} = \mathbf{0}, \tag{2.9}$$

where $n = NT$, $q_n = n - r_n$, and λ_n is the smallest root of

$$|\frac{1}{n}\mathbf{G}^{(f)} - \lambda_n \frac{1}{q_n}\mathbf{H}^{(f)}| = 0. \tag{2.10}$$

The solution to (2.10) gives the minimum of the variance ratio

$$R_n = \frac{(1, -\boldsymbol{\theta}')\mathbf{G}^{(f)}(1, -\boldsymbol{\theta}')'}{(1, -\boldsymbol{\theta}')\mathbf{H}^{(f)}(1, -\boldsymbol{\theta}')'}. \tag{2.11}$$

We note that the LIML estimation in our formulation does not depend on the particular distribution for disturbances although the original derivation by Anderson and Rubin (1949, 1950) assumed the normal disturbances and it could be interpreted as a semi-parametric estimation method.

3 Asymptotic Properties of Estimators

In this section we shall state our main results on the asymptotic properties of estimators when both N and T go to infinity. For this purpose we first state a set of assumptions.

Assumptions :

(A1) $\{\mathbf{v}_{it}\}$ $(i = 1, ..., N; \ t = 1, ..., T)$ are i.i.d. across time and individuals and independent of η_i and $\mathbf{y}_{i0}$, with $\mathcal{E}[\mathbf{v}_{it}] = \mathbf{0}$, $\mathcal{E}[\mathbf{v}_{it}\mathbf{v}'_{it}] = \mathbf{\Omega}$, and $\mathcal{E}[\|\mathbf{v}_{it}\|^8]$ exsits.

(A2) The initial observations satisfy

$$\mathbf{y}_{i0} = (\mathbf{I}_2 - \mathbf{\Pi})^{-1}\boldsymbol{\pi}_i + \mathbf{w}_{i0} \qquad\qquad (i = 1, ..., N),$$

where $\mathbf{w}_{i0}$ is independent of η_i and i.i.d. with the steady state distribution of the homogenous process, so that we can write $\mathbf{w}_{i0} = \sum_{j=0}^{\infty} \mathbf{\Pi}^j \mathbf{v}_{i(0-j)}$.

(A3) η_i are i.i.d. across individuals with $\mathcal{E}[\eta_i] = 0$, $Var[\eta_i] = \sigma_\eta^2$, and finite fourth order moment.

(A4) The true parameters satisfy that $|\gamma_1| < 1$, $|\gamma_2| < 1$, $\gamma_2 \neq 0$, and $\gamma_1 \neq \gamma_2$.

The conditions from (A1) to (A3) and the stationarity condition of (A4) are analogue to the assumptions used in Alvarez and Alrellano (2003). They can be certainly relaxed, but with some complications of our derivations. The condition $\gamma_2 \neq 0$ is the rank condition for the identification of β_2. It is mathematically convenient to assume $\gamma_1 \neq \gamma_2$ for analyzing the dynamic process of $\mathbf{y}_{it}$. We also assume that the limit of T/N is equal to $1/2$ or less. This condition is necessary to define the GMM and LIML estimators appropriately, or insure the nonsingularity of a matrix $\mathbf{Z}'_{T-1}\mathbf{Z}_{T-1}$.

Let

$$\boldsymbol{\pi} = (\mathbf{I}_2 - \mathbf{\Pi})^{-1}\mathbf{B}^{-1}\begin{pmatrix} 1 \\ \delta \end{pmatrix} \tag{3.1}$$

and the underlying stationary process be

$$\mathbf{w}_{it} = \mathbf{y}_{it} - \boldsymbol{\pi}\eta_i . \tag{3.2}$$

Then we write the auto-covariance matrices of $\{\mathbf{w}_{it}\}$ as $\mathbf{\Gamma}_h = \mathcal{E}[\mathbf{w}_{it}\mathbf{w}'_{it+h}]$ for $h \geq 0$ under the stationary assumption, which are given by

$$\mathbf{\Gamma}_h = \sum_{i=0}^{\infty} \mathbf{\Pi}^i \mathbf{\Omega} \mathbf{\Pi}'^{i+h} \tag{3.3}$$

and $\mathbf{\Omega} = (\omega_{gh})$ $(g, h = 1, 2)$.

For the WG and the GMM estimators, we have the next result. The proof will be given in Section 6 (Sections 6.1, 6.2 and 6.4).

Theorem 1 : Let Assumptions (A1)-(A-4) hold.
(i) As $T \to \infty$, regardless of whether N is fixed or tends to infinity,

$$\hat{\boldsymbol{\theta}}_{WG} - \begin{pmatrix} \beta_2 \\ \gamma_1 \end{pmatrix} \xrightarrow{p} \left[\boldsymbol{\Phi}^* + \mathcal{E}[v_{it}^{(2)2}] \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1,0) \right]^{-1} \left[\begin{matrix} \mathcal{E}[v_{it}^{(2)} u_{it}^{(1)}] \\ 0 \end{matrix} \right]. \tag{3.4}$$

(ii) Assume $T/N \to c$ and $0 \le c \le 1/2$ as N and $T \to \infty$. Then

$$\hat{\boldsymbol{\theta}}_{GMM} - \begin{pmatrix} \beta_2 \\ \gamma_1 \end{pmatrix} \xrightarrow{p} \left[\boldsymbol{\Phi}^* + c\mathcal{E}[v_{it}^{(2)2}] \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1,0) \right]^{-1} \left[\begin{matrix} c\mathcal{E}[v_{it}^{(2)} u_{it}^{(1)}] \\ 0 \end{matrix} \right]. \tag{3.5}$$

where $\boldsymbol{\Phi}^* = \mathbf{D}' \boldsymbol{\Gamma}_0 \mathbf{D}$ and

$$\mathbf{D} = \begin{bmatrix} 0 & 1 \\ \gamma_2 & 0 \end{bmatrix}.$$

(iii) When $c = 0$, we additionally assume that $0 \le \lim_{N,T\to\infty}(T^3/N) = d_a < \infty$.
Then

$$\sqrt{NT} \left[\hat{\boldsymbol{\theta}}_{GMM} - \begin{pmatrix} \beta_2 \\ \gamma_1 \end{pmatrix} \right] \xrightarrow{d} \mathcal{N}(\mathbf{b}_0^{(a)}, \sigma^2 \boldsymbol{\Phi}^{*-1}), \tag{3.6}$$

where

$$\mathbf{b}_0^{(a)} = d_a^{1/2} \boldsymbol{\Phi}^{*-1} \left[\begin{matrix} (0,1)\boldsymbol{\Omega}\boldsymbol{\beta} \\ \mathbf{0} \end{matrix} \right]. \tag{3.7}$$

When $\mathcal{E}[v_{it}^{(2)} u_{it}^{(1)}] \neq 0$ and we have an endogenous variable, the WG estimator is generally inconsistent, even though T tends to infinity, and also the GMM estimator becomes inconsistent if $c \neq 0$.

For the LIML estimator we shall give lengthy arguments for deriving its asymptotic behaviors when both N and T go to infinity while T/N tends to a positive constant. We summarize our results whose proof will be given in Section 6.3.

Theorem 2 : Let Assumptions (A1)-(A4) hold. Assume N and $T \to \infty$ and $T/N \to c$ $(0 \le c \le 1/2)$. Then $\hat{\boldsymbol{\theta}}_{LI} \xrightarrow{p} \boldsymbol{\theta}$ and

$$\sqrt{NT} \left[\hat{\boldsymbol{\theta}}_{LI} - \begin{pmatrix} \beta_2 \\ \gamma_1 \end{pmatrix} \right] \xrightarrow{d} \mathcal{N}(\mathbf{b}, \boldsymbol{\Psi}^*), \tag{3.8}$$

where

$$\mathbf{b} \;=\; (\frac{-\sqrt{c}}{1-c})\boldsymbol{\Phi}^{*-1}\mathbf{D}'(\mathbf{I}_2 - \boldsymbol{\Pi})^{-1}\boldsymbol{\Omega}\boldsymbol{\beta}, \tag{3.9}$$

$$\boldsymbol{\Psi}^* \;=\; \boldsymbol{\Phi}^{*-1}\left[\sigma^2\boldsymbol{\Phi}^* + (c_*|\boldsymbol{\Omega}| + \Xi_4)\begin{pmatrix}1\\0\end{pmatrix}(1,0) + \boldsymbol{\Xi}_3 + \boldsymbol{\Xi}_3'\right]\boldsymbol{\Phi}^{*-1}, \tag{3.10}$$

$$\boldsymbol{\Xi}_3 \;=\; (\frac{\mathcal{E}[u_{it}^{(1)2}u_{it}^{\perp}]}{1-c}\lim_{n\to\infty}\frac{1}{n}\mathbf{D}'\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{w}_{t-1}'\mathbf{d}_t], \mathbf{0}), \tag{3.11}$$

$$\Xi_4 \;=\; \frac{\mathcal{E}[(u_{it}^{(1)2}-\sigma^2)u_{it}^{\perp 2}]}{(1-c)^2}(\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{d}_t'\mathbf{d}_t] - c^2), \tag{3.12}$$

and $\boldsymbol{\beta} = (1, -\beta_2)'$, $\sigma^2 = \mathcal{E}[u_{it}^{(1)2}]$, $c_* = c/(1-c)$, $\mathbf{d}_t$ denotes the $N \times 1$ vector containing the diagonal elements of $\mathbf{M}_t = \mathbf{Z}_t(\mathbf{Z}_t'\mathbf{Z}_t)^{-1}\mathbf{Z}_t'$ and $\mathbf{u}_{it}^{\perp}$ is defined by

$$u_{it}^{\perp} = (0,1)[\mathbf{I}_2 - \frac{\boldsymbol{\Omega}\boldsymbol{\beta}\boldsymbol{\beta}'}{\boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{\beta}}]\mathbf{v}_{it} , \tag{3.13}$$

provided that $\boldsymbol{\Xi}_3$ and Ξ_4 are well-defined.

The asymptotic covariance (3.10) of the LIML estimator has the same structure as the recent result by Anderson et al. (2008b). However, we have an extra asymptotic bias term, which depends on the limiting behavior of T/N. This is due to the effects of the forward-filtering in our formulation of dynamic panel structural equations. When the disturbances are normally distributed, the asymptotic covariance becomes

$$\boldsymbol{\Psi}^* = \boldsymbol{\Phi}^{*-1}\left[\sigma^2\boldsymbol{\Phi}^* + c_*|\boldsymbol{\Omega}|\right]\boldsymbol{\Phi}^{*-1}, \tag{3.14}$$

which is much simpler than the general case of (3.10). We expect that the additional two terms in (3.10) are often small in comparison with leading two terms. Because we have many orthogonal conditions, we have the second term of (3.10) when $c \neq 0$.

4   On Finite Sample Properties

It is important to investigate the finite sample properties of estimators partly because they are not necessarily similar to their asymptotic properties. One simple example would be the fact that the exact moments of some estimators do not necessarily exist. (In that case it may be meaningless to compare the exact MSEs of alternative estimators and their Monte Carlo analogues.) Hence we have investigated the distribution functions of several estimators in the normalized form

given by

$$\frac{\sqrt{NT}}{\sigma} \begin{bmatrix} 1/\sqrt{\phi^{11}} & 0 \\ 0 & 1/\sqrt{\phi^{22}} \end{bmatrix} \begin{bmatrix} \hat{\beta}_2 - \beta_2 \\ \hat{\gamma}_1 - \gamma_1 \end{bmatrix}, \qquad (4.15)$$

where ϕ^{11} and ϕ^{22} are the (1,1)-th element and (2.2)-th element of $\boldsymbol{\Phi}^{*-1}$, respectively. We have chosen this standardization because the limiting distribution of the LIML estimator in the form of (3.8) is $N_2(\mathbf{0}, \mathbf{I}_2)$ when $c = 0$.

We have conducted our numerical investigations in a systematic way. We first set the unknown parameters such as $(\beta_2, \gamma_1) = (.5, .5)$, $\gamma_2 = .3$. $\delta = Var[\eta_i]/\sigma^2 = 1$, $Var[v_{it}^{(g)}] = 1$ $(g = 1, 2)$ and $Cov[v_{it}^{(1)} v_{it}^{(2)}] = .3$. Then we generate large number of normal random variables by simulations and calculate the empirical distribution function in the form of (4.15). We repeat 5,000 replications for each case and the smoothing technique to estimate the empirical distribution functions. The details of simulations are similar to those explained by Anderson, Kunitomo and Matsushita (2005, 2008a). We shall report only the results for $(N, T) = (75, 25), (150, 50)$ and $(150, 50)$ as the typical cases among a large number of experiments.

When N and T are large, the WG estimator is badly biased. The GMM estimator is badly biased unless T is much smaller than N and T^3/N converges to a constant as the minimum requirement. We have confirmed these asymptotic behaviors in Figures 2,4,6 and 8. Figures 1,3,5 and 7 show the distribution function of the LIML estimator in a particular normalization. We have found that the distributions are significantly biased and also the normalization by the limiting covariance matrix is not appropriate because the circles in figures are the standard normal distribution function N(0,1) in these figures. Then we have drawn their distribution functions by first removing bias term and then using the normalized factor given by *Theorem 2*, that is,

$$\sqrt{NT} \begin{bmatrix} \psi_{11}^{-1/2} & 0 \\ 0 & \psi_{22}^{-1/2} \end{bmatrix} \begin{bmatrix} \begin{pmatrix} \hat{\beta}_2 - \beta_2 \\ \hat{\gamma}_1 - \gamma_1 \end{pmatrix} - \frac{1}{\sqrt{NT}} \mathbf{b} \end{bmatrix}, \qquad (4.16)$$

where ψ_{11} and ψ_{22} are the (1,1)-th element and the (2,2)-th element of $\boldsymbol{\Psi}^*$, respectively.
The resulting curves are called the LIML distributin with large-K asymptotics. From these figures, we have confirmed that the limiting normal distributions approximate the finite sample distribution functions of the LIML estimator quite well as *Theorem 1* and *Theorem 2* stated.

There are immediate implications. First, the GMM estimator is badly biased when T is large and it should not be used unless T is very small. (The WG estimator is badly biased even when T is small.) Second, in order to use the limiting normal distribution of the LIML estimator for statistical inferences, it is

important to adjust the asymptotic bias and the asymptotic variance formulas in *Theorem 2*. Since we have the explicit formulas for the bias and the covariance, it is straightforward to use them in practical applications.

In order to make comparison with the results reported by Alvarez and Arellano (2003) and Anderson et al. (2008a), we have conducted the similar calculations without endogeneity by setting $\beta_2 = 0$ and without dynamic effect by setting $\gamma_1 = 0$, respectively. We give the distribution function of the standardized WG, the GMM and the LIML estimators as Figures 11, 12,15, 16 for the first case. We have confirmed the numerical results of Alvarez and Arellano (2003) in the sense that it is important to adjust the bias terms in the limiting distributions of the GMM and LIML estimators. We give the distribution function of the standardized WG, the GMM and the LIML estimators as Figures 9, 10, 13, 14 for the second case. The numerical results are similar to those reported by Anderson et al. (2008a) on the TSLS, the GMM and the LIML estimators although there are some differences due to the bias terms in the panel LIML estimation.

5   Some Concluding Remarks

In this paper we have investigated the finite sample and asymptotic properties of the WG, the GMM and the LIML estimators for coefficients in a particular dynamic panel structural equation, that is, the model used by Blundell and Bond (2000) with one endogenous variable. We have investigated the conditions for the consistency and the asymptotic normality of the WG, GMM, and LIML estimators when both N and T go to infinity. We have derived the asymptotic distributions and the asymptotic bias terms of the GMM and the LIML estimators explicitly. Although we have a finite number of observations in actual applications, we have confirmed that our asymptotic results agree with their finite sample properties based on a large number of Monte Carlo experiments. When N and T are reasonably large, our results show the asymptotic robustness of the panel LIML estimation with many instruments, which agree with the recent results obtained by Anderson et al. (2008a, b) for the standard structural equation estimation. We have pointed out that it is possible to use the bias correction of the LIML estimator for general dynamic panel structural equations if necessary.

Finally, as I have mentioned in Introduction, the results reported in this paper can be generalized to more general dynamic panel structural equations with some complications. Some results on the asymptotic properties of estimators and testing procedures have been developed by Akashi and Kunitomo (2010), and Akashi (2008) in a more general framework, respectively. They have suggested that the

essential characteristics of good performance of the LIML estimation in dynamic panel structural equations reported in this paper remain the same.

6 Mathematical Details

This section gathers the mathematical derivations of our results in Section 3. The most parts of our derivations are rather straight-forward applications and some extensions of Alvarez and Allerano (2003) and Anderson et al. (2008b). For the sake of completeness, we give some details.

6.1 Within-Groups

First, we consider the dynamic process $\mathbf{w}_{it} = \mathbf{y}_{it} - \boldsymbol{\pi}\eta_i$. Under the condition $\gamma_1 \neq \gamma_2$, $\boldsymbol{\Pi}^h$ can be decomposed as

$$
\begin{aligned}
\boldsymbol{\Pi}^h &= \begin{pmatrix} 1 & \frac{\beta_2\gamma_2}{\gamma_2-\gamma_1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma_1^h & 0 \\ 0 & \gamma_2^h \end{pmatrix} \begin{pmatrix} 1 & \frac{\beta_2\gamma_2}{\gamma_2-\gamma_1} \\ 0 & 1 \end{pmatrix}^{-1} \\
&= \begin{pmatrix} \gamma_1^h & \phi_\beta(\gamma_1^h - \gamma_2^h) \\ 0 & \gamma_2^h \end{pmatrix},
\end{aligned}
\tag{6.1}
$$

where $\phi_\beta = \beta_2\gamma_2/(\gamma_1-\gamma_2)$. Then the auto-covariance matrices $\boldsymbol{\Gamma}_h = \mathcal{E}[\mathbf{w}_{it}\mathbf{w}'_{it+h}]$ ($h = 0, 1, ...$) are given by $\boldsymbol{\Gamma}_h = \sum_{i=0}^{\infty} \boldsymbol{\Pi}^i \boldsymbol{\Omega} \boldsymbol{\Pi}'^{i+h}$ and $\boldsymbol{\Omega} = \mathcal{E}[\mathbf{v}_{it}\mathbf{v}'_{it}]$. By using the relation $\boldsymbol{\Gamma}_h - \boldsymbol{\Pi}\boldsymbol{\Gamma}_h\boldsymbol{\Pi}' = \boldsymbol{\Omega}\boldsymbol{\Pi}^{h'}$, we find that $vec[\boldsymbol{\Gamma}_h] = (\mathbf{I}_4 - \boldsymbol{\Pi} \otimes \boldsymbol{\Pi})^{-1}vec[\boldsymbol{\Omega}\boldsymbol{\Pi}^{h'}]$ and $(\mathbf{I}_4 - \boldsymbol{\Pi} \otimes \boldsymbol{\Pi})$ is a block-diagonal matrix. By a direct calculation, the elements of $\boldsymbol{\Gamma}_h$ are given by

$$
\begin{aligned}
\gamma_{h(1,1)} &= \frac{1}{(1-\gamma_1^2)(1-\gamma_1\gamma_2)}[(1-\gamma_1\gamma_2)(\omega_{11}+\omega_{12}\phi_\beta)+(\gamma_1\beta_2\gamma_2)(\omega_{21}+\omega_{22}\phi_\beta)]\gamma_1^h \\
&\quad + \frac{1}{(1-\gamma_1^2)(1-\gamma_1\gamma_2)(1-\gamma_2^2)}[(\omega_{12}-\omega_{22}\phi_\beta)(1-\gamma_2^2)(\gamma_1\beta_2\gamma_2) \\
&\quad + (\beta_2^2\gamma_2^2)(1+\gamma_1\gamma_2)-(1-\gamma_1\gamma_2)(1-\gamma_2^2)(\omega_{12}\phi_\beta)]\gamma_2^h, \\
\gamma_{h(1,2)} &= \frac{1}{(1-\gamma_1\gamma_2)(1-\gamma_2^2)}[(1-\gamma_2^2)\omega_{12}+\beta_2\gamma_2^2\omega_{22}]\gamma_2^h, \\
\gamma_{h(2,1)} &= \frac{1}{(1-\gamma_1\gamma_2)}([\omega_{21}+\omega_{22}\phi_\beta]\gamma_1^h+[\frac{\beta_2^2\gamma_2^2\omega_{22}}{1-\gamma_2^2}-\phi_\beta\omega_{22}]\gamma_2^h), \\
\gamma_{h(2,2)} &= (\frac{\omega_{22}}{1-\gamma_2^2})\gamma_2^h.
\end{aligned}
$$

For the within-groups estimator, we write

$$
\mathcal{E}[\sum_{t=1}^{T-1} \mathbf{Y}_{t\cdot t-1}^{f'} \mathbf{u}_t^{(1,f)}] = \mathcal{E}[\begin{pmatrix} \sum_{i=1}^{N} \mathbf{y}_i^{(2)'} \mathbf{Q}_T \mathbf{u}_i \\ \sum_{i=1}^{N} \mathbf{y}_{i(-1)}^{(1)'} \mathbf{Q}_T \mathbf{u}_i \end{pmatrix}],
\tag{6.2}
$$

11

where $\mathbf{y}_i^{(2)'} = (y_{i1}^{(2)}, ..., y_{iT}^{(2)})$, $\mathbf{y}_{i(-1)}^{(1)'} = (y_{i0}^{(1)}, ..., y_{i(T-1)}^{(1)})$ and $\mathbf{u}_i' = (u_{i1}^{(1)}, ..., u_{iT}^{(1)})$. Then

$$\mathcal{E}[\frac{1}{N}\sum_{i=1}^{N}\mathbf{y}_{i(-1)}^{(1)'}\mathbf{Q}_T\,\mathbf{u}_i] = \mathcal{E}[\mathbf{y}_{i(-1)}^{(1)'}\,\mathbf{u}_i] - \frac{1}{T}\boldsymbol{\iota}_T'\mathcal{E}[\mathbf{y}_{i(-1)}^{(1)}\mathbf{u}_i']\boldsymbol{\iota}_T\,, \tag{6.3}$$

where the (t,s) elements of $\mathcal{E}[\mathbf{y}_{i(-1)}^{(1)}\mathbf{u}_i']$ are $\gamma_1^t\mathcal{E}[v_{it}^{(1)}u_{it}^{(1)}] + \phi_\beta(\gamma_1^t - \gamma_2^t)\mathcal{E}[v_{it}^{(2)}u_{it}^{(1)}]$ if $t > s$ and 0 otherwise.

By evaluating each elements, we obtain

$$\mathcal{E}[\sum_{i=1}^{N}\mathbf{y}_{i(-1)}^{(1)'}\mathbf{Q}_T\,\mathbf{u}_i] = -\frac{N}{T}[(\mathcal{E}[v_{it}^{(1)}u_{it}^{(1)}] + \phi_\beta\mathcal{E}[v_{it}^{(2)}u_{it}^{(1)}])(\frac{1}{1-\gamma_1})(T - \frac{1-\gamma_1^T}{1-\gamma_1})$$

$$- \phi_\beta\mathcal{E}[v_{it}^{(2)}u_{it}^{(1)}](\frac{1}{1-\gamma_2})(T - \frac{1-\gamma_2^T}{1-\gamma_2})], \tag{6.4}$$

$$\mathcal{E}[\sum_{i=1}^{N}\mathbf{y}_i^{(2)'}\mathbf{Q}_T\,\mathbf{u}_i] = N\mathcal{E}[v_{it}^{(2)}u_{it}^{(1)}](T - \frac{1}{1-\gamma_2}[1 - (\frac{\gamma_2}{T})(\frac{1-\gamma_2^T}{1-\gamma_2})]). \tag{6.5}$$

Next, we need to evaluate the variances and we write

$$Var[\frac{1}{\sqrt{NT}}\begin{pmatrix}\sum_{i=1}^{N}\mathbf{y}_i^{(2)'}\mathbf{Q}_T\,\mathbf{u}_i \\ \sum_{i=1}^{N}\mathbf{y}_{i(-1)}^{(1)'}\mathbf{Q}_T\,\mathbf{u}_i\end{pmatrix}]$$

$$= Var[\frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T}\begin{pmatrix}w_{it}^{(2)} \\ w_{it-1}^{(1)}\end{pmatrix}u_{it}^{(1)} - \sqrt{\frac{T}{N}}\sum_{i=1}^{N}\begin{pmatrix}\bar{w}_i^{(2)} \\ \bar{w}_{i(-1)}^{(1)}\end{pmatrix}\bar{u}_i]\,, \tag{6.6}$$

where

$$\bar{w}_i^{(2)} = \frac{1}{T}\sum_{t=1}^{T}w_{it}^{(2)},\ \ \bar{w}_{i(-1)}^{(1)} = \frac{1}{T}\sum_{t=0}^{T-1}w_{it}^{(1)},\ \ \bar{u}_i = \frac{1}{T}\sum_{t=1}^{T}u_{it}^{(1)}. \tag{6.7}$$

In order to evaluate the first element of the leading term in (6.6), we use the relation that $Var[\sum_t w_{it}^{(2)}u_{it}^{(1)}] = \mathcal{E}[(\sum_t w_{it}^{(2)}u_{it}^{(1)})^2] - (T\mathcal{E}[v_{it}^{(2)}u_{it}^{(1)}])^2$. Then for the first term of this relation

$$\mathcal{E}[w_{it}^{(2)2}u_{it}^{(1)2}] = \mathcal{E}[u_{it}^{(1)2}v_{it}^{(2)2}] + \mathcal{E}[u_{it}^{(1)2}]\mathcal{E}[v_{it}^{(2)2}]\frac{\gamma_2^2}{(1-\gamma_2^2)},\ (t=s) \tag{6.8}$$

$$\mathcal{E}[w_{is}^{(2)}u_{is}^{(1)}w_{it}^{(2)}u_{it}^{(1)}] = \mathcal{E}[w_{i(s\wedge t)}^{(2)}u_{i(s\wedge t)}^{(1)}]\mathcal{E}[v_{i(s\vee t)}^{(2)}u_{i(s\vee t)}^{(1)}] \tag{6.9}$$

$$= (\mathcal{E}[v_{it}^{(2)}u_{it}^{(1)}])^2,\ (t \neq s).$$

For the second element of the leading term of (6.6), we use $\mathcal{E}[u_{it-j}^{(1)}w_{it-1-j}^{(1)}] = 0$ for $j \neq 0$ and

$$\mathcal{E}[(\sum_{t=1}^{T}u_{it}^{(1)}w_{it-1}^{(1)})^2] = T\mathcal{E}[u_{it}^{(1)2}w_{it-1}^{(1)2}] = T\mathcal{E}[u_{it}^{(1)2}]\mathcal{E}[w_{it-1}^{(1)2}]. \tag{6.10}$$

For the second term of (6.6), we have

$$Var[\sqrt{\frac{T}{N}} \sum_{i=1}^{N} \begin{pmatrix} \bar{w}_i^{(2)} \\ \bar{w}_{i(-1)}^{(1)} \end{pmatrix} \bar{u}_i] = TVar[\begin{pmatrix} \bar{w}_i^{(2)} \\ \bar{w}_{i(-1)}^{(1)} \end{pmatrix} \bar{u}_i] = O(\frac{1}{T}), \qquad (6.11)$$

since $Var[(.)\bar{u}_i^{(1)}] = O(1/T^2)$ (See Page 1139 of Alvaretz and Arellano (2003)). Then

$$Var[\frac{1}{\sqrt{NT}} \begin{pmatrix} \sum_{i=1}^{N} \mathbf{y}_i^{(2)'} \mathbf{Q}_T \mathbf{u}_i \\ \sum_{i=1}^{N} \mathbf{y}_{i(-1)}^{(1)'} \mathbf{Q}_T \mathbf{u}_i \end{pmatrix}] \qquad (6.12)$$

$$= \begin{pmatrix} \mathcal{E}[u_{it}^{(1)2} v_{it}^{(2)2}] + \mathcal{E}[u_{it}^{(1)2}]\mathcal{E}[v_{it}^{(2)2}](\frac{\gamma_2^2}{1-\gamma_2^2}) - (\mathcal{E}[u_{it}^{(1)} v_{it}^{(2)}])^2 \\ \mathcal{E}[u_{it}^{(1)2}]\Gamma_{0(1,1)} \end{pmatrix} + O(\frac{1}{T}).$$

Third, we need to evaluate the expectation of quadratic forms

$$\mathcal{E}[\frac{1}{NT} \sum_{t=1}^{T-1} \mathbf{Y}_{t\cdot t-1}^{f'} \mathbf{Y}_{t\cdot t-1}^{f}]$$

$$= \begin{pmatrix} \gamma_{0(2,2)} & \gamma_{1(1,2)} \\ \gamma_{1(1,2)} & \gamma_{0(1,1)} \end{pmatrix} + \mathcal{E}[\begin{pmatrix} \bar{w}_i^{(2)2} & \bar{w}_i^{(2)}\bar{w}_{i(-1)}^{(1)} \\ \bar{w}_{i(-1)}^{(1)}\bar{w}_i^{(2)} & \bar{w}_{i(-1)}^{(1)2} \end{pmatrix}], \qquad (6.13)$$

where

$$\mathcal{E}[\bar{w}_{i(-1)}^{(1)2}] = \frac{1}{T^2}[\sum_{j=0}^{T-1}\sum_{h=0}^{j}\Gamma_h + \sum_{j=1}^{T-1}\sum_{h=1}^{j}\Gamma_h']_{(2,2)}. \qquad (6.14)$$

The first double sum $\mathbf{S}_{0T} = \sum_{j=0}^{T-1}\sum_{h=0}^{j}\Gamma_0\Pi'^h$ can be represented as

$$\begin{aligned} \mathbf{S}_{0T} &= \Gamma_0\mathbf{S}_T \qquad (6.15) \\ &= \Gamma_0[\Pi'(\Pi^{T'} - \mathbf{I}_2)(\Pi' - \mathbf{I}_2)^{-1} - T\mathbf{I}_2](\Pi' - \mathbf{I}_2)^{-1}, \end{aligned}$$

and the elements of $\mathbf{S}_T$ are given by

$$\begin{aligned} \mathbf{S}_{T(g,g)} &= \frac{[\pi_{ii}(1 - \pi_{ii}^T)(1 - \pi_{ii}) - T](\pi_{ii} - 1)}{(1 - \pi_{11})^2(1 - \pi_{22})^2} \quad (g = 1, 2) \\ \mathbf{S}_{T(2,1)} &= (\frac{1}{(1 - \pi_{11})^2(1 - \pi_{22})^2})[([\pi_{22}(\pi_{11}^T - 1) + \pi_{12}\phi_\beta(\pi_{11}^T - \pi_{22}^T)](\pi_{11} - 1) \\ &\quad - \pi_{22}\pi_{12}(\pi_{11}^T - 1))(\pi_{11} - 1) - (\pi_{22}(1 - \pi_{22}^T)(1 - \pi_{22}) - T)\pi_{12}], \\ \mathbf{S}_{T(1,2)} &= 0, \end{aligned}$$

where $\Pi = \{\pi_{gh}\}$ $(g, h = 1, 2)$. Also define $\mathbf{S}_{1T} = \Gamma_1\mathbf{S}_T$, then we use the relation

$$\mathcal{E}[\bar{w}_{i(-1)}^{(1)2}] = [(\mathbf{S}_{0T} + \mathbf{S}_{1(T-1)}')/T^2]_{(2,2)}, \quad \mathcal{E}[\bar{w}_i^{(2)2}] = [(\mathbf{S}_{0T} + \mathbf{S}_{1(T-1)}')/T^2]_{(1,1)}. \quad (6.16)$$

13

Similarly, we can show that $\mathcal{E}[\bar{w}_i^{(2)}\bar{w}_{i(-1)}^{(1)}] = [(\mathbf{S}_{1T} + \mathbf{S}'_{0(T-1)})/T^2]_{(1,2)}$. Hence we find that

$$\mathcal{E}[\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t\cdot t-1}^{f'}\mathbf{Y}_{t\cdot t-1}^{f}] = \begin{pmatrix} \gamma_{0(2,2)} & \gamma_{1(1,2)} \\ \gamma_{1(1,2)} & \gamma_{0(1,1)} \end{pmatrix} + O(\frac{1}{T}) \, . \tag{6.17}$$

Moreover, we shall evaluate the covariance matrix

$$\begin{aligned} Var[\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{Y}_{t\cdot t-1}^{f'}\mathbf{Y}_{t\cdot t-1}^{f}] &= \frac{1}{N}Var[\frac{1}{T}\sum_{t}\begin{pmatrix} w_{it}^{(2)2} & w_i^{(2)}w_{i(t-1)}^{(1)} \\ w_{i(t-1)}^{(1)}w_i^{(2)} & w_{i(t-1)}^{(1)2} \end{pmatrix} \\ &\quad - \begin{pmatrix} \bar{w}_i^{(2)2} & \bar{w}_i^{(2)}\bar{w}_{i(-1)}^{(1)} \\ \bar{w}_{i(-1)}^{(1)}\bar{w}_i^{(2)} & \bar{w}_{i(-1)}^{(1)2} \end{pmatrix}] \\ &= o(1) \, , \end{aligned}$$

where the second terms are $O(N^{-1}T^{-2})$ by using the same arguments as for (6.11). In order to show that the first terms are $O(N^{-1}T^{-1})$, we decompose $w_{it}^{(1)}$ as the sum of the two AR(1) processes, or $w_{it}^{(11)} + w_{it}^{(12)}$, those coefficients are γ_1 and γ_2, respectively. Then

$$Var[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}w_{i(t-1)}^{(2)2}] \rightarrow (\frac{1+\gamma_1^2}{1-\gamma_1^2})Var[w_{it}^{(2)2}], \tag{6.18}$$

$$Var[\frac{1}{\sqrt{T}}\sum_{t=1}^{T}w_{it}^{(1)2}] \rightarrow (\frac{1+\gamma_1^2}{1-\gamma_1^2})Cov[w_{it}^{(1)2},w_{it}^{(11)2}] + 2(\frac{1+\gamma_1^2\gamma_2^2}{1-\gamma_1^2\gamma_2^2}) \tag{6.19}$$

$$\times Cov[w_{it}^{(1)2},w_{it}^{(11)}w_{it}^{(12)}] + (\frac{1+\gamma_2^2}{1-\gamma_2^2})Cov[w_{it}^{(1)2},w_{it}^{(12)2}],$$

and $Var[(1/\sqrt{T})\sum_{t=1}^{T}w_{it}^{(2)}w_{i(t-1)}^{(1)}] \rightarrow O(1)$. Also we have

$$Var[\frac{1}{T}\sum_{t=1}^{T}w_{i(t-1)}^{(2)}w_{i(t-1)}^{(1)}] \rightarrow 0, \tag{6.20}$$

which will be used later. It is because $Var[(1/\sqrt{T})\sum_{t=1}^{T}w_{it}^{(2)}w_{it}^{(1)}] = Var[(1/\sqrt{T})\sum_{t=1}^{T}w_{it}^{(2)}(\beta_2 w_{it}^{(2)} + \gamma_1 w_{i(t-1)}^{(1)} + u_{it}^{(1)})]$ and that the right-hand side's variance terms are $O(1)$ as $T \rightarrow \infty$.

Proof of (3.4) : We have shown the suffcient condition fors mean-square convergence to the limit of (6.2) and (6.13), and therefore the convergence in probability follows. **Q.E.D.**

6.2   GMM (Derivation of (3.4))

We prepare two lemmas for our derivation. The first one is a direct application of Lemma C1 of Alvaretz and Arellano (2003).

Lemma 1 : Let $\mathbf{d}_t$ and $\mathbf{d}_s$ be $N \times 1$ vectors containing the diagonal elements of $\mathbf{M}_t$ and $\mathbf{M}_s$, respectively, such that $tr(\mathbf{M}_t) = \mathbf{d}_t' \boldsymbol{\iota}_N = 2t$, $tr(\mathbf{M}_s) = \mathbf{d}_s' \boldsymbol{\iota}_N = 2s$ and $\mathbf{d}_t' \mathbf{d}_s \leq 2 \max(t, s)$. Then, under (A1) and (A3), for $l \geq r \geq t$, $p \geq q \geq s$, $t \geq s$,

$$Cov[\boldsymbol{\epsilon}_l^{(a)'} \mathbf{M}_t \boldsymbol{\epsilon}_r^{(b)}, \boldsymbol{\epsilon}_p^{(a)'} \mathbf{M}_s \boldsymbol{\epsilon}_q^{(b)}] \qquad (6.21)$$

$$= \begin{cases} (m^{(3)} + m^{(2)})(2s) + m^{(0)} \mathcal{E}[\mathbf{d}_t' \mathbf{d}_s] & \text{if } l = r = p = q, \\ \mathcal{E}[\epsilon_{it}^{(a)2} \epsilon_{it}^{(b)}] \mathcal{E}[\mathbf{d}_t' \mathbf{M}_s \boldsymbol{\epsilon}_q^{(b)}] & \text{if } l = r = p \neq q < t, \\ m^{(3)}(2s) & \text{if } l = p \neq r = q, \\ 0 & \text{otherwise,} \end{cases}$$

where $|\mathcal{E}[\mathbf{d}_t' \mathbf{M}_s \boldsymbol{\epsilon}_q^{(b)}]| \leq 2(st \mathcal{E}[\epsilon_{it}^{(b)2}])^{1/2}$,

$$m^{(1)} = m^{(1)}(\boldsymbol{\epsilon}_t^{(a)}, \boldsymbol{\epsilon}_t^{(b)}) = \mathcal{E}[\epsilon_{it}^{(a)2} \epsilon_{it}^{(b)2}], \qquad (6.22)$$
$$m^{(2)} = m^{(2)}(\boldsymbol{\epsilon}_t^{(a)}, \boldsymbol{\epsilon}_t^{(b)}) = (\mathcal{E}[\epsilon_{it}^{(a)} \epsilon_{it}^{(b)}])^2,$$
$$m^{(3)} = m^{(3)}(\boldsymbol{\epsilon}_t^{(a)}, \boldsymbol{\epsilon}_t^{(b)}) = \mathcal{E}[\epsilon_{it}^{(a)2}] \mathcal{E}[\epsilon_{it}^{(b)2}],$$
$$m^{(0)} = m^{(0)}(\boldsymbol{\epsilon}_t^{(a)}, \boldsymbol{\epsilon}_t^{(b)}) = m^{(1)} - 2m^{(2)} - m^{(3)},$$

and $(\boldsymbol{\epsilon}_t^{(a)}, \boldsymbol{\epsilon}_t^{(b)})$ takes any pair of $N \times 1$ vectors from random variables $(v_{it}^{(g)})$ $(g = 1, 2)$ such as $(\boldsymbol{v}_t^{(2)}, \boldsymbol{u}_t)$.

Proof: We shall show that

$$Cov_t[\boldsymbol{\epsilon}_l^{(a)'} \mathbf{M}_t \boldsymbol{\epsilon}_r^{(b)}, \boldsymbol{\epsilon}_p^{(a)'} \mathbf{M}_s \boldsymbol{\epsilon}_q^{(b)}] \qquad (6.23)$$

$$= \begin{cases} (m^{(3)} + m^{(2)})(2s) + m^{(0)} \mathbf{d}_t' \mathbf{d}_s & \text{if } l = r = p = q, \\ \mathcal{E}_t[\epsilon_{it}^{(a)2} \epsilon_{it}^{(b)}] \mathbf{d}_t' \mathbf{M}_s \boldsymbol{\epsilon}_q^{(b)} & \text{if } l = r = p \neq q < t, \\ m^{(3)}(2s) & \text{if } l = p \neq r = q, \\ 0 & \text{otherwise,} \end{cases}$$

where $\mathcal{E}_t[.]$ denotes an expectation conditional on η_i and $\{\epsilon_{i(t-j)}^{(a)}, \epsilon_{i(t-j)}^{(b)}\}_{j=1}^{\infty}$. We shall use

$$Cov_t[\boldsymbol{\epsilon}_l^{(a)'} \mathbf{M}_t \boldsymbol{\epsilon}_r^{(b)}, \boldsymbol{\epsilon}_p^{(a)'} \mathbf{M}_s \boldsymbol{\epsilon}_q^{(b)}]$$
$$= \mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)'} \mathbf{M}_t \boldsymbol{\epsilon}_r^{(b)} \boldsymbol{\epsilon}_p^{(a)'} \mathbf{M}_s \boldsymbol{\epsilon}_q^{(b)}] - \mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)'} \mathbf{M}_t \boldsymbol{\epsilon}_r^{(b)}] \mathcal{E}_t[\boldsymbol{\epsilon}_p^{(a)'} \mathbf{M}_s \boldsymbol{\epsilon}_q^{(b)}] . \qquad (6.24)$$

If $p < t$, then $\boldsymbol{\epsilon}_p^{(a)'} \mathbf{M}_s \boldsymbol{\epsilon}_q^{(b)}$ is constant and the covariance vanishes. The conditional means are given by

$$\mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)'} \mathbf{M}_t \boldsymbol{\epsilon}_r^{(b)}] = tr(\mathbf{M}_t \mathcal{E}_t[\boldsymbol{\epsilon}_l^{(b)} \boldsymbol{\epsilon}_r^{(a)'}]) \begin{cases} \mathcal{E}[\epsilon_{it}^{(a)} \epsilon_{it}^{(b)}] tr(\mathbf{M}_t) & \text{if } l = r, \\ 0 & \text{if } l \neq r, \end{cases} \qquad (6.25)$$

15

$$\mathcal{E}_t[\boldsymbol{\epsilon}_q^{(a)'}\mathbf{M}_s\boldsymbol{\epsilon}_p^{(b)}] = tr(\mathbf{M}_s\mathcal{E}_t[\boldsymbol{\epsilon}_q^{(b)}\boldsymbol{\epsilon}_p^{(a)'}]) \begin{cases} \mathcal{E}[\epsilon_{it}^{(a)}\epsilon_{it}^{(b)}]tr(\mathbf{M}_s) & \text{if } q = p, \\ 0 & \text{if } q \neq p. \end{cases} \quad (6.26)$$

As for the leading term we have

$$\mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)'}\mathbf{M}_t\boldsymbol{\epsilon}_r^{(b)}\boldsymbol{\epsilon}_p^{(a)'}\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}] \quad (6.27)$$

$$= \begin{cases} \mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)'}\mathbf{M}_t\boldsymbol{\epsilon}_r^{(b)}\boldsymbol{\epsilon}_p^{(a)'}\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}] & \text{if } l = r = p = q, \\ \mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)'}\mathbf{M}_t\boldsymbol{\epsilon}_l^{(b)}\boldsymbol{\epsilon}_l^{(a)'}]\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)} & \text{if } l = r = p \neq q < t, \\ tr(\mathbf{M}_t\mathcal{E}_t[\boldsymbol{\epsilon}_r^{(b)}\boldsymbol{\epsilon}_r^{(b)'}]\mathbf{M}_s\mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)}\boldsymbol{\epsilon}_l^{(a)'}]) & \text{if } l = p \neq r = q, \\ 0 & \text{otherwise.} \end{cases}$$

Thus there is only a nonzero mean-product subtraction in covariances with $l = r = p = q$. For the first type of nonzero terms

$$\mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)'}\mathbf{M}_t\boldsymbol{\epsilon}_l^{(b)}\boldsymbol{\epsilon}_l^{(a)'}\mathbf{M}_s\boldsymbol{\epsilon}_l^{(b)}] \quad (6.28)$$

$$= \sum_i\sum_j\sum_k\sum_\ell m_{ij}^{(t)}m_{k\ell}^{(s)}\mathcal{E}_t[\epsilon_{il}^{(a)}\epsilon_{jl}^{(b)}\epsilon_{kl}^{(a)}\epsilon_{\ell l}^{(b)}]$$

$$= \mathcal{E}[\epsilon_{it}^{(a)2}\epsilon_{it}^{(b)2}]\mathbf{d}_t'\mathbf{d}_s + \sum_i\sum_{k\neq i}m_{ii}^{(t)}m_{kk}^{(s)}(\mathcal{E}[\epsilon_{it}^{(a)}\epsilon_{it}^{(b)}])^2$$

$$+ \sum_i\sum_{j\neq i}m_{ij}^{(t)}m_{ij}^{(s)}\mathcal{E}[\epsilon_{it}^{(a)2}]\mathcal{E}[\epsilon_{it}^{(b)2}] + \sum_i\sum_{j\neq i}m_{ij}^{(t)}m_{ji}^{(s)}(\mathcal{E}[\epsilon_{it}^{(a)}\epsilon_{it}^{(b)}])^2$$

$$= m^{(0)}\mathbf{d}_t'\mathbf{d}_s + m^{(2)}tr(\mathbf{M}_t)tr(\mathbf{M}_s) + (m^{(3)} + m^{(2)})tr(\mathbf{M}_t\mathbf{M}_s),$$

where $m_{ij}^{(t)}$ and $m_{k\ell}^{(s)}$ denote the elements of $\mathbf{M}_t$ and $\mathbf{M}_s$, respectively. Then by using (6.25), the result follows.

For the second type, we have

$$\mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)'}\mathbf{M}_t\boldsymbol{\epsilon}_l^{(b)}\boldsymbol{\epsilon}_l^{(a)'}]\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)} = \mathcal{E}[\epsilon_{it}^{(a)2}\epsilon_{it}^{(b)}]\mathbf{d}_t'\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}, \quad (6.29)$$

and

$$tr(\mathbf{M}_t\mathcal{E}_t[\boldsymbol{\epsilon}_r^{(b)}\boldsymbol{\epsilon}_r^{(b)'}]\mathbf{M}_s\mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)}\boldsymbol{\epsilon}_l^{(a)'}]) = \mathcal{E}[\epsilon_{it}^{(a)2}]\mathcal{E}[\epsilon_{it}^{(b)2}]tr(\mathbf{M}_t\mathbf{M}_s). \quad (6.30)$$

Given (6.23), we find the unconditional covariance given by

$$Cov[\boldsymbol{\epsilon}_l^{(a)'}\mathbf{M}_t\boldsymbol{\epsilon}_r^{(b)}, \boldsymbol{\epsilon}_p^{(a)'}\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}] \quad (6.31)$$

$$= \mathcal{E}[Cov_t[\boldsymbol{\epsilon}_l^{(a)'}\mathbf{M}_t\boldsymbol{\epsilon}_r^{(b)}, \boldsymbol{\epsilon}_p^{(a)'}\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}]] + Cov[\mathcal{E}_t[\boldsymbol{\epsilon}_l^{(a)'}\mathbf{M}_t\boldsymbol{\epsilon}_r^{(b)}], \mathcal{E}_t[\boldsymbol{\epsilon}_p^{(a)'}\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}]],$$

but the second term vanishes. To prove $(\mathcal{E}[\mathbf{d}_t'\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}])^2 \leq 4ts\mathcal{E}[\epsilon_{it}^{(b)2}]$, we use the inequalities

$$(\mathbf{d}_t'\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)})^2 \leq (\mathbf{d}_t'\mathbf{M}_s\mathbf{d}_t)(\boldsymbol{\epsilon}_q^{(b)'}\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}) \leq \mathbf{d}_t'\mathbf{d}_t(\boldsymbol{\epsilon}_q^{(b)'}\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}) \leq 2t(\boldsymbol{\epsilon}_q^{(b)'}\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}),$$

and hence

$$\mathcal{E}[(\mathbf{d}_t'\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)})^2] \leq 2t\mathcal{E}[\boldsymbol{\epsilon}_q^{(b)'}\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}] = 4ts\mathcal{E}[\epsilon_{it}^{(b)2}]. \tag{6.32}$$

Since $\mathcal{E}[(\mathbf{d}_t'\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)})^2] = Var[\mathbf{d}_t'\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}] + (\mathcal{E}[\mathbf{d}_t'\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}])^2$, we have $(\mathcal{E}[\mathbf{d}_t'\mathbf{M}_s\boldsymbol{\epsilon}_q^{(b)}])^2 \leq 4ts\mathcal{E}[\epsilon_{it}^{(b)2}]$. **Q.E.D.**

Lemma 2 : Under Assumptions (A1)-(A4), assume both N and $T \to \infty$ while $T/N \to c$ $(0 \leq c \leq 1/2)$. Then

$$\frac{1}{NT}\sum_{t=1}^{T-1}\begin{pmatrix}\mathbf{w}_t^{(2)'}\\\mathbf{w}_{t-1}^{(1)'}\end{pmatrix}\mathbf{M}_t(\mathbf{w}_t^{(2)},\mathbf{w}_{t-1}^{(1)}) \overset{m.s.}{\to} \boldsymbol{\Phi}^* + c\mathcal{E}[v_{it}^{(2)2}]\begin{pmatrix}1\\0\end{pmatrix}(1,0), \tag{6.33}$$

where

$$\boldsymbol{\Phi}^* = \mathbf{D}'\boldsymbol{\Gamma}_0\mathbf{D}, \quad \mathbf{D} = \begin{pmatrix}0 & 1\\\gamma_2 & 0\end{pmatrix}. \tag{6.34}$$

Proof : Let $\mathbf{M}_\mu' = (\boldsymbol{\mu}_1,...,\boldsymbol{\mu}_N)$, $\boldsymbol{\mu}_i = \eta_i\boldsymbol{\pi}$ and we construct the $N \times 2$ error matrix of the linear projection of $\mathbf{M}_\mu$ on $\mathbf{Z}_t^*$,

$$(\boldsymbol{\mu}_t^{*(1)}, \boldsymbol{\mu}_t^{*(2)}) = \mathbf{M}_\mu - \mathbf{Z}_t^*(\boldsymbol{\gamma}_t^{(1)},\boldsymbol{\gamma}_t^{(2)}), \tag{6.35}$$

where $\boldsymbol{\gamma}_t^{(g)} = (\mathcal{E}[\mathbf{z}_{it}^*\mathbf{z}_{it}^{*'}])^{-1}\mathcal{E}[\mathbf{z}_{it}^*\boldsymbol{\mu}_{i(g,1)}]$ $(g = 1,2)$ are $(2t) \times 1$ vectors and $\mathbf{Z}_t^* = \mathbf{Z}_t\boldsymbol{\Gamma}_{0t}$. We take $\boldsymbol{\Gamma}_{0t}$ as the $2t \times 2t$ block-diagonal matrix whose 2×2 diagonal blocks are lower triangular matrix $\mathbf{L}^{-1'}$ such that $\boldsymbol{\Gamma}_0 = \mathbf{LL}'$. Let also $\mathbf{V}_t$ be the $2t \times 2t$ partitioned symmetric matrix as

$$\mathbf{V}_t = \begin{pmatrix}\mathbf{I}_2 & & & &\\\boldsymbol{\Phi} & \mathbf{I}_2 & & &\\\vdots & \vdots & \ddots & &\\\boldsymbol{\Phi}^{t-2} & \boldsymbol{\Phi}^{t-3} & \dots & \mathbf{I}_2 &\\\boldsymbol{\Phi}^{t-1} & \boldsymbol{\Phi}^{t-2} & \dots & \boldsymbol{\Phi} & \mathbf{I}_2\end{pmatrix}, \tag{6.36}$$

where $\boldsymbol{\Phi} = \mathbf{L}^{-1}\boldsymbol{\Pi}\mathbf{L}$ and $(2t) \times 1$ vector $\mathbf{l}_t = (\boldsymbol{\pi}'\mathbf{L}^{-1'}, ..., \boldsymbol{\pi}'\mathbf{L}^{-1'})'$. Then we have

$$\begin{aligned}\boldsymbol{\gamma}_t^{(g)} &= (\sigma_\eta^2\mathbf{l}_t\mathbf{l}_t' + \mathbf{V}_t)^{-1}(\boldsymbol{\pi}_{(g,1)}\sigma_\eta^2\mathbf{l}_t) \tag{6.37}\\[2mm]&= \frac{\boldsymbol{\pi}_{(g,1)}\sigma_\eta^2}{1 + \sigma_\eta^2\mathbf{l}_t'\mathbf{V}_t^{-1}\mathbf{l}_t}\mathbf{V}_t^{-1}\mathbf{l}_t \ (g = 1,2),\end{aligned}$$

17

and the inverse matrix of $\mathbf{V}_t$ is a block-tridiagonal symmetric matrix such that

$$\mathbf{V}_t^{-1} = \begin{pmatrix} \mathbf{P}^{-1} & & & & \\ -\mathbf{\Phi}\mathbf{P}^{-1} & \mathbf{P}^{-1} + \mathbf{\Phi}\mathbf{P}^{-1}\mathbf{\Phi}' & & & \\ \vdots & \vdots & \ddots & & \\ \mathbf{O}_2 & \mathbf{O}_2 & \cdots & \mathbf{P}^{-1} + \mathbf{\Phi}\mathbf{P}^{-1}\mathbf{\Phi}' & \\ \mathbf{O}_2 & \mathbf{O}_2 & \cdots & -\mathbf{\Phi}\mathbf{P}^{-1} & \mathbf{I}_2 + \mathbf{\Phi}\mathbf{P}^{-1}\mathbf{\Phi}' \end{pmatrix} \tag{6.38}$$

where $\mathbf{P} = \mathbf{I}_2 - \mathbf{\Phi}'\mathbf{\Phi}$. Hence, for $g = 1, 2$, we find that

$$\begin{aligned}
\mu_{it}^{*(g)} &= \mu_i^{(g)} - \mathbf{z}_{it}^{*'}\boldsymbol{\gamma}_t^{(g)} \tag{6.39} \\
&= \mu_i^{(g)} - \eta_i \mathbf{l}_t'\boldsymbol{\gamma}_t - \left[(\mathbf{L}^{-1}\mathbf{w}_{i0})', ..., (\mathbf{L}^{-1}\mathbf{w}_{it-1})'\right]\boldsymbol{\gamma}_t \\
&= \left[\frac{\boldsymbol{\pi}_{(g,1)}}{1 + \sigma_\eta^2 \mathbf{l}_t'\mathbf{V}_t^{-1}\mathbf{l}_t}\right]\left[\eta_i - \sigma_\eta^2 \mathbf{w}_{it}^{*'}\mathbf{V}_t^{-1}\mathbf{l}_t\right] \\
&= \boldsymbol{\pi}_{(g,1)}\mu_{it}^* \ (, say), \\
\mathcal{E}[\mu_{it}^{*(g)2}] &= \frac{(\boldsymbol{\pi}_{(g,1)})^2 \sigma_\eta^2}{1 + \sigma_\eta^2 \mathbf{l}_t'\mathbf{V}_t^{-1}\mathbf{l}_t} , \tag{6.40}
\end{aligned}$$

where $\mathbf{w}_t^{*'} = (\mathbf{w}_0', \cdots, \mathbf{w}_{t-1}')$.

Since $\boldsymbol{\pi}'\mathbf{L}^{-1'}(\mathbf{I}_2 - \mathbf{\Phi})\mathbf{P}^{-1/2} \neq \mathbf{0}'$ and

$$\begin{aligned}
\lim_{t\to\infty} \frac{\mathbf{l}_t'\mathbf{V}_t^{-1}\mathbf{l}_t}{t} &= \boldsymbol{\pi}'\mathbf{L}^{-1'}\left[\mathbf{P}^{-1} + \mathbf{\Phi}\mathbf{P}^{-1}\mathbf{\Phi}' - \mathbf{P}^{-1}\mathbf{\Phi}' - \mathbf{\Phi}\mathbf{P}^{-1}\right]\mathbf{L}^{-1}\boldsymbol{\pi} \tag{6.41} \\
&= \left[\mathbf{P}^{-1/2'}(\mathbf{I}_2 - \mathbf{\Phi}')\mathbf{L}^{-1}\boldsymbol{\pi}\right]'\left[\mathbf{P}^{-1/2'}(\mathbf{I}_2 - \mathbf{\Phi}')\mathbf{L}^{-1}\boldsymbol{\pi}\right] \\
&= O(1) ,
\end{aligned}$$

we have established that $\mathbf{l}_t'\mathbf{V}_t^{-1}\mathbf{l}_t = O(t)$ and $\mathcal{E}[\mu_{it}^{*(g)2}] = O(t^{-1})$.

Furthermore, we evaluate the fourth-order moments of $\mu_{it}^{*(g)}$ $(g = 1, 2)$. Because of the form (6.39), we have $\mathcal{E}[\mu_{it}^{*(g)4}] = O(t^{-4}) \times O(t^2) = O(t^{-2})$ $(g = 1, 2)$.

Next, we shall consider the decomposition

$$\begin{aligned}
\mathbf{w}_{t-1}^{(g)'}\mathbf{M}_t\mathbf{w}_{t-1}^{(h)} &= \mathbf{w}_{t-1}^{(g)'}\mathbf{w}_{t-1}^{(h)} - \mathbf{w}_{t-1}^{(g)'}(\mathbf{I}_N - \mathbf{M}_t)\mathbf{w}_{t-1}^{(h)} \tag{6.42} \\
&= \mathbf{w}_{t-1}^{(g)'}\mathbf{w}_{t-1}^{(h)} - \boldsymbol{\mu}_t^{*(g)'}(\mathbf{I}_N - \mathbf{M}_t)\boldsymbol{\mu}_t^{*(h)} \ (g, h = 1, 2),
\end{aligned}$$

where the second equality follows from $\mathbf{w}_{t-1} = \mathbf{y}_{t-1} - \mathbf{Z}_t^*(\boldsymbol{\gamma}_t^{(1)}, \boldsymbol{\gamma}_t^{(2)}) - (\boldsymbol{\mu}_t^{*(1)}, \boldsymbol{\mu}_t^{*(2)})$, $\mathbf{M}_t = \mathbf{Z}_t^*(\mathbf{Z}_t^{*'}\mathbf{Z}_t^*)^{-1}\mathbf{Z}_t^{*'}$ and $(\mathbf{I}_N - \mathbf{M}_t)[\mathbf{y}_{t-1} - \mathbf{Z}_t^*(\boldsymbol{\gamma}_t^{(1)}, \boldsymbol{\gamma}_t^{(2)})] = \mathbf{0}$. Therefore, for $g, h = 1, 2$,

$$\frac{1}{NT}\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{w}_{t-1}^{(g)'}\mathbf{M}_t\mathbf{w}_{t-1}^{(h)}] = \mathcal{E}[\mathbf{w}_{it-1}^{(g)}\mathbf{w}_{it-1}^{(h)}] + \frac{1}{NT}\sum_{t=1}^{T-1}\mathcal{E}[\boldsymbol{\mu}_t^{*(g)'}(\mathbf{I}_N - \mathbf{M}_t)\boldsymbol{\mu}_t^{*(h)}] \tag{6.43}$$

Moreover,

$$|\boldsymbol{\mu}_t^{*(g)'}(\mathbf{I}_N - \mathbf{M}_t)\boldsymbol{\mu}_t^{*(h)}| \leq |\boldsymbol{\pi}_{(g,1)}\boldsymbol{\pi}_{(h,1)}|(\boldsymbol{\mu}_t^{*'}\boldsymbol{\mu}_t^*), \ (g, h = 1, 2) \tag{6.44}$$

18

since $(\mathbf{I}_N - \mathbf{M}_t)$ is idempotent. Then

$$\frac{1}{NT}|\sum_{t=1}^{T-1}\mathcal{E}[\boldsymbol{\mu}_t^{*(g)'}(\mathbf{I}_N - \mathbf{M}_t)\boldsymbol{\mu}_t^{*(h)'}]| \leq \frac{|\boldsymbol{\pi}_{(g,1)}\boldsymbol{\pi}_{(h,1)}|}{NT}\sum_{t=1}^{T-1}\mathcal{E}[\boldsymbol{\mu}_t^{*'}\boldsymbol{\mu}_t^*] \quad (6.45)$$

$$= \frac{|\boldsymbol{\pi}_{(g,1)}\boldsymbol{\pi}_{(h,1)}|}{T}\sum_{t=1}^{T-1}\mathcal{E}[\mu_{it}^{*2}]$$

$$= \frac{1}{T}O(\log T)\ .$$

Hence

$$\frac{1}{NT}\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{w}_{t-1}'\mathbf{M}_t\mathbf{w}_{t-1}] - \mathcal{E}[\mathbf{w}_{i(t-1)}\mathbf{w}_{i(t-1)}'] \to 0\ . \quad (6.46)$$

To establish the mean-square convergence, by using the relations from (6.18) to (6.20), we have

$$Var[\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{w}_{t-1}'\mathbf{w}_{t-1}] = \frac{1}{N}Var[\frac{1}{T}\sum_{t=1}^{T-1}\mathbf{w}_{it-1}\mathbf{w}_{it-1}'] = \frac{1}{N}O(\frac{1}{T})\ , \quad (6.47)$$

and

$$Var[\frac{|\boldsymbol{\pi}_{(g,1)}\boldsymbol{\pi}_{(h,1)}|}{NT}\sum_{t=1}^{T-1}\boldsymbol{\mu}_t^{*'}\boldsymbol{\mu}_t^*]$$

$$= \frac{(\boldsymbol{\pi}_{(g,1)}\boldsymbol{\pi}_{(h,1)})^2}{N}Var[\frac{1}{T}\sum_{t=1}^{T-1}\mu_{it}^{*2}] \quad (6.48)$$

$$= \frac{(\boldsymbol{\pi}_{(g,1)}\boldsymbol{\pi}_{(h,1)})^2}{N}\left[\frac{1}{T^2}\sum_t Var[\mu_{it}^{*2}] + \frac{2}{T^2}\sum_s\sum_{t>s}Cov[\mu_{it}^{*2},\mu_{is}^{*2}]\right]$$

$$\leq \frac{(\boldsymbol{\pi}_{(g,1)}\boldsymbol{\pi}_{(h,1)})^2}{N}\left[\frac{1}{T^2}\sum_t O(\frac{1}{t^2}) + \frac{2}{T^2}\sum_s\sum_{t>s}O(\frac{1}{t})O(\frac{1}{s})\right]$$

$$= o(1)\ .$$

We have $\mathcal{E}[(1/NT)\sum_{t=1}^{T-1}\mathbf{w}_{t-1}^{(g)'}\mathbf{M}_t\mathbf{v}_t^{(2)}] = 0$ for $g = 1, 2$ and then

$$\mathcal{E}[\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{v}_t^{(2)'}\mathbf{M}_t\mathbf{v}_t^{(2)}] = \frac{1}{NT}\sum_{t=1}^{T-1}(2t)\mathcal{E}[v_{it}^{(2)2}] \to c\mathcal{E}[v_{it}^{(2)2}]. \quad (6.49)$$

Also by using $\mathbf{w}_t^{(2)} = \gamma_2 \mathbf{w}_{t-1}^{(2)} + \mathbf{v}_t^{(2)}$, it is sufficient to show that

$$
\begin{aligned}
Var[\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{w}_{t-1}^{(g)'}\mathbf{M}_t\mathbf{v}_t^{(2)}] &= \frac{\mathcal{E}[v_{it}^{(2)2}]}{N^2T^2}\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{w}_{t-1}^{(g)'}\mathbf{M}_t\mathbf{w}_{t-1}^{(g)}] \\
&= o(1) \; (g=1,2), \hspace{3cm} (6.50)\\
Var[\frac{1}{NT}\sum_{t=1}^{T-1}\mathbf{v}_t^{(2)'}\mathbf{M}_t\mathbf{v}_t^{(2)}] &\leq \frac{1}{N^2T^2}\sum_{t=1}^{T-1}\Big[(\mathcal{E}[v_{it}^{(2)2}])^2 + (\mathcal{E}[v_{it}^{(2)4}] - 3\mathcal{E}[v_{it}^{(2)2}])\Big](2t) \\
&= o(1) \,. \hspace{4.5cm} (6.51)
\end{aligned}
$$

Since $Cov[\mathbf{w}_{t-1}^{(g)'}\mathbf{M}_t\mathbf{v}_t^{(2)}, \mathbf{w}_{s-1}^{(g)}\mathbf{M}_s\mathbf{v}_s^{(2)}] = 0$ and $Cov[\mathbf{v}_t^{(2)'}\mathbf{M}_t\mathbf{v}_t^{(2)}, \mathbf{v}_s^{(2)}\mathbf{M}_s\mathbf{v}_s^{(2)}] = 0$ hold for $t > s$ $(g=1,2)$, the inequality holds due to *Lemma 1*. **Q.E.D.**

We use the decomposition

$$
\sum_{t=1}^{T-1}\mathbf{y}_{t-1}^{(1,f)'}\mathbf{M}_t\mathbf{u}_t^{(f)} = \sum_{t=1}^{T-1}c_t(\mathbf{w}_{t-1}^{(1)} - \frac{1}{T-t}\tilde{\mathbf{w}}_{t-1}^{(1)})'\mathbf{M}_t\mathbf{u}_t^{(f)} - \sum_{t=1}^{T-1}c_t\tilde{\mathbf{v}}_{tT}^{(1)'}\mathbf{M}_t\mathbf{u}_t^{(f)}, \; (6.52)
$$

where

$$
\begin{aligned}
\mathbf{w}_{t-1}^{(1)} &= (w_{1t-1}^{(1)}, ..., w_{Nt-1}^{(1)})', & (6.53)\\
\mathbf{u}_t^{(f)} &= (u_{1t}^{(1,f)}, ..., u_{Nt}^{(1,f)})', & (6.54)\\
\tilde{\mathbf{v}}_{tT}^{(1)} &= \frac{1}{T-t}[\phi_{T-t}^{(\gamma_1)}\mathbf{v}_t^{(1)} + ... + \phi_1^{(\gamma_1)}\mathbf{v}_{T-1}^{(1)}] & (6.55)\\
&\quad + \frac{\phi_\beta}{T-t}[(\phi_{T-t}^{(\gamma_1)} - \phi_{T-t}^{(\gamma_2)})\mathbf{v}_t^{(2)} + ... + (\phi_1^{(\gamma_1)} - \phi_1^{(\gamma_2)})\mathbf{v}_{T-1}^{(2)}] \\
&= \tilde{\mathbf{v}}_{tT}^{(11)} + \tilde{\mathbf{v}}_{tT}^{(12)} \; (, say), \\
\phi_j^{(x)} &= \frac{1-x^j}{1-x}, \;\; (x = \gamma_1, \gamma_2), & (6.56)\\
\tilde{\mathbf{w}}_{t-1}^{(1)} &= \sum_{j=0}^{\infty}[(\mathbf{v}_{t-1-j}^{(1)}, \mathbf{v}_{t-1-j}^{(2)})(\sum_{h=1}^{T-t}\mathbf{\Pi}_{(1)}^{j+h})'], & (6.57)
\end{aligned}
$$

and $\mathbf{\Pi}_{(1)}^{j+h}$ denotes the first row of $\mathbf{\Pi}^{j+h}$. Since only the second term in the right-hand side of (6.52) has nonzero mean (the same calculation of (A46) and (A47) of Alvaretz and Arellano (2003)), we obtain

$$
\begin{aligned}
\mathcal{E}[\sum_{t=1}^{T-1}\mathbf{y}_{t-1}^{(1,f)'}\mathbf{M}_t\mathbf{u}_t^{(f)}] &= 2T\Big(\frac{\mathcal{E}[\phi_\beta u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_2}[1 - \frac{1}{T(1-\gamma_2)}\sum_{t=1}^{T}\frac{(1-\gamma_2^t)}{t}] & (6.58)\\
&\quad - \frac{\mathcal{E}[u_{it}^{(1)}v_{it}^{(1)} + \phi_\beta u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_1}[1 - \frac{1}{T(1-\gamma_1)}\sum_{t=1}^{T}\frac{(1-\gamma_1^t)}{t}]\Big).
\end{aligned}
$$

Similarly, we decompose $\mathbf{y}_t^{(2,f)}$ as

$$\sum_{t=1}^{T-1} \mathbf{y}_t^{(2,f)'} \mathbf{M}_t \mathbf{u}_t^{(f)} = \sum_{t=1}^{T-1} c_t (\gamma_2 - \frac{\gamma_2^2 \phi_{T-t}^{(\gamma_2)}}{T-t}) \mathbf{w}_{t-1}^{(2)'} \mathbf{M}_t \mathbf{u}_t^{(f)} \qquad (6.59)$$

$$+ \sum_{t=1}^{T-1} c_t [(1 - \frac{\gamma_2 \phi_{T-t}^{(\gamma_2)}}{T-t}) \mathbf{v}_t^{(2)} - (\frac{1}{T-t})(\phi_{T-t}^{(\gamma_2)} \mathbf{v}_{t+1}^{(2)} + ... + \phi_1^{(\gamma_2)} \mathbf{v}_T^{(2)})]' \mathbf{M}_t \mathbf{u}_t^{(f)} .$$

Again we use the facts that (i) only the second term has non-zero mean, (ii) $\mathcal{E}[\mathbf{v}_{t+j}^{(2)'} \mathbf{M}_t \mathbf{u}_{t+k}^{(1)}] = (2t)\mathcal{E}[u_{it}^{(1)} v_{it}^{(2)}]$ for $j = k$ and zero otherwise, and (iii) $\phi_1^{(\gamma_2)} + ... + \phi_{T-t-1}^{(\gamma_2)} = (T - t - \phi_{T-t}^{(\gamma_2)})/(1 - \gamma_2)$. Then for each t, we have

$$\mathcal{E}[\mathbf{y}_t^{(2,f)'} \mathbf{M}_t \mathbf{u}_t^{(f)}] \qquad (6.60)$$

$$= c_t^2 (2t) \mathcal{E}[u_{it}^{(1)} v_{it}^{(2)}](1 - \frac{\gamma_2 \phi_{T-t}^{(\gamma_2)}}{T-t} + \frac{\phi_{T-t}^{(\gamma_2)} + ... + \phi_1^{(\gamma_2)}}{(T-t)^2})$$

$$= \frac{(2t)\mathcal{E}[u_{it}^{(1)} v_{it}^{(2)}]}{T-t+1}((T-t) + \frac{1}{1-\gamma_2} - \gamma_2 \phi_{T-t}^{(\gamma_2)} - (\frac{\gamma_2}{1-\gamma_2})(\frac{\phi_{T-t}^{(\gamma_2)}}{T-t})) .$$

Furthermore, by taking the sum of (6.59), we have

$$\mathcal{E}[\sum_{t=1}^{T-1} \mathbf{y}_t^{(2,f)'} \mathbf{M}_t \mathbf{u}_t^{(f)}] \qquad (6.61)$$

$$= 2\mathcal{E}[u_{it}^{(1)} v_{it}^{(2)}] \left(\frac{(T+2)(T-1)}{2} + (\frac{\gamma_2}{1-\gamma_2})(T+1)\sum_{s=2}^{T} \frac{1}{s} \right.$$

$$- (\frac{\gamma_2}{1-\gamma_2})(T-2) - (\frac{\gamma_2}{1-\gamma_2})[(T+1)\sum_{s=2}^{T} \frac{\phi_{s-1}^{(\gamma_2)}}{s(s-1)} - \sum_{s=2}^{T} \frac{\phi_{s-1}^{(\gamma_2)}}{(s-1)}]$$

$$\left. - \gamma_2[(T+1)\sum_{s=2}^{T} \frac{\phi_{s-1}^{(\gamma_2)}}{s} - \sum_{s=2}^{T} \phi_{s-1}^{(\gamma_2)}] \right)$$

$$= O(T^2) .$$

Proof of (3.5) : The convergence in probability to the limit of (6.58), (6.61) and the first factor of (3.5) are established by using (6.83) and (6.62) in the next subsection, respectively. **Q.E.D.**

6.3 The LIML Estimator

Proof of Theorem 2 : First, we need the convergence result on $(1/n)\mathbf{G}^{(f)}$ and $(1/q_n)\mathbf{H}^{(f)}$. We use the similar arguments as Akashi and Kunitomo (2010)

with Lemmas 1 and 2. Then we have

$$\frac{1}{n}\mathbf{G}^{(f)} \xrightarrow{p} \mathbf{G}_0 = \begin{pmatrix} \boldsymbol{\theta}' \\ \mathbf{I}_2 \end{pmatrix} \boldsymbol{\Phi}^*(\boldsymbol{\theta}, \mathbf{I}_2) + c \begin{pmatrix} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix} = \boldsymbol{\Phi}_\theta + c \begin{pmatrix} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix} \ (, say), (6.62)$$

where $\boldsymbol{\Phi}^* = \mathbf{D}'\boldsymbol{\Gamma}_0\mathbf{D}$.
Similarly,

$$\frac{1}{q_n}\mathbf{H}^{(f)} \xrightarrow{p} \mathbf{H}_0 = \begin{pmatrix} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix}. \tag{6.63}$$

Then we have

$$\left| \boldsymbol{\Phi}_\theta + (c - plim_{n\to\infty}\lambda_n) \begin{pmatrix} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0}' & 0 \end{pmatrix} \right| = 0. \tag{6.64}$$

By the singularity of $\boldsymbol{\Phi}_\theta$, we first notice that c is a solution. If $c > plim_{n\to\infty}\lambda_n$, then $|\,.\,| > 0$. Hence $\lambda_n \xrightarrow{p} c$. For the consistency, we use the representation

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_{LI} - \boldsymbol{\theta} &= (\frac{1}{n}\sum_{t=1}^{T-1}\mathbf{Y}_{t\cdot t-1}^{(f)'}\mathbf{M}_t\mathbf{Y}_{t\cdot t-1}^{(f)'} - \lambda_n[\frac{1}{q_n}\sum_{t=1}^{T-1}\mathbf{Y}_{t\cdot t-1}^{(f)'}(\mathbf{I}_N - \mathbf{M}_t)\mathbf{Y}_{t\cdot t-1}^{(f)'}])^{-1} \\
&\quad \times (\frac{1}{n}\sum_{t=1}^{T-1}\mathbf{Y}_{t\cdot t-1}^{(f)'}M_t\mathbf{u}_t^{(f)} - \lambda_n[\frac{1}{q_n}\sum_{t=1}^{T-1}\mathbf{Y}_{t\cdot t-1}^{(f)'}(\mathbf{I}_N - \mathbf{M}_t)\mathbf{u}_t^{(f)}]). \quad (6.65)
\end{aligned}$$

By (6.4) and (6.61), we find

$$(\frac{1}{n}\sum_{t=1}^{T-1}\mathbf{Y}_{t\cdot t-1}^{(f)'}\mathbf{M}_t\mathbf{u}_t^{(f)} - \lambda_n[\frac{1}{q_n}\sum_{t=1}^{T-1}\mathbf{Y}_{t\cdot t-1}^{(f)'}(\mathbf{I}_N - \mathbf{M}_t)\mathbf{u}_t^{(f)}]) \tag{6.66}$$

$$\xrightarrow{p} \begin{pmatrix} c\mathcal{E}[v_{it}^{(2)}u_{it}^{(1)}] - c[\frac{\mathcal{E}[v_{it}^{(2)}u_{it}^{(1)}] - c\mathcal{E}[v_{it}^{(2)}u_{it}^{(1)}]}{1-c}] \\ 0 - c0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

In order to prove the asymptotic normality of the LIML estimator, we shall utilize the next expression in several steps.

$$\boldsymbol{\Phi}^*\sqrt{n}\begin{pmatrix} \hat{\beta}_{2.LI} - \beta_2 \\ \hat{\gamma}_{1.LI} - \gamma_1 \end{pmatrix} \tag{6.67}$$

$$= \frac{1}{\sqrt{n}}\mathbf{D}'\sum_{t=1}^{T-1}\mathbf{Y}_{t-1}^{(f)'}\mathbf{N}_t\mathbf{u}_t^{(f)} + \frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\begin{pmatrix} \mathbf{u}_t^{(\perp,f)'} \\ \mathbf{0}' \end{pmatrix}\mathbf{N}_t\mathbf{u}_t^{(f)} + o_p(1)$$

$$= \frac{1}{\sqrt{n}}\mathbf{D}'\sum_{t=1}^{T-1}\begin{pmatrix} \mathbf{w}_{t-1}^{(1)'} \\ \mathbf{w}_{t-1}^{(2)'} \end{pmatrix}\mathbf{u}_t + \frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\begin{pmatrix} \mathbf{u}_t^{\perp'} \\ \mathbf{0}' \end{pmatrix}\mathbf{N}_t\mathbf{u}_t + O(1) + o_p(1)$$

$$= \mathbf{A}_{1n} + \mathbf{A}_{2n} + O(1) + o_p(1) \ (, say),$$

where

$$(\mathbf{u}_t^{(\perp,f)\prime}, \mathbf{u}_t^{\perp\prime}) = (0,1)[\mathbf{I}_2 - \frac{\boldsymbol{\Omega}\boldsymbol{\beta}\boldsymbol{\beta}'}{\boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{\beta}}](\mathbf{v}_t^{(f)\prime}, \mathbf{v}_t').\tag{6.68}$$

[**Step 1**] : Define $\mathbf{G}_1^{(f)}$, $\mathbf{H}_1^{(f)}$, $\lambda_{1n}^{(f)}$ and $\mathbf{b}_1$ by $\mathbf{G}_1^{(f)} = \sqrt{n}[(1/n)\mathbf{G}^{(f)} - \mathbf{G}_0]$, $\mathbf{H}_1^{(f)} = \sqrt{q_n}[(1/q_n)\mathbf{H}^{(f)} - \mathbf{H}_0]$, $\lambda_{1n}^{(f)} = \sqrt{n}(\lambda_n - c)$ and

$$\mathbf{b}_1 = \sqrt{n}\begin{pmatrix} 0 \\ \hat{\boldsymbol{\theta}} - \boldsymbol{\theta} \end{pmatrix}.\tag{6.69}$$

By substituting these variables into (2.9), we have

$$[\mathbf{G}_0 - c\mathbf{H}_0]\begin{pmatrix} 1 \\ -\boldsymbol{\theta} \end{pmatrix} + \frac{1}{\sqrt{n}}[\mathbf{G}_1^{(f)} - \lambda_{1n}^{(f)}\mathbf{H}_0]\begin{pmatrix} 1 \\ -\boldsymbol{\theta} \end{pmatrix} + \frac{1}{\sqrt{n}}[\mathbf{G}_0 - c\mathbf{H}_0]\mathbf{b}_1$$

$$-\frac{1}{\sqrt{q_n}}[c\mathbf{H}_1^{(f)}]\begin{pmatrix} 1 \\ -\boldsymbol{\theta} \end{pmatrix} = o_p(\frac{1}{\sqrt{n}}).\tag{6.70}$$

By multiplying $(1, -\boldsymbol{\theta})$ to (6.70) from the left, we have

$$\lambda_{1n}^{(f)} = \frac{(1, -\boldsymbol{\theta}')[\mathbf{G}_1^{(f)} - \sqrt{cc_*}\mathbf{H}_1^{(f)}](1, -\boldsymbol{\theta}')'}{(1, -\boldsymbol{\theta}')\mathbf{H}_0(1, -\boldsymbol{\theta}')'} + o_p(1),\tag{6.71}$$

where $c_* = c/(1-c)$. Then by using the rank relation of $\boldsymbol{\Phi}_\theta(1, -\boldsymbol{\theta}')' = \mathbf{0}$, we have

$$\boldsymbol{\Phi}_\theta\mathbf{b}_1 = [\mathbf{G}_1^{(f)} - \lambda_{1n}^{(f)}\mathbf{H}_0 - \sqrt{cc_*}\mathbf{H}_1^{(f)}]\begin{pmatrix} 1 \\ -\boldsymbol{\theta} \end{pmatrix} + o_p(1).\tag{6.72}$$

Also by multiplying $(0, \mathbf{I}_2)$ to (6.72) from the left and substituting $\lambda_{1n}^{(f)}$ for (6.72), we find

$$\boldsymbol{\Phi}^*\sqrt{n}\begin{pmatrix} \hat{\beta}_{2LI} - \beta_2 \\ \hat{\gamma}_{1LI} - \gamma_1 \end{pmatrix}\tag{6.73}$$

$$= (0, \mathbf{I}_2)[\mathbf{G}_1^{(f)} - \lambda_{1n}^{(f)}\mathbf{H}_0 - \sqrt{cc_*}\mathbf{H}_1^{(f)}]\begin{pmatrix} 1 \\ -\boldsymbol{\theta} \end{pmatrix} + o_p(1)$$

$$= (0, \mathbf{I}_2)[\mathbf{I}_3 - \frac{1}{\boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{\beta}}\begin{pmatrix} \boldsymbol{\Omega}\boldsymbol{\beta} \\ 0 \end{pmatrix}(1, -\boldsymbol{\theta}')][\mathbf{G}_1^{(f)} - \sqrt{cc_*}\mathbf{H}_1^{(f)}]\begin{pmatrix} 1 \\ -\boldsymbol{\theta} \end{pmatrix} + o_p(1),$$

where $\boldsymbol{\beta} = (1, -\beta_2)'$.

In order to evaluate $[\mathbf{G}_1^{(f)} - \sqrt{cc_*}\mathbf{H}_1^{(f)}](1, -\boldsymbol{\theta}')'$, we decompose $\mathbf{G}^{(f)}$ as

$$\mathbf{G}^{(f)} = \mathbf{G}^{(f,1)} + \mathbf{G}^{(f,2)} + \mathbf{G}^{(f,2)\prime} + \mathbf{G}^{(f,3)},\tag{6.74}$$

23

where

$$\mathbf{G}^{(f,1)} = \mathbf{D}^{*\prime} \sum_{t=1}^{T-1} \mathbf{Y}_{t-1}^{(f)\prime} \mathbf{M}_t \mathbf{Y}_{t-1}^{(f)} \mathbf{D}^*,$$

$$\mathbf{G}^{(f,2)} = \mathbf{D}^{*\prime} \sum_{t=1}^{T-1} \mathbf{Y}_{t-1}^{(f)\prime} \mathbf{M}_t (\mathbf{v}_t^{(f)}, \mathbf{0}),$$

$$\mathbf{G}^{(f,3)} = \sum_{t=1}^{T-1} \begin{pmatrix} \mathbf{v}_t^{(f)\prime} \\ \mathbf{0}' \end{pmatrix} \mathbf{M}_t (\mathbf{v}_t^{(f)}, \mathbf{0}),$$

$\mathbf{D}^* = \left[(\pi_{11}, \pi_{12})', \mathbf{D} \right]$ and $\mathbf{Y}_{t-1}^{(f)} = (\mathbf{y}_{t-1}^{(1,f)}, \mathbf{y}_{t-1}^{(2,f)})$. Then by using the relation $\mathbf{D}^*(1, -\boldsymbol{\theta}')' = \mathbf{0}$, we write

$$[\mathbf{G}_1^{(f)} - \sqrt{cc_*}\mathbf{H}_1^{(f)}] \begin{pmatrix} 1 \\ -\boldsymbol{\theta} \end{pmatrix} \tag{6.75}$$

$$= \frac{1}{\sqrt{n}} \mathbf{D}^{*\prime} \sum_{t=1}^{T-1} \mathbf{Y}_{t-1}^{(f)\prime} \mathbf{M}_t \mathbf{u}_t^{(f)} + \frac{1}{\sqrt{n}} [\sum_{t=1}^{T-1} \begin{pmatrix} \mathbf{v}_t^{(f)\prime} \\ \mathbf{0}' \end{pmatrix} \mathbf{M}_t \mathbf{u}_t^{(f)} - r_n \begin{pmatrix} \boldsymbol{\Omega}\boldsymbol{\beta} \\ \mathbf{0}' \end{pmatrix}]$$

$$- \frac{\sqrt{cc_*}}{\sqrt{q_n}} \mathbf{D}^{*\prime} \sum_{t=1}^{T-1} \mathbf{Y}_{t-1}^{(f)\prime} [\mathbf{I}_N - \mathbf{M}_t] \mathbf{u}_t^{(f)}$$

$$- \frac{\sqrt{cc_*}}{\sqrt{q_n}} [\sum_{t=1}^{T-1} \begin{pmatrix} \mathbf{v}_t^{(f)\prime} \\ \mathbf{0}' \end{pmatrix} [\mathbf{I}_N - \mathbf{M}_t] \mathbf{u}_t^{(f)} - q_n \begin{pmatrix} \boldsymbol{\Omega}\boldsymbol{\beta} \\ \mathbf{0}' \end{pmatrix}].$$

Also we use the relation $\sqrt{cc_*}/\sqrt{q_n} - c_*/\sqrt{n} = o(1)$ and set

$$\mathbf{N}_t = \mathbf{M}_t - c_*(\mathbf{I}_N - \mathbf{M}_t) = \frac{1}{1-c}[\mathbf{M}_t - c\mathbf{I}_N]. \tag{6.76}$$

Then,

$$\boldsymbol{\Phi}^* \sqrt{n} \begin{pmatrix} \hat{\beta}_{2LI} - \beta_2 \\ \hat{\gamma}_{1LI} - \gamma_1 \end{pmatrix} = \frac{1}{\sqrt{n}} \mathbf{D}' \sum_{t=1}^{T-1} \mathbf{Y}_{t-1}^{(f)\prime} \mathbf{N}_t \mathbf{u}_t^{(f)} + \frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \begin{pmatrix} \mathbf{u}_t^{(\perp,f)\prime} \\ \mathbf{0}' \end{pmatrix} \mathbf{N}_t \mathbf{u}_t^{(f)} + o_p(1).$$

and then Step 1 is established.

[**Step 2**] : Let $\bar{\mathbf{u}}_{tT} = (\mathbf{u}_t + ... + \mathbf{u}_T)/(T - t + 1)$. To evaluate the sampling error

of $\hat{\gamma}_{1.LI}$, we further decompose the two terms of (6.52) as [6]

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{y}_{t-1}^{(1,f)'} \mathbf{N}_t \mathbf{u}_t^{(f)}$$

$$= \frac{1}{1-c} \left[\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}^{(1)'} \mathbf{M}_t \mathbf{u}_t - \Upsilon_{11n}^{(1)} - \Upsilon_{12n}^{(1)} \right) - (\Upsilon_{21n}^{(1)} - \Upsilon_{22n}^{(1)}) \right]$$

$$- c_* \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}^{(1)'} \mathbf{u}_t - \Upsilon_{3n}^{(1)} \right), \tag{6.77}$$

where

$$\Upsilon_{11n}^{(1)} = \frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}^{(1)'} \mathbf{M}_t \bar{\mathbf{u}}_{tT}, \tag{6.78}$$

$$\Upsilon_{12n}^{(1)} = \frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \frac{c_t}{T-t} \tilde{\mathbf{w}}_{t-1}^{(1)'} \mathbf{M}_t \mathbf{u}_t^{(f)}, \tag{6.79}$$

$$\Upsilon_{21n}^{(1)} = \frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \tilde{\mathbf{v}}_{tT}^{(1)'} \mathbf{M}_t \mathbf{u}_t, \tag{6.80}$$

$$\Upsilon_{22n}^{(1)} = \frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \tilde{\mathbf{v}}_{tT}^{(1)'} \mathbf{M}_t \bar{\mathbf{u}}_{tT}, \tag{6.81}$$

$$\Upsilon_{3n}^{(1)} = \sqrt{\frac{T}{N}} \sum_{i=1}^{N} \bar{\mathbf{w}}_{i(-1)}^{(1)} \bar{\mathbf{u}}_i. \tag{6.82}$$

By using *Lemma 2*, the leading term of (6.77) and $(1/\sqrt{n}) \sum_t \mathbf{w}_{t-1}^{(2)} \mathbf{M}_t \mathbf{u}_t$ satisfy

$$Var\left[\frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}^{(g)'} \mathbf{M}_t \mathbf{u}_t \right] = \frac{\mathcal{E}[u_{it}^{(1)2}]}{NT} \sum_{t=1}^{T-1} \mathcal{E}[\mathbf{w}_{t-1}^{(g)'} \mathbf{M}_t \mathbf{w}_{t-1}^{(g)}] = O(1). \tag{6.83}$$

Next, we shall show that the variances of $\Upsilon_{11n}^{(1)}$, $\Upsilon_{12n}^{(1)}$, $\Upsilon_{21n}^{(1)}$ and $\Upsilon_{22n}^{(1)}$ tend to zero. First, we notice that $Var[\Upsilon_{3n}^{(1)}] \to 0$ from (6.11). Second,

$$Var[\Upsilon_{11n}^{(1)}] = \frac{1}{NT} \sum_{t=1}^{T-1} \sum_{s=1}^{T-1} \mathcal{E}[\mathbf{w}_{t-1}^{(1)'} \mathbf{M}_t \bar{\mathbf{u}}_{tT} \bar{\mathbf{u}}_{sT}' \mathbf{M}_s \mathbf{w}_{s-1}^{(1)}]. \tag{6.84}$$

---

[6]Note that $\mathbf{u}_t^{(f)} = (\mathbf{u}_t - \bar{\mathbf{u}}_{tT})/c_t$

For $t \geq s$,

$$\mathcal{E}[\mathbf{w}_{t-1}^{(1)'}\mathbf{M}_t\bar{\mathbf{u}}_{tT}\bar{\mathbf{u}}_{sT}'\mathbf{M}_s\mathbf{w}_{s-1}^{(1)}] \tag{6.85}$$

$$= \mathcal{E}[\mathbf{w}_{t-1}^{(1)'}\mathbf{M}_t\mathcal{E}_t[\bar{\mathbf{u}}_{tT}\bar{\mathbf{u}}_{sT}']\mathbf{M}_s\mathbf{w}_{s-1}^{(1)}]$$

$$= \frac{\mathcal{E}[u_{it}^{(1)2}]\mathcal{E}[\mathbf{w}_{t-1}^{(1)'}\mathbf{M}_s\mathbf{w}_{s-1}^{(1)}]}{(T-s+1)}$$

$$= \frac{\mathcal{E}[u_{it}^{(1)2}]}{(T-s+1)}\mathcal{E}[E_s[\mathbf{w}_{t-1}^{(1)'}]\mathbf{M}_s\mathbf{w}_{s-1}^{(1)}]$$

$$= \frac{\mathcal{E}[u_{it}^{(1)2}]}{(T-s+1)}\mathcal{E}[([\mathbf{\Pi}^{t-s}]_{(1,1)}\mathbf{w}_{s-1}^{(1)'} + [\mathbf{\Pi}^{t-s}]_{(1,2)}\mathbf{w}_{s-1}^{(2)'})\mathbf{M}_s\mathbf{w}_{s-1}^{(1)}],$$

where $[\mathbf{A}]_{(i,j)}$ is the (i,j)-th element of $\mathbf{A}$.

From the covariance-stationarity, we have $|\mathcal{E}[\mathbf{w}_{s-1}^{(1)'}\mathbf{M}_s\mathbf{w}_{s-1}^{(2)}]| \leq (\mathcal{E}[(\mathbf{w}_0^{(1)'}\mathbf{w}_0^{(1)})(\mathbf{w}_0^{(2)'}\mathbf{w}_0^{(2)})])^{1/2}$.

By using the relation $(\mathcal{E}[(\mathbf{w}_0^{(1)'}\mathbf{w}_0^{(1)})(\mathbf{w}_0^{(2)'}\mathbf{w}_0^{(2)})])^{1/2} = O(N)$,

$$Var[\Upsilon_{11n}^{(1)}] \leq \frac{\mathcal{E}[u_{it}^{(1)2}](\mathcal{E}[(\mathbf{w}_0^{(1)'}\mathbf{w}_0^{(1)})(\mathbf{w}_0^{(2)'}\mathbf{w}_0^{(2)})])^{1/2}}{NT} \tag{6.86}$$

$$\times [(\frac{1}{T} + ... + \frac{1}{2}) + (1 + |\phi_\beta|)S_{T1}(|\gamma_1|) + |\phi_\beta|S_{T1}(|\gamma_2|)]$$

$$= \frac{(\mathcal{E}[(\mathbf{w}_0^{(1)'}\mathbf{w}_0^{(1)})(\mathbf{w}_0^{(2)'}\mathbf{w}_0^{(2)})])^{1/2}}{N}\frac{\mathcal{E}[u_{it}^{(1)2}]}{T}$$

$$\times [(\frac{1+|\gamma_1|}{1-|\gamma_1|} + |\phi_\beta| - 2)S_{T2}(1) - \frac{2}{1-|\gamma_1|}S_{T2}(|\gamma_1|) + |\phi_\beta|S_{T2}(|\gamma_2|)]$$

$$= \frac{O(\log T)}{T} \to 0,$$

where

$$S_{T1}(x) = \frac{2}{T}(x + ... + x^{T-2}) + \frac{2}{T-1}(x + ... + x^{T-3}) + ... + \frac{2}{3}x,$$

$$S_{T2}(x) = \frac{x}{2} + ... + \frac{x^{T-1}}{T}.$$

Then,

$$Var[\Upsilon_{12n}^{(1)}] = \frac{1}{NT}Var[\sum_{t=1}^{T-1}\frac{c_t}{T-t}\tilde{\mathbf{w}}_{t-1}^{(1)'}\mathbf{M}_t\mathbf{u}_t^{(f)}] \tag{6.87}$$

$$= \frac{1}{NT}\sum_{t=1}^{T-1}\frac{c_t^2}{(T-t)^2}Var[\tilde{\mathbf{w}}_{t-1}^{(1)'}\mathbf{M}_t\mathbf{u}_t^{(f)}]$$

$$= \frac{\mathcal{E}[u_{it}^{(1)2}]}{NT}\sum_{t=1}^{T-1}\frac{c_t^2}{(T-t)^2}\mathcal{E}[\tilde{\mathbf{w}}_{t-1}^{(1)'}\mathbf{M}_t\tilde{\mathbf{w}}_{t-1}^{(1)}]$$

26

because

$$Var[\tilde{\mathbf{w}}_{t-1}^{(1)'}\mathbf{M}_t\mathbf{u}_t^{(f)}] = \mathcal{E}[\tilde{\mathbf{w}}_{t-1}^{(1)'}\mathbf{M}_t\mathbf{u}_t^{(f)}\mathbf{u}_t^{(f)'}\mathbf{M}_t\tilde{\mathbf{w}}_{t-1}^{(1)}] \tag{6.88}$$
$$= \mathcal{E}[\tilde{\mathbf{w}}_{t-1}^{(1)'}\mathbf{M}_t\mathcal{E}_t[\mathbf{u}_t^{(f)}\mathbf{u}_t^{(f)'}]\mathbf{M}_t\tilde{\mathbf{w}}_{t-1}^{(1)}]$$
$$= \mathcal{E}[u_{it}^{(1)2}]\mathcal{E}[\tilde{\mathbf{w}}_{t-1}^{(1)'}\mathbf{M}_t\tilde{\mathbf{w}}_{t-1}^{(1)}],$$

and the covariance terms are zero. In effect, for $t > s$, we have

$$Cov[\tilde{\mathbf{w}}_{t-1}^{(1)'}\mathbf{M}_t\mathbf{u}_t^{(f)},\ \tilde{\mathbf{w}}_{s-1}^{(1)'}\mathbf{M}_s\mathbf{u}_s^{(f)}] = \mathcal{E}[\tilde{\mathbf{w}}_{t-1}^{(1)'}\mathbf{M}_t\mathcal{E}_t[\mathbf{u}_t^{(f)}\mathbf{u}_s^{(f)'}]\mathbf{M}_s\tilde{\mathbf{w}}_{s-1}^{(1)}] = 0, \tag{6.89}$$

since $\mathcal{E}_t[\mathbf{u}_t^{(f)}\mathbf{u}_s^{(f)'}] = \mathbf{O}_N$. Moreover,

$$Var[\Upsilon_{12n}^{(1)}] \leq \frac{\mathcal{E}[u_{it}^{(1)2}]}{NT}\sum_{t=1}^{T-1}\frac{c_t^2}{(T-t)^2}\mathcal{E}[\tilde{\mathbf{w}}_{t-1}^{(1)'}\tilde{\mathbf{w}}_{t-1}^{(1)}] \tag{6.90}$$
$$= \frac{\mathcal{E}[u_{it}^{(1)2}]}{NT}\sum_{t=1}^{T-1}\frac{c_t^2 N}{(T-t)^2}[\sum_{j=0}^{\infty}(\sum_{h=1}^{T-t}\mathbf{\Pi}_{(1)}^{j+h})\mathbf{\Omega}(\sum_{h=1}^{T-t}\mathbf{\Pi}_{(1)}^{j+h})']$$
$$= o(1),$$

since the term $|[\sum_{j=0}^{\infty}\ \cdot\ ]|$ in the left-hand side is bounded by $c^{(1)}|\phi_{T-t}^{(\gamma_1)}|^2+c^{(1,2)}|\phi_{T-t}^{(\gamma_1)}||\phi_{T-t}^{(\gamma_2)}|$ $+c^{(2)}|\phi_{T-t}^{(\gamma_2)}|^2$ for some positive constants $c^{(1)}, c^{(1,2)}$ and $c^{(2)}$.

We turn to evaluate the variance of $\Upsilon_{21n}^{(1)}$. By using Lemma 1, the only non zero terms to be considered are $a_{0n}^{(11)}$, $a_{1n}^{(11)}$, $a_{0n}^{(12)}$ and $a_{1n}^{(12)}$, we have

$$Var[\Upsilon_{21n}^{(1)}] = \frac{1}{NT}(Var[\sum_{t=1}^{T-1}\mathbf{u}_t'\mathbf{M}_t\tilde{\mathbf{v}}_{tT}^{(11)}] + 2Cov[\sum_{t=1}^{T-1}\mathbf{u}_t'\mathbf{M}_t\tilde{\mathbf{v}}_{tT}^{(11)}, \sum_{t=1}^{T-1}\mathbf{u}_t'\mathbf{M}_t\tilde{\mathbf{v}}_{tT}^{(12)}]$$
$$+ Var[\sum_{t=1}^{T-1}\mathbf{u}_t'\mathbf{M}_t\tilde{\mathbf{v}}_{tT}^{(12)}])$$
$$= (a_{0n}^{(11)} + a_{1n}^{(11)}) + \frac{2}{NT}Cov[\sum_{t=1}^{T-1}\mathbf{u}_t'\mathbf{M}_t\tilde{\mathbf{v}}_{tT}^{(11)}, \sum_{t=1}^{T-1}\mathbf{u}_t'\mathbf{M}_t\tilde{\mathbf{v}}_{tT}^{(12)}]$$
$$+ (a_{0n}^{(12)} + a_{1n}^{(12)}), \tag{6.91}$$

where

$$a_{0n}^{(11)} = \frac{1}{NT}\sum_{t=1}^{T-1}\frac{1}{(T-t)^2}(\phi_{T-t}^{(\gamma_1)2}Var[\mathbf{u}_t'\mathbf{M}_t\mathbf{v}_t^{(1)}] + ... + \phi_1^{(\gamma_1)2}Var[\mathbf{u}_t'\mathbf{M}_t\mathbf{v}_{T-1}^{(1)}]),$$

$$a_{0n}^{(12)} = \frac{\phi_\beta^2}{NT}\sum_{t=1}^{T-1}\frac{1}{(T-t)^2}((\phi_{T-t}^{(\gamma_1)} - \phi_{T-t}^{(\gamma_2)})^2 Var[\mathbf{u}_t'\mathbf{M}_t\mathbf{v}_t^{(2)}]$$
$$+ ... + (\phi_1^{(\gamma_1)} - \phi_1^{(\gamma_2)})^2 Var[\mathbf{u}_t'\mathbf{M}_t\mathbf{v}_{T-1}^{(2)}]),$$

$$a_{1n}^{(11)} = \frac{2}{NT}\sum_{t=1}^{T-2}(\frac{\phi_{T-t-1}^{(\gamma_1)2}Cov[\mathbf{u}_t'\mathbf{M}_t\mathbf{v}_{t+1}^{(1)},\mathbf{u}_{t+1}'\mathbf{M}_{t+1}\mathbf{v}_{t+1}^{(1)}]}{(T-t)(T-t-1)}$$
$$+ ... + \frac{\phi_1^{(\gamma_1)2}Cov[\mathbf{u}_t'\mathbf{M}_t\mathbf{v}_{T-1}^{(1)},\mathbf{u}_{T-1}'\mathbf{M}_{T-1}\mathbf{v}_{T-1}^{(1)}]}{(T-t)}),$$

$$a_{1n}^{(12)} = \frac{2\phi_\beta^2}{NT}\sum_{t=1}^{T-2}(\frac{(\phi_{T-t-1}^{(\gamma_1)} - \phi_{T-t-1}^{(\gamma_2)})^2 Cov[\mathbf{u}_t'\mathbf{M}_t\mathbf{v}_{t+1}^{(2)},\mathbf{u}_{t+1}'\mathbf{M}_{t+1}\mathbf{v}_{t+1}^{(2)}]}{(T-t)(T-t-1)}$$
$$+ ... + \frac{(\phi_1^{(\gamma_1)} - \phi_1^{(\gamma_2)})^2 Cov[\mathbf{u}_t'\mathbf{M}_t\mathbf{v}_{T-1}^{(2)},\mathbf{u}_{T-1}'\mathbf{M}_{T-1}\mathbf{v}_{T-1}^{(2)}]}{(T-t)}).$$

By using *Lemma 1* and the fact that $\phi_j^{(\gamma_1)2} < 4/(1-\gamma_1)^2$ for all j, we have

$$a_{0n}^{(11)} = \frac{1}{NT}\sum_{t=1}^{T-1}\frac{1}{(T-t)^2}(\phi_{T-t}^{(\gamma_1)2}[(m^{(3)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)}) + m^{(2)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)}))(2t) \qquad (6.92)$$
$$+ m^{(0)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)})\mathcal{E}[\mathbf{d}_t'\mathbf{d}_t]] + (\phi_{T-t-1}^{(\gamma_1)2} + ... + \phi_1^{(\gamma_1)2})[m^{(3)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)})(2t)])$$
$$\leq \frac{1}{NT}\sum_{t=1}^{T-1}\frac{2t}{(T-t)^2}(\phi_{T-t}^{(\gamma_1)2}[m^{(3)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)}) + m^{(2)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)})$$
$$+ |m^{(0)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)})|] + (\phi_{T-t-1}^{(\gamma_1)2} + ... + \phi_1^{(\gamma_1)2})m^{(3)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)}))$$
$$\leq \frac{4\cdot 2}{(1-\gamma_1)^2}\frac{1}{NT}\sum_{t=1}^{T-1}\frac{t}{(T-t)^2}([m^{(3)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)}) + m^{(2)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)})$$
$$+ |m^{(0)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)})|] + (T-t-1)m^{(3)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)})) \qquad (6.93)$$
$$= O(\frac{\log T}{N}),$$

where $m^{(k)}(\mathbf{u}_t^{(1)},\mathbf{v}_t^{(1)})$ $(k=0,2,3)$ are defined as the same way as (6.22). The last equality follows from

$$\sum_{t=1}^{T-1}\frac{t(T-t-1)}{(T-t)^2} = \sum_{s=1}^{T-1}\frac{(T-s)(s-1)}{s^2} = T \times O(\log T). \qquad (6.94)$$

28

Moreover, from the fact that $|\mathcal{E}[\mathbf{d}'_{t+h}\mathbf{M}_t\mathbf{u}_t^{(1)}]| \leq 2(t+h)(\mathcal{E}[u_{it}^{(1)2}])^{1/2}$, we have

$$
\begin{aligned}
|a_{1n}^{(11)}| &= \frac{2}{NT}|\sum_{t=1}^{T-2}(\frac{\phi_{T-t-1}^{(\gamma_1)2}\mathcal{E}[v_{it}^{(1)2}u_{it}^{(1)}]\mathcal{E}[\mathbf{d}'_{t+1}\mathbf{M}_t\mathbf{u}_t^{(1)}]}{(T-t)(T-t-1)} \\
&\quad +...+ \frac{\phi_1^{(\gamma_1)2}\mathcal{E}[v_{it}^{(1)2}u_{it}^{(1)}]\mathcal{E}[\mathbf{d}'_{T-1}\mathbf{M}_t\mathbf{u}_t^{(1)}]}{(T-t)})| \\
&\leq \frac{4|\mathcal{E}[v_{it}^{(1)2}u_{it}^{(1)}]|(2^2\mathcal{E}[u_{it}^{(1)2}])^{1/2}}{(1-\gamma_1)^2}\frac{2}{NT}\sum_{t=1}^{T-2}\frac{1}{(T-t)}(\frac{t+1}{T-t-1}+...+\frac{T-1}{1}) \\
&\leq \frac{4|\mathcal{E}[v_{it}^{(1)2}u_{it}^{(1)}]|(2^2\mathcal{E}[u_{it}^{(1)2}])^{1/2}}{(1-\gamma_1)^2}\frac{2}{NT}(\frac{1}{2}+...+\frac{1}{T-1}) \\
&\quad \times [T(\frac{1}{2}+...+\frac{1}{T-1})+1] \\
&= O(\frac{(\log T)^2}{N}).
\end{aligned}
$$

Also $a_{0n}^{(12)} = O(\log T/N)$ and $a_{1n}^{(12)} = O((\log T)^2/N)$ are analogous to $a_{0n}^{(11)}$ and $a_{1n}^{(11)}$, respectively.

Finally, we consider the variance of $\Upsilon_{22n}^{(1)}$. By the same arguments as used for *Lemma 1*,

$$
\begin{aligned}
Var[\tilde{\mathbf{v}}_{tT}^{(1)'}\mathbf{M}_t\bar{\mathbf{u}}_{tT}] &= [m^{(3)}(\tilde{\mathbf{v}}_{tT}^{(1)},\bar{\mathbf{u}}_{tT})+m^{(2)}(\tilde{\mathbf{v}}_{tT}^{(1)},\bar{\mathbf{u}}_{tT})](2t)+m^{(0)}(\tilde{\mathbf{v}}_{tT}^{(1)},\bar{\mathbf{u}}_{tT})\mathcal{E}[\mathbf{d}'_t\mathbf{d}_t] \\
&\leq (2t)[m^{(1)}(\tilde{\mathbf{v}}_{tT}^{(1)},\bar{\mathbf{u}}_{tT})+m^{(3)}(\tilde{\mathbf{v}}_{tT}^{(1)},\bar{\mathbf{u}}_{tT})] \qquad (6.96)
\end{aligned}
$$

since $\mathcal{E}[\mathbf{d}'_t\mathbf{d}_t] \leq 2t$, $m^{(1)}(\tilde{\mathbf{v}}_{tT}^{(1)},\bar{\mathbf{u}}_{tT})-m^{(2)}(\tilde{\mathbf{v}}_{tT}^{(1)},\bar{\mathbf{u}}_{tT}) \geq 0$ and $m^{(k)}$ $(k=1,2,3)$ are non-negative.

Moreover,

$$
\begin{aligned}
m^{(3)}(\tilde{\mathbf{v}}_{tT}^{(1)},\bar{\mathbf{u}}_{tT}) &= Var[([\phi_{T-t}^{(\gamma_1)}v_{it}^{(1)}+\phi_\beta(\phi_{T-t}^{(\gamma_1)}-\phi_{T-t}^{(\gamma_2)})v_{it}^{(2)}] \qquad (6.97) \\
&\quad +...+[\phi_1^{(\gamma_1)}v_{iT-1}^{(1)}+\phi_\beta(\phi_1^{(\gamma_1)}-\phi_1^{(\gamma_2)})v_{iT-1}^{(2)}])/(T-t)] \\
&\quad \times Var[(u_{it}^{(1)}+...+u_{iT}^{(1)})/(T-t+1)] \\
&= O((\frac{1}{T-t})^2).
\end{aligned}
$$

It is because the terms $\boldsymbol{\phi}'_{T-t}\mathbf{v}_{it}$ and $\boldsymbol{\phi}'_{T-s}\mathbf{v}_{is}$ are independent for $t \neq s$, where $\boldsymbol{\phi}'_{T-t} = (\phi_{T-t}^{(\gamma_1)},\ \phi_\beta(\phi_{T-t}^{(\gamma_1)}-\phi_{T-t}^{(\gamma_2)}))$, and each $\|\boldsymbol{\phi}_{T-t}\|$ is bounded. As for the second term we have

$$
\begin{aligned}
m^{(1)}(\tilde{\mathbf{v}}_{tT}^{(1)},\bar{\mathbf{u}}_{tT}) &= \frac{1}{(T-t)^2(T-t+1)^2}\mathcal{E}[(\boldsymbol{\phi}'_{T-t}v_{it}+...+\boldsymbol{\phi}'_1v_{iT-1})^2 \quad (6.98) \\
&\quad \times (u_{it}^{(1)}+...+u_{iT}^{(1)})^2] \\
&= O(\frac{1}{(T-t)^2})\ ,
\end{aligned}
$$

since each terms of $\mathcal{E}[(\phi'_{T-t}v_{it})^2 u_{it}^{(1)2}]$ and $\mathcal{E}[(\phi'_{T-t}v_{it})^2]\mathcal{E}[u_{is}^{(1)2}]$ ($t \neq s$) appeared are $O((T-t+1)(T-t))$ in the numerator. Thus

$$Var[\tilde{\mathbf{v}}_{tT}^{(1)'}\mathbf{M}_t\bar{\mathbf{u}}_{tT}] = O(\frac{t}{(T-t)^2}). \tag{6.99}$$

The variance of $\Upsilon_{22n}^{(1)}$ is given by

$$Var[\Upsilon_{22n}^{(1)}] = \frac{1}{NT}Var[\sum_{t=1}^{T-1}\tilde{\mathbf{v}}_{tT}^{(1)'}\mathbf{M}_t\bar{\mathbf{u}}_{tT}] = b_{0n}^{(1)} + b_{1n}^{(1)}, \tag{6.100}$$

where

$$
\begin{aligned}
b_{0n}^{(1)} &= \frac{1}{NT}\sum_{t=1}^{T-1}Var[\tilde{\mathbf{v}}_{tT}^{(1)'}\mathbf{M}_t\bar{\mathbf{u}}_{tT}] \\
&= \frac{1}{NT}O(\sum_{t=1}^{T-1}\frac{t}{(T-t)^2}) = O(\frac{1}{N}), \tag{6.101} \\
b_{1n}^{(1)} &= \frac{2}{NT}\sum_{s}\sum_{t>s}Cov[\tilde{\mathbf{v}}_{tT}^{(1)'}\mathbf{M}_t\bar{\mathbf{u}}_{tT}, \tilde{\mathbf{v}}_{sT}^{(1)'}\mathbf{M}_s\bar{\mathbf{u}}_{sT}]. \tag{6.102}
\end{aligned}
$$

Since $Cov[\tilde{\mathbf{v}}_{tT}^{(1)'}\mathbf{M}_t\bar{\mathbf{u}}_{tT}, \tilde{\mathbf{v}}_{sT}^{(1)'}\mathbf{M}_s\bar{\mathbf{u}}_{sT}] \leq (Var[\tilde{\mathbf{v}}_{tT}^{(1)'}\mathbf{M}_t\bar{\mathbf{u}}_{tT}])^{1/2}(Var[\tilde{\mathbf{v}}_{sT}^{(1)'}\mathbf{M}_s\bar{\mathbf{u}}_{sT}])^{1/2}$, we can evaluate

$$
\begin{aligned}
|b_{1n}^{(1)}| &\leq \frac{2}{NT}\sum_{s}\sum_{t>s}|Cov[\tilde{\mathbf{v}}_{tT}^{(1)'}\mathbf{M}_t\bar{\mathbf{u}}_{tT}, \tilde{\mathbf{v}}_{sT}^{(1)'}\mathbf{M}_s\bar{\mathbf{u}}_{sT}]| \tag{6.103} \\
&\leq \frac{2}{NT}\sum_{s}\sum_{t>s}O(\frac{\sqrt{t}}{T-t})O(\frac{\sqrt{s}}{T-s}) \\
&\leq \frac{2}{NT}\sum_{t}O(\frac{t}{T-t})\sum_{s}O(\frac{1}{T-s}) \\
&= O(\frac{(\log T)^2}{N}).
\end{aligned}
$$

We turn to consider the sampling error of $\hat{\beta}_{2\,LI}$. First,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\mathbf{y}_{t-1}^{(2,f)'}\mathbf{N}_t\mathbf{u}_t^{(f)} \xrightarrow{p} \frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\mathbf{y}_{t-1}^{(2)'}\mathbf{N}_t\mathbf{u}_t + O(1) + o_p(1) \tag{6.104}$$

by using a result of Alvarez and Arellano (2003).

For the term $(1/\sqrt{n})\sum_t\mathbf{u}_t^{(\perp,f)'}\mathbf{N}_t\mathbf{u}_t^{(f)}$, by using $\mathbf{v}_t^{(g,f)} = (\mathbf{v}_t^{(g)} - \bar{\mathbf{v}}_{tT})/c_t$ and

$$\mathbf{u}_t^{(f)} = (\mathbf{u}_t - \bar{\mathbf{u}}_{tT})/c_t, \ (g=1,2),$$

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{v}_t^{(g,f)'} \mathbf{N}_t \mathbf{u}_t^{(f)} \tag{6.105}$$

$$= \quad \frac{1}{1-c} \Big[\Big(\frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{v}_t^{(g)'} \mathbf{M}_t \mathbf{u}_t + \frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \Big(\frac{1}{T-t} \Big) \mathbf{v}_t^{(g)'} \mathbf{M}_t \mathbf{u}_t \Big) - c_t^{-2} \mathbf{v}_t^{(g)'} \mathbf{M}_t \bar{\mathbf{u}}_{tT}$$

$$- c_t^{-2} \bar{\mathbf{v}}_{tT}^{(g)'} \mathbf{M}_t \mathbf{u}_t + c_t^{-2} \bar{\mathbf{v}}_{tT}^{(g)'} \mathbf{M}_t \bar{\mathbf{u}}_{tT} \Big] - c_* \Big(\frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{v}_t^{(g)'} \mathbf{u}_t - \sqrt{\frac{T}{N}} \sum_{i=1}^{N} \bar{\mathbf{v}}_i^{(g)} \bar{\mathbf{u}}_i \Big).$$

By using *Lemma 1*, $Var[\mathbf{v}_t^{(g)'} \mathbf{M}_t \mathbf{u}_t] \leq O(t)$ and $Cov[\mathbf{v}_t^{(g)'} \mathbf{M}_t \mathbf{u}_t, \mathbf{v}_s^{(g)'} \mathbf{M}_s \mathbf{u}_s] = 0$, we have

$$Var\Big[\frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \Big(\frac{1}{T-t} \Big) \mathbf{v}_t^{(g)'} \mathbf{M}_t \mathbf{u}_t \Big] \leq \frac{1}{NT} \sum_{t=1}^{T-1} \Big(\frac{1}{T-t} \Big)^2 O(t) = o(1) . \tag{6.106}$$

The third, fourth and fifth terms of (6.105) are analogous to $\Upsilon_{21n}^{(1)}$, $\Upsilon_{22n}^{(1)}$, respectively, and their variances tend to zero. The last term also tends to zero by the similar arguments as for (6.11).

Also we have the relation that

$$\frac{1}{1-c} \frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}' \mathbf{M}_t \mathbf{u}_t - \frac{c_*}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}' \mathbf{u}_t = \frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}' \mathbf{u}_t + o_p(1). \tag{6.107}$$

Since

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}' \mathbf{M}_t \mathbf{u}_t \quad = \quad \frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}' \mathbf{u}_t - \frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}' (\mathbf{I}_N - \mathbf{M}_t) \mathbf{u}_t \tag{6.108}$$

and *Lemma 2*, for $g = 1, 2$,

$$Var\Big[\frac{1}{\sqrt{n}} \sum_{t=1}^{T-1} \mathbf{w}_{t-1}^{(g)'} (\mathbf{I}_N - \mathbf{M}_t) \mathbf{u}_t \Big] \quad = \quad \frac{1}{NT} \sum_{t=1}^{T-1} Var[\mathbf{w}_{t-1}^{(g)'} (\mathbf{I}_N - \mathbf{M}_t) \mathbf{u}_t] \tag{6.109}$$

$$= \quad \frac{\mathcal{E}[u_{it}^{(1)2}]}{NT} \sum_{t=1}^{T-1} \mathcal{E}[\mathbf{w}_{t-1}^{(g)'} (\mathbf{I}_N - \mathbf{M}_t) \mathbf{w}_{t-1}^{(g)}]$$

$$= \quad \frac{\mathcal{E}[u_{it}^{(1)2}]}{NT} \sum_{t=1}^{T-1} \mathcal{E}[\boldsymbol{\mu}_t^{*(g)'} (\mathbf{I}_N - \mathbf{M}_t) \boldsymbol{\mu}_t^{*(g)}]$$

$$= \quad \frac{1}{T} O(\log T) .$$

For the above arguments, the leading order which has to tend to zero is $(\log T)^2/N$. Provided $T/N \to c \geq 0$, we have $(\log T)^2/N \sim c(\log T)^2/T \to 0$. It is because

31

$\lim_{T \to \infty} c(\log T)^2/T = \lim_{T \to \infty} 2c(\log T/T) = 0$, then Step 2 has been established.

[**Step 3**] : We shall evaluate the asymptotic variance-covariance terms of (6.68). First, by using the stationarity and direct calculations, we have

$$\mathcal{E}[\mathbf{A}_{1n}\mathbf{A}'_{1n}] = \frac{\sigma^2}{n}\mathbf{D}'\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{w}'_{t-1}\mathbf{w}_{t-1}]\mathbf{D} \to \sigma^2\mathbf{\Phi}^* ,\qquad (6.110)$$

where $\sigma^2 = \mathcal{E}[u_{it}^{(1)2}]$.
Under Assumptions (A1)-(A4) stated, we have

$$\frac{1}{n}\sum_{t=1}^{T-1}\mathbf{w}'_{t-1}\mathbf{d}_t \xrightarrow{p} \lim_{n \to \infty}\frac{1}{n}\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{w}'_{t-1}\mathbf{d}_t] ,\qquad (6.111)$$

and

$$\frac{1}{n}\sum_{t=1}^{T-1}\mathbf{d}'_t\mathbf{d}_t \xrightarrow{p} \lim_{n \to \infty}\frac{1}{n}\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{d}'_t\mathbf{d}_t] ,\qquad (6.112)$$

where $\mathbf{w}'_{t-1} = (w_{1(t-1)}, ..., w_{N(t-1)})$, $\{w_{it}\}$ are defined by (3.1) and $\mathbf{d}_t$ denotes the $N \times 1$ vector containing the diagonal elements of $\mathbf{M}_t = \mathbf{Z}_t(\mathbf{Z}'_t\mathbf{Z}_t)^{-1}\mathbf{Z}'_t$.

Then by the facts that the $i-th$ element of $\mathcal{E}_t[\mathbf{u}_t\mathbf{u}_t^{\perp'}\mathbf{N}_t\mathbf{u}_t]$ is equal to $N_{t(i,i)}\mathcal{E}[u_{it}^{(1)2}u_{it}^{\perp}]$ and $\mathcal{E}[w_{it-1}^{(g)}] = 0$ $(g = 1, 2)$, we have

$$\begin{aligned}\mathcal{E}[\mathbf{A}_{1n}\mathbf{A}'_{2n}] &= \left[\frac{1}{n}\mathbf{D}'\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{w}'_{t-1}\mathcal{E}_t[\mathbf{u}_t\mathbf{u}_t^{\perp'}\mathbf{N}_t\mathbf{u}_t]], \mathbf{0}\right] &(6.113)\\ &= \left[\frac{\mathcal{E}(u_{it}^{(1)2}u_{it}^{\perp})}{1-c}\frac{1}{n}\mathbf{D}'\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{w}'_{t-1}(\mathbf{d}_t - c\boldsymbol{\iota}_N)], \mathbf{0}\right]\\ &\to \left[\frac{\mathcal{E}[u_{it}^{(1)2}u_{it}^{\perp}]}{1-c}\lim_{n \to \infty}\frac{1}{n}\mathbf{D}'\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{w}'_{t-1}\mathbf{d}_t], \mathbf{0}\right].\end{aligned}$$

Next, we use the decomposition following Kunitomo and Akashi (2010) as

$$\mathcal{E}[\mathbf{A}_{2n}\mathbf{A}'_{2n}] = \frac{1}{n}\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{u}_t^{\perp'}\mathbf{N}_t[\sigma^2\mathbf{I}_N + (\mathbf{u}_t\mathbf{u}'_t - \sigma^2\mathbf{I}_N)]\mathbf{N}_t\mathbf{u}_t^{\perp}]\begin{pmatrix}1\\0\end{pmatrix}(1,0)\quad(6.114)$$

Then the first term converges to

$$\frac{1}{n}\sum_{t=1}^{T-1}tr(\mathbf{N}_t^2)\sigma^2\mathcal{E}[u_{it}^{\perp 2}] \to (c_*\sigma^2)\mathcal{E}[u_{it}^{\perp 2}] = c_*|\mathbf{\Omega}|,\qquad (6.115)$$

where $\mathcal{E}[u_{it}^{\perp 2}] = (1/\sigma^2)[\boldsymbol{\Omega}\sigma^2 - \boldsymbol{\Omega}\boldsymbol{\beta}\boldsymbol{\beta}'\boldsymbol{\Omega}]_{(2,2)}$ and $[\,.\,]_{(2,2)} = |\boldsymbol{\Omega}|$. For the second term, we have

$$\frac{1}{n}\sum_{t=1}^{T-1}\mathcal{E}[\mathbf{u}_t^{\perp'}\mathbf{N}_t[\mathbf{u}_t\mathbf{u}_t' - \sigma^2\mathbf{I}_N]\mathbf{N}_t\mathbf{u}_t^{\perp}] \tag{6.116}$$

$$= \frac{1}{n}\sum_{t=1}^{T-1}\mathcal{E}[\sum_{i,j=1}^{N}[\mathbf{u}_t\mathbf{u}_t' - \sigma^2\mathbf{I}_N]_{(i,j)}(u_{1t}^{\perp}\mathbf{N}_{t(i,1)} + ... + u_{Nt}^{\perp}\mathbf{N}_{t(i,N)})$$

$$\times (u_{1t}^{\perp}\mathbf{N}_{t(j,1)} + ... + u_{Nt}^{\perp}\mathbf{N}_{t(j,N)})]$$

$$= \frac{1}{n}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\mathcal{E}[(\mathbf{N}_{t(i,i)})^2\mathcal{E}_t[(u_{it}^2 - \sigma^2)u_{it}^{\perp 2}]]$$

$$= \frac{\mathcal{E}[(u_{it}^{(1)2} - \sigma^2)u_{it}^{\perp 2}]}{(1-c)^2}\frac{1}{n}\sum_{t=1}^{T-1}\sum_{i=1}^{N}\mathcal{E}[(\mathbf{M}_{t(i,i)} - c)^2]$$

$$= \frac{\mathcal{E}[(u_{it}^{(1)2} - \sigma^2)u_{it}^{\perp 2}]}{(1-c)^2}\frac{1}{n}\sum_{t=1}^{T-1}(\mathcal{E}[\mathbf{d}_t'\mathbf{d}_t] - 2tr(\mathbf{M}_t)c + Nc^2),$$

the second equality is from the facts that $\mathcal{E}_t[u_{it}^{\perp}u_{jt}] = \mathcal{E}[u_{it}^{\perp}u_{jt}] = 0$ for any i, j and that $\mathcal{E}[[\mathbf{u}_t\mathbf{u}_t' - \sigma^2\mathbf{I}_N]_{(i,i)}]\mathcal{E}[u_{jt}^{\perp 2}] = 0$ for $i \neq j$.

For any N, define 2×1 vectors as the martingale difference sequence

$$\mathbf{A}_t = \frac{1}{\sqrt{N}}\left[\mathbf{D}'\begin{pmatrix}\mathbf{w}_{t-1}^{(1)'}\\\mathbf{w}_{t-1}^{(2)'}\end{pmatrix}\mathbf{u}_t + \begin{pmatrix}\mathbf{u}_t^{\perp'}\\\mathbf{0}'\end{pmatrix}\mathbf{N}_t\mathbf{u}_t\right] = \mathbf{A}_{1t} + \mathbf{A}_{2t}\ (,say). \tag{6.117}$$

To apply the martingale central limit theorem, for any 2×1 vector $\mathbf{a}$, we check the condition that $(1/n)\sum_t\mathbf{a}'\mathcal{E}_t[(\mathbf{A}_{1t} + \mathbf{A}_{2t})(\mathbf{A}_{1t} + \mathbf{A}_{2t})']\mathbf{a} \xrightarrow{p} \lim_{n\to\infty}\mathbf{a}'\mathcal{E}[(\mathbf{A}_{1t} + \mathbf{A}_{2t})(\mathbf{A}_{1t} + \mathbf{A}_{2t})']\mathbf{a}$. As for the relevant Lyapounov conditions hold from the result of Kunitomo and Akashi (2010).

By using the facts that $(1/n)\sum_t\mathbf{w}_{t-1}'\mathbf{w}_{t-1} \xrightarrow{p} \boldsymbol{\Gamma}_0$, $(1/n)\sum_t\mathbf{w}_{t-1}'\boldsymbol{\iota}_N \xrightarrow{p} \mathbf{0}$, (6.110) and (6.111),

$$\frac{1}{T}\sum_{t=1}^{T-1}\mathcal{E}_t[\mathbf{A}_{1t}\mathbf{A}_{1t}'] \xrightarrow{p} \lim_{n\to\infty}\mathcal{E}[\mathbf{A}_{1t}\mathbf{A}_{1t}'], \quad \frac{1}{T}\sum_{t=1}^{T-1}\mathcal{E}_t[\mathbf{A}_{1t}\mathbf{A}_{2t}'] \xrightarrow{p} \lim_{n\to\infty}\mathcal{E}[\mathbf{A}_{1t}\mathbf{A}_{2t}'],$$

and $(1/T)\sum_{t=1}^{T-1}\mathcal{E}_t[\mathbf{A}_{2t}\mathbf{A}_{2t}'] \xrightarrow{p} \lim_{n\to\infty}\mathcal{E}[\mathbf{A}_{2t}\mathbf{A}_{2t}']$.

[**Step 4**] : Finally, we evaluate the asymptotic bias in the right-hand side of (6.66). It is a collection of the terms which the mean-square convergences to non-zero means, and it can be evaluated as

$$\lim_{n\to\infty}\boldsymbol{\mu}_n = \lim_{n\to\infty}\mathcal{E}[\frac{1}{\sqrt{n}}\mathbf{D}'\sum_{t=1}^{T-1}\mathbf{Y}_{t-1}^{(f)'}\mathbf{N}_t\mathbf{u}_t^{(f)} + \frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\begin{pmatrix}\mathbf{u}^{(f,\perp)'}\\\mathbf{0}'\end{pmatrix}\mathbf{N}_t\mathbf{u}_t^{(f)}]. \tag{6.118}$$

33

We first notice that

$$\mathcal{E}[\frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\mathbf{u}_t^{(\perp,f)'}\mathbf{N}_t\mathbf{u}_t^{(f)}] = \frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\mathcal{E}[tr(\mathbf{N}_t\mathcal{E}[\mathbf{u}_t^{(f)}\mathbf{u}_t^{(\perp,f)'}])] = 0. \qquad (6.119)$$

From the result of (6.4) and (6.58),

$$\lim_{n\to\infty}\mathcal{E}[\frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\mathbf{y}_{t-1}^{(1,f)'}\mathbf{N}_t\mathbf{u}_t^{(f)}] \qquad (6.120)$$

$$= \lim_{n\to\infty}\frac{1}{1-c}\frac{2T}{\sqrt{NT}}(\frac{\mathcal{E}[\phi_\beta u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_2} - \frac{\mathcal{E}[u_{it}^{(1)}v_{it}^{(1)} + \phi_\beta u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_1})$$

$$- \lim_{n\to\infty}\frac{c}{1-c}\frac{N}{\sqrt{NT}}(\frac{\mathcal{E}[\phi_\beta u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_2} - \frac{\mathcal{E}[u_{it}^{(1)}v_{it}^{(1)} + \phi_\beta u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_1})$$

$$= (\frac{\sqrt{c}}{1-c})(\frac{\mathcal{E}[\phi_\beta u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_2} - \frac{\mathcal{E}[u_{it}^{(1)}v_{it}^{(1)} + \phi_\beta u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_1})$$

$$= (\frac{-\sqrt{c}}{1-c})(\frac{\mathcal{E}[u_{it}^{(1)}v_{it}^{(1)}]}{1-\gamma_1} + \frac{\beta_2\gamma_2\mathcal{E}[u_{it}^{(1)}v_{it}^{(2)}]}{(1-\gamma_1)(1-\gamma_2)}).$$

By the similar calculation as for (6.58),

$$\mathcal{E}[\mathbf{y}_{t-1}^{(2,f)'}\mathbf{M}_t\mathbf{u}_t^{(f)}] = -\frac{2T\mathcal{E}[u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_2}[1 - \frac{1}{T(1-\gamma_2)}\sum_{t=1}^{T}\frac{(1-\gamma_2^t)}{t}], \qquad (6.121)$$

and from the result of (6.5),

$$\lim_{n\to\infty}\mathcal{E}[\frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\mathbf{y}_{t-1}^{(2,f)'}\mathbf{N}_t\mathbf{u}_t^{(f)}] = \lim_{n\to\infty}\frac{1}{1-c}\frac{2T}{\sqrt{NT}}(-\frac{\mathcal{E}[u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_2}) \qquad (6.122)$$

$$- \lim_{n\to\infty}\frac{c}{1-c}\frac{N}{\sqrt{NT}}(-\frac{\mathcal{E}[u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_2})$$

$$= (\frac{-\sqrt{c}}{1-c})(\frac{\mathcal{E}[u_{it}^{(1)}v_{it}^{(2)}]}{1-\gamma_2}).$$

Hence we summarize these results as

$$\lim_{n\to\infty}\mu_n = (\frac{-\sqrt{c}}{1-c})\mathbf{D}'(\mathbf{I}_2 - \mathbf{\Pi})^{-1}\mathbf{\Omega}\boldsymbol{\beta}, \qquad (6.123)$$

where $\mathbf{\Omega}\boldsymbol{\beta} = \mathcal{E}[\mathbf{v}_{it}u_{it}^{(1)}]$. Thus, we can define the asymptotic bias of $\sqrt{n}(\hat{\boldsymbol{\theta}}_{LI} - \boldsymbol{\theta})$ by

$$\mathbf{b} = \mathbf{\Phi}^{*-1}\lim_{n\to\infty}\boldsymbol{\mu}_n. \qquad (6.124)$$

6.4 GMM (Derivation of (3.5))

We now turn to the asymptotic covariance and the bias of the GMM estimator when $c = 0$. (The general case has been treated in Akashi and Kunitomo (2010).) The normalized GMM estimators are asymptotically equivalent to

$$\mathbf{G}_0\sqrt{n}(\hat{\boldsymbol{\theta}}_{GM} - \boldsymbol{\theta})$$

$$= \frac{1}{\sqrt{n}}\mathbf{D}'\sum_{t=1}^{T-1}\begin{pmatrix}\mathbf{w}_t^{(1)}\\\mathbf{w}_t^{(2)}\end{pmatrix}\mathbf{u}_t + \frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\begin{pmatrix}\mathbf{V}_t^{(2)}\\\mathbf{O}\end{pmatrix}\mathbf{M}_t\mathbf{u}_t + o_p(1) .$$

By using *Lemma 1*,

$$Var\Big[\frac{1}{\sqrt{n}}\sum_{t=1}^{T-1}\mathbf{V}_t^{(2)'}\mathbf{M}_t\mathbf{u}_t\Big] = \frac{1}{NT}\sum_{t=1}^{T-1}Var[\mathbf{V}_t^{(2)'}\mathbf{M}_t\mathbf{u}_t]$$

$$= O(c) . \tag{6.125}$$

Then the asymptotic covariance-variance matrix becomes $\mathbf{G}_0^{-1}(\sigma^2\boldsymbol{\Phi}^*)\mathbf{G}_0^{-1} = \sigma^2\boldsymbol{\Phi}^{*-1}$. Under the condition $\sum_t tr(\mathbf{M}_t)\,/(\sqrt{NT}) < \infty$, its asymptotic bias is given by

$$\mathbf{b}_0^{(a)} = \lim_{N,T\to\infty}\Big[\frac{1}{\sqrt{NT}}\sum_{t=1}^{T-1}tr(\mathbf{M}_t)\Big]\boldsymbol{\Phi}^{*-1}\begin{pmatrix}(1,0)\mathcal{E}[\mathbf{v}_{it}u_{it}^{(1)}]\\\mathbf{0}\end{pmatrix} . \tag{6.126}$$

The derivation of the asymptotic normality of the GMM estimator is similar to the one for the LIML estimator and it is omitted.

Q.E.D.

References

[1] Akashi, K. (2008), "*t*-Test in Dynamic Panel Structural Equations," *Unpublished Manuscript.*

[2] Akashi, K. and Kunitomo, N. (2010), "The Limited Information Maximum Likelihood Approach to Dynamic Panel Structural Equations," Discussion Paper CIRJE-F-708, Graduate School of Economics, University of Tokyo, (http://www.e.u-tokyo.ac.jp/cirje/research/dp/2010).

[3] Alvarez, J and M. Arellano (2003), "The Time Series and Cross Section Asymptotics of Dynamic Panel Data Estimators," *Econometrica*, Vol. 71, 1121-1159.

[4] Anderson, T.W. and C. Hsiao (1981), "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association*, Vol. 76, 598-606.

[5] Anderson, T.W. and C. Hsiao (1982), "Formulation and Estimation of Dynamic Models with Panel Data," *Journal of Econometrics*, Vol. 18, 47-82.

[6] Anderson, T.W. and H. Rubin (1949), "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *Annals of Mathematical Statistics*, Vol. 20, 46-63.

[7] Anderson, T.W. and H. Rubin (1950), "The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations," it Annals of Mathematical Statistics, Vol. 21, 570-582.

[8] Anderson, T.W., N. Kunitomo, and Y. Matsushita (2005), "A New Light from Old Wisdoms : Alternative Estimation Methods of Simultaneous Equations with Possibly Many Instruments," Discussion Paper CIRJE-F-321, Graduate School of Economics, University of Tokyo.

[9] Anderson, T.W., N. Kunitomo, and Y. Matsushita (2008a), "On Finite Sample Properties of Alternative Estimators of Coefficients in a Structural Equation with Many Instruments," Discussion Paper CIRJE-F-576, Graduate School of Economics, University of Tokyo, forthcoming in *the Journal of Econometrics.*

[10] Anderson, T.W., N. Kunitomo and Y. Matsushita (2008b), "On the Asymptotic Optimality of the LIML Estimator with Possibly Many Instruments," Discussion Paper CIRJE-F-542, Graduate School of Economics, University of Tokyo, forthcoming in *the Journal of Econometrics.*

[11] Arellano, M. (2003), *Panel Data Econometrics*, Oxford University Press.

[12] Arellano, M. and S.R. Bond (1991), "Some Tests of Specification for Panel Data : Monte Carlo Evidence and an Application to Employment Equation," *Reviwe of Economics Studies*, Vol. 58, 277-297.

[13] Baltagi, Badi H. (2005), *Econometric Analysis of Panel Data*, 3rd Edition, John-Wiley.

[14] Blundell, R. and S. Bond (2000), "GMM Estimation with Persistent Panel Data : An Application to Production Function," *Econometric Review*, Vol. 19-3, 321-340.

[15] Hsiao, Cheng (2003), *Analysis of Panel Data*, 2nd Edition, Cambridge University Press.

APPENDIX : Some Figures

In Figures the distribution functions of the WG, the GMM and the LIML estimators are shown with the large sample normalization (i.e. the case of $c = 0$) and the large-K normalization. The limiting distributions for the LIML estimator in the large-K asymptotics are $N_2(\mathbf{0}, \mathbf{I}_2)$ and its marginal distributions are $N(0,1)$ as $n \to \infty$, which are denoted as "o". For the sake of comparisons, the distribution functions of the WG and the GMM estimators are normalized in the same way and presented in figures. The parameters of our settings and the details of numerical computation method are similar to those explained in Anderson et al. (2005, 2008a). The BB model stands for the one by Blundel and Bond (2000) and the AA model stands for the one by Alvarez and Arellano (2003). The just identified model is the simultaneous equation with $\gamma_1 = 0$.
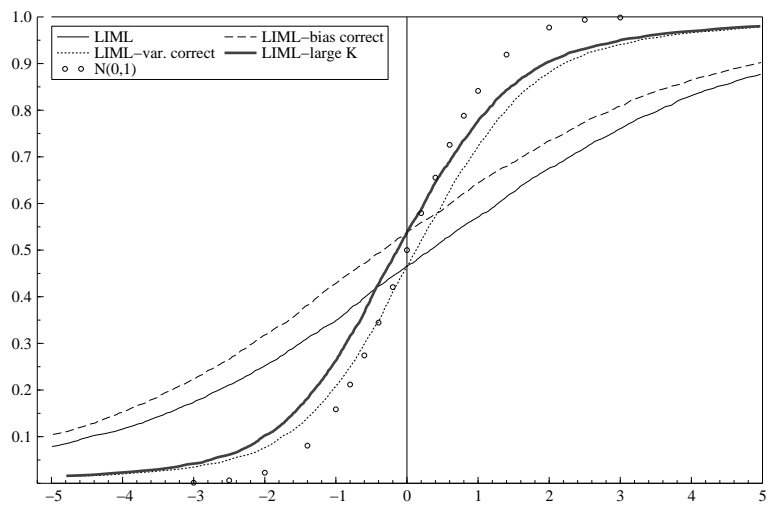
Figure 1: β_2 (BB-model) : $N = 75,\ T = 25,\ c = 1/3$
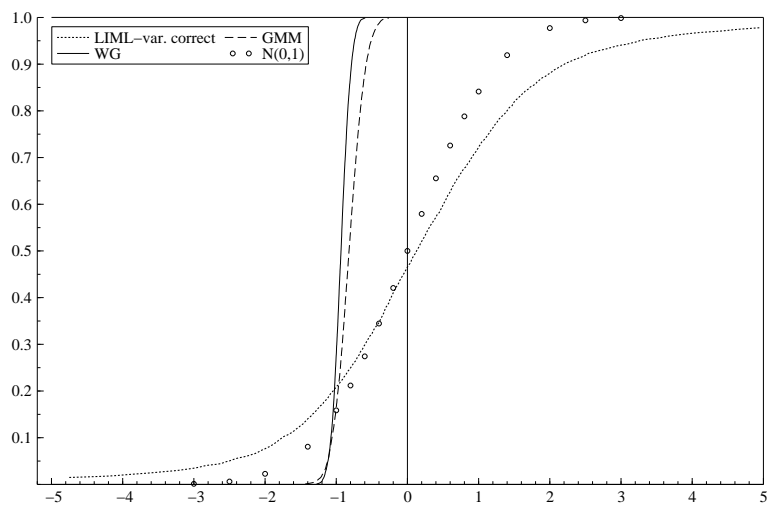
Figure 2: β_2 (BB-model) : $N = 75,\ T = 25\ ,c = 1/3$
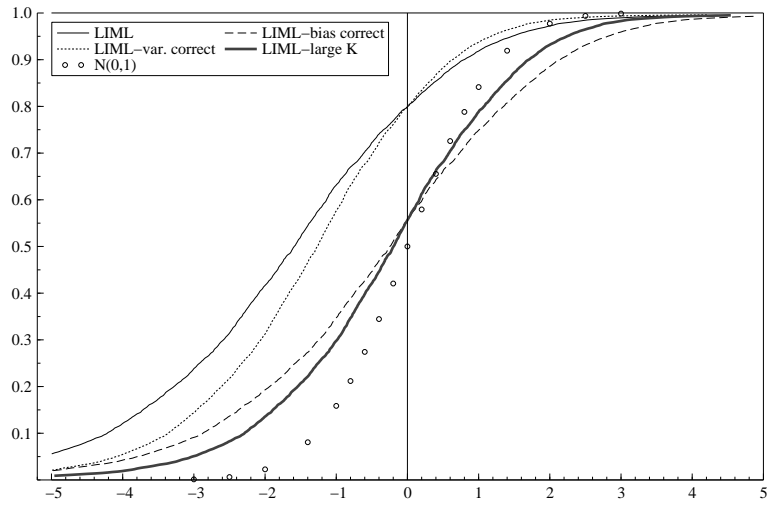
Figure 3: γ_1 (BB-model) : $N = 75$, $T = 25$, $c = 1/3$
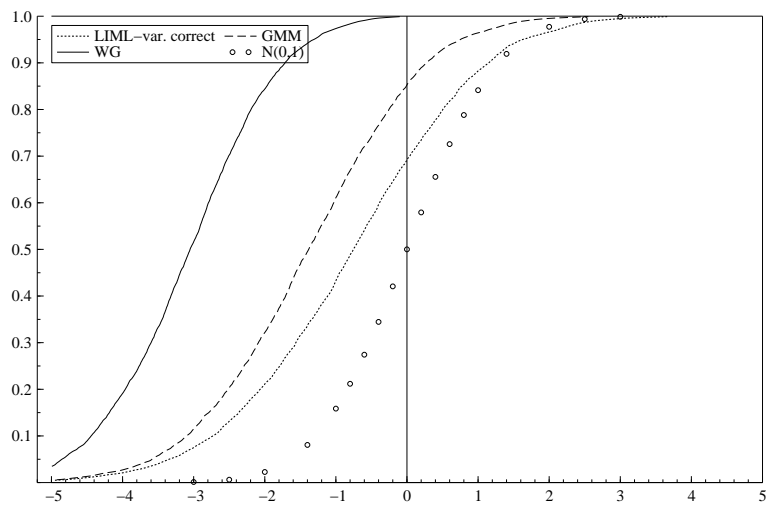
Figure 4: γ_1 (BB-model) : $N = 75$, $T = 25$, $c = 1/3$
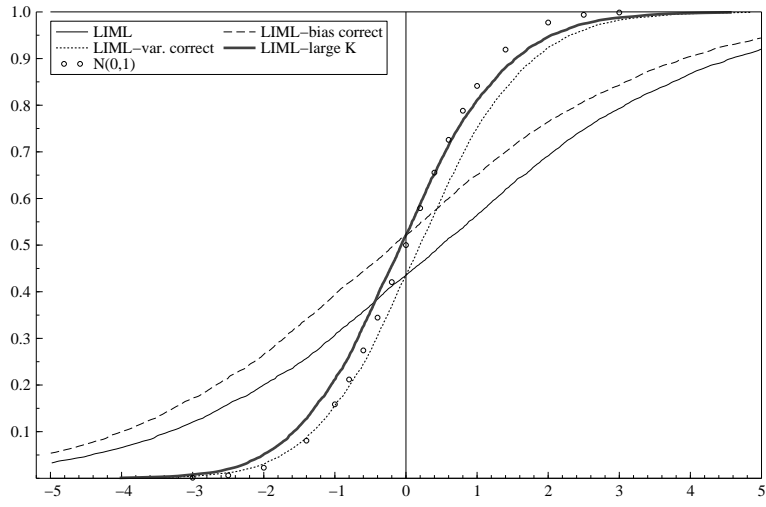
Figure 5: β_2 (BB-model) : $N = 150$, $T = 50$, $c = 1/3$

Figure 6: β_2 (BB-model) : $N = 150$, $T = 50$, $c = 1/3$

41
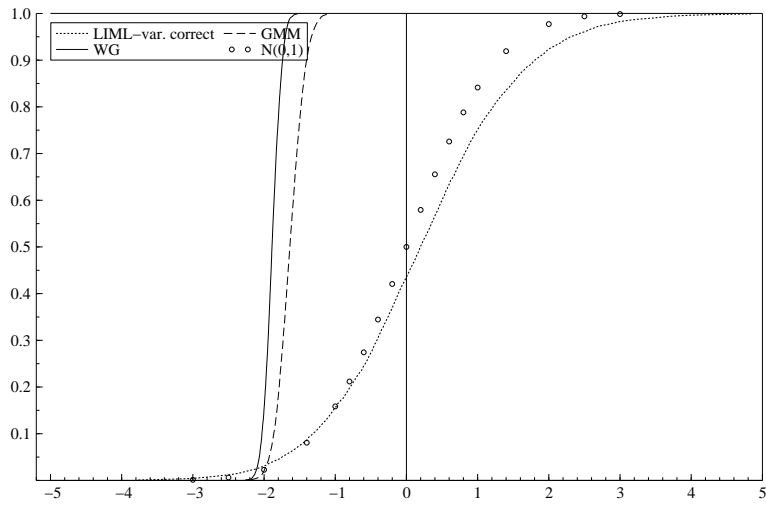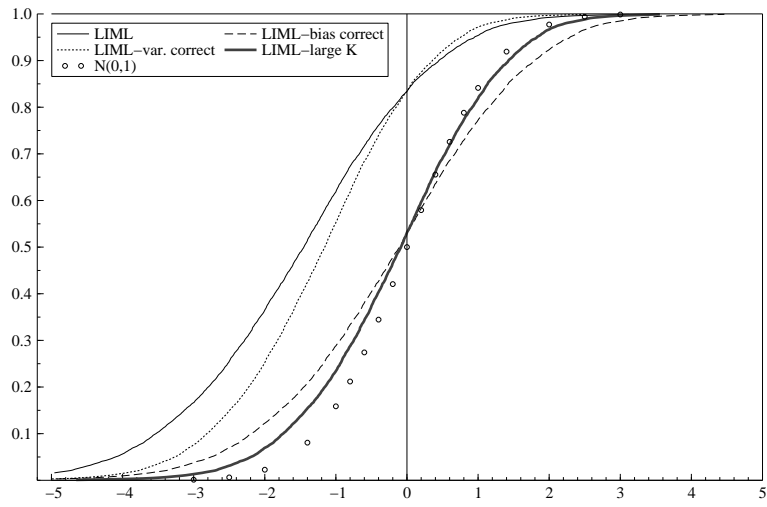
Figure 7: γ_1 (BB-model) : $N = 150$, $T = 50$, $c = 1/3$
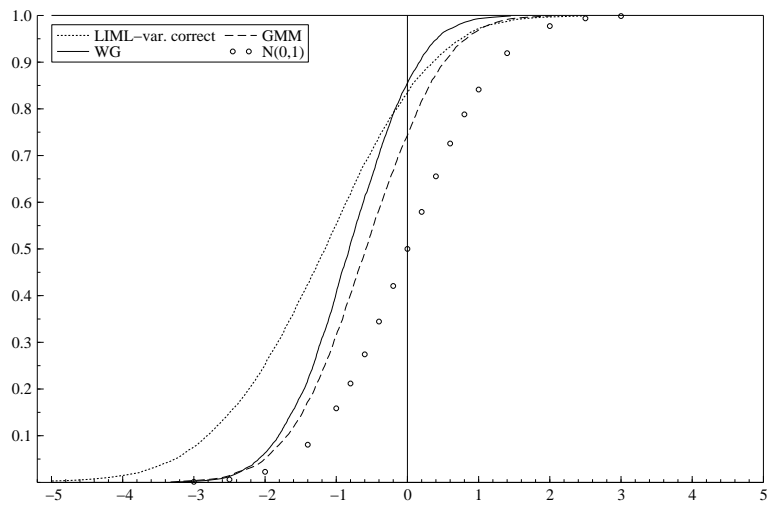
Figure 8: γ_1 (BB-model) : $N = 150$, $T = 50$, $c = 1/3$
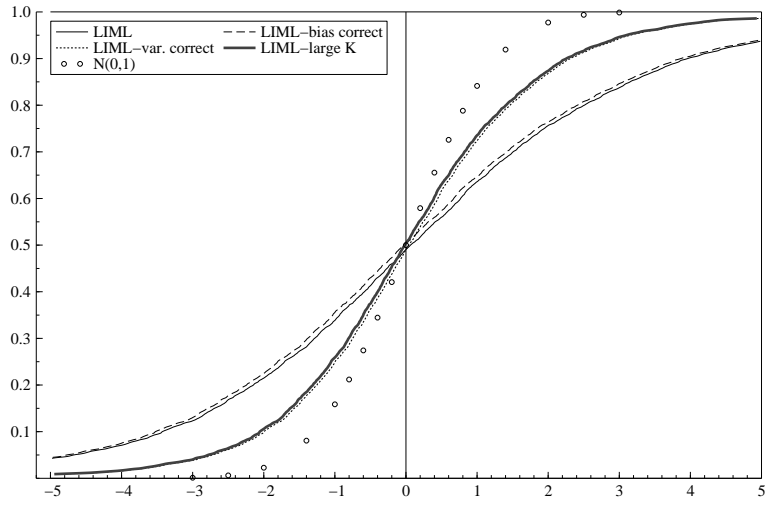
Figure 9: β_2 (Just-identified Case) : $N = 75$, $T = 25$, $c = 1/6$
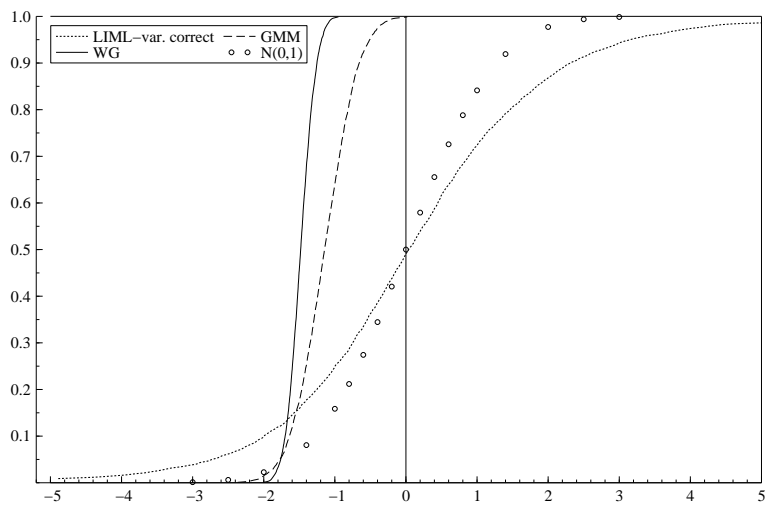
Figure 10: β_2 (Just-identified Case) : $N = 75$, $T = 25$, $c = 1/6$
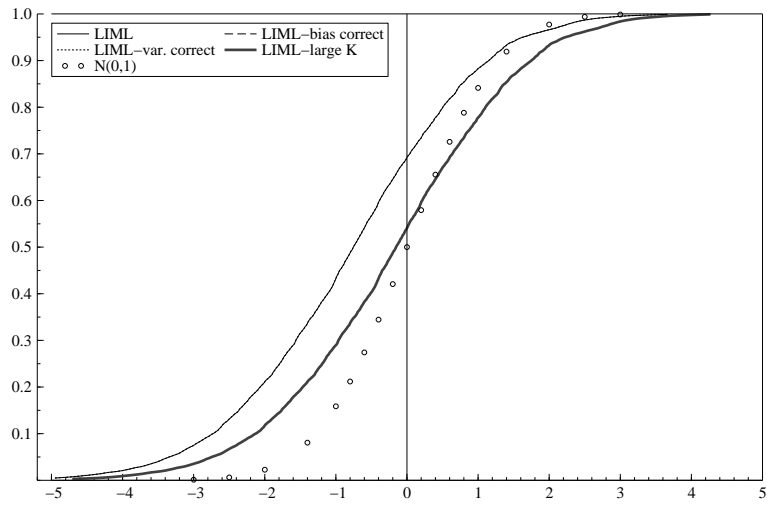
Figure 11: γ_1 (AA-model) : $N = 75$, $T = 25$, $c = 1/6$
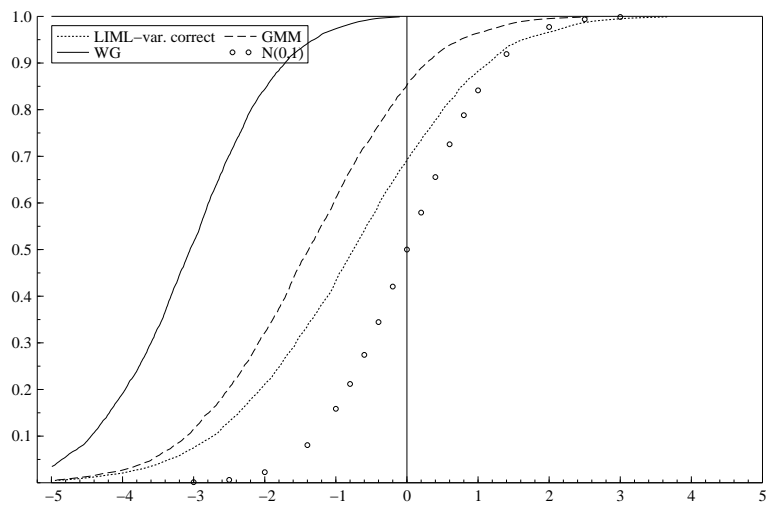
Figure 12: γ_1 (AA-model) : $N = 75$, $T = 25$, $c = 1/6$
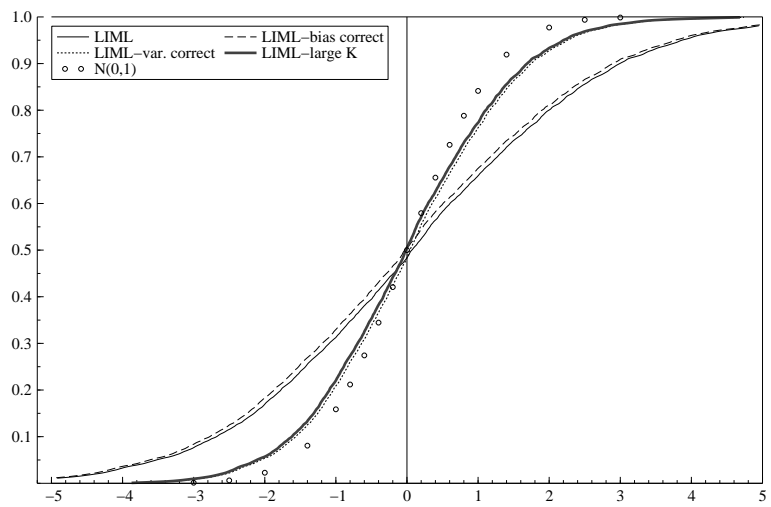
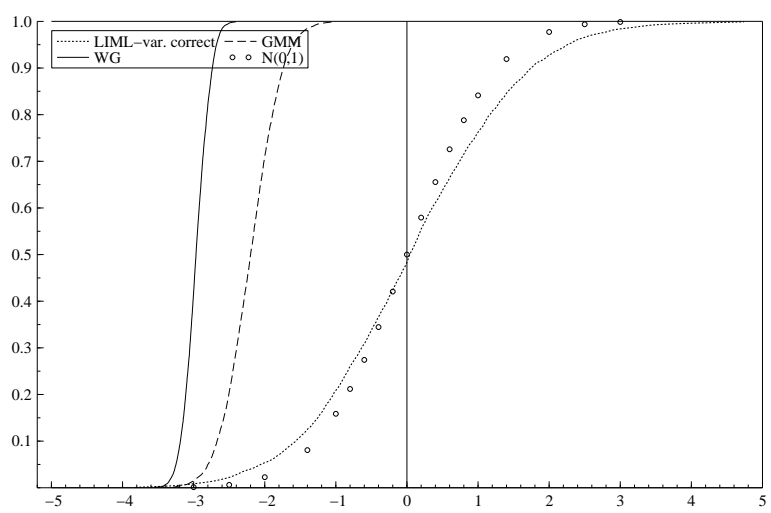Figure 13: β_2 (Just-identified Case) : $N = 150,\ T = 50,\ c = 1/6$

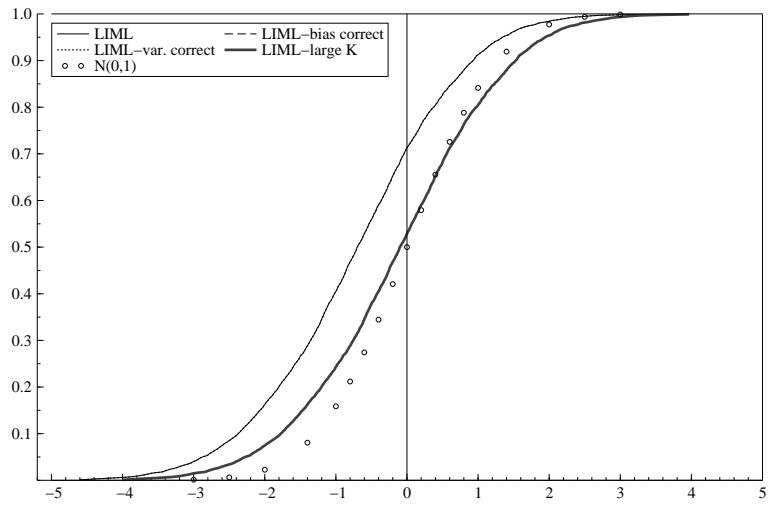Figure 14: β_2 (Just-identified Case) : $N = 150,\ T = 50\ ,c = 1/6$
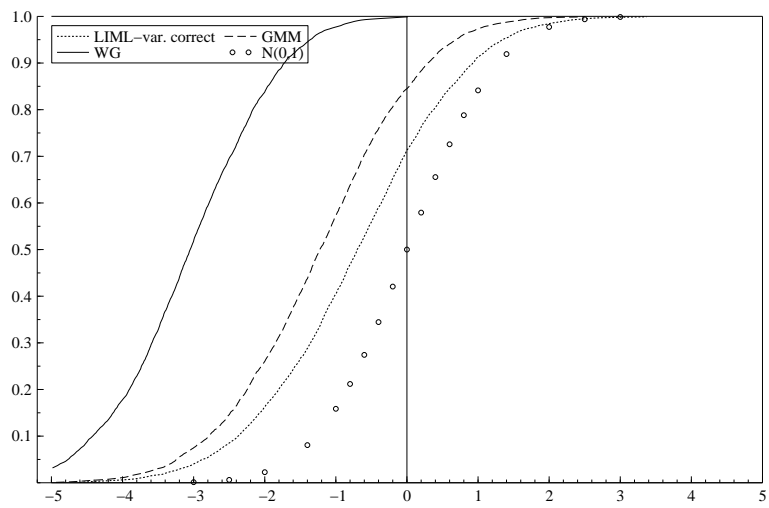
Figure 15: γ_1 (AA-model) : $N = 150$, $T = 50$, $c = 1/6$

Figure 16: γ_1 (AA-model) : $N = 150$, $T = 50$, $c = 1/6$