Supplemental Appendix for "On Standard Inference for GMM with Seeming Local Identification Failure"

Ji Hyung Lee * Zhipeng Liao†

December, 2014

Abstract

This supplemental appendix provides some auxiliary materials for "On Standard Inference for GMM with Seeming Local Identification Failure" (Lee and Liao, 2014; cited as LL in this appendix).

1 GMM Inference in Common CH Factor Model

In this section we investigate new GMM estimation and tests for the common CH features proposed in LL. Following Dovonon and Renault (2013) (cited as DR hereafter) and LL, we explicitly use parameter space $\Theta^*$ such that

$$\theta^* = \left( \theta_1, \ldots, \theta_{n-1}, 1 - \sum_{i=1}^{n-1} \theta_i \right)' = \left( \theta', 1 - \sum_{i=1}^{n-1} \theta_i \right)' = G_2 \theta + l_n, \quad (1.1)$$

for any $\theta$ in some parameter space $\Theta \subset \mathbb{R}^p$, where $p = n - 1$ and

$$G_2 = \begin{pmatrix}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1 \\
-1 & -1 & \cdots & -1
\end{pmatrix}_{n \times (n-1)}$$

$$= \begin{pmatrix}
I_p \\
-1_{1 \times p}
\end{pmatrix}_{n \times (n-1)}$$

where $I_p$ is a $p \times p$ identity matrix and $1_{1 \times p}$ is an $1 \times p$ (row) vector of ones.

To observe the moment conditions in (2.4) of LL for the common CH model, we write

$$m_t(\theta) = \begin{bmatrix}
\psi_t(\theta) \\
g_t(\theta)
\end{bmatrix} = \begin{bmatrix}
(z_t - \mu_z) \theta^*_s \left[ Y_{t+1} Y_{t+1}' - E(Y_{t+1} Y_{t+1}') \right] \theta_s \\
((z_t - \mu_z) \otimes I_p) G_2' \left[ Y_{t+1} Y_{t+1}' - E(Y_{t+1} Y_{t+1}') \right] \theta_s
\end{bmatrix}, \quad (1.2)$$

*Department of Economics, University of Washington, 336 Savery Hall, Box 353330, Seattle, WA 98195. Email: jihyung2@uw.edu
†Department of Economics, UC Los Angeles, 8379 Bunche Hall, Mail Stop: 147703, Los Angeles, CA 90095. Email: zhipeng.liao@econ.ucla.edu
where \( z_t \) is an \( H \times 1 \) vector of instrumental variables and \( \mu_z \) denotes its mean. The existence of the common CH features implies

\[
E [ m_t (\theta_0) ] = 0 \text{ for some } \theta_0 \in \Theta.
\]

(1.3)

By definition,

\[
H = \frac{\partial E [ g_t (\theta) ]}{\partial \theta} = E \left[ ((z_t - \mu_z) \otimes I_p) G'_2 Y_{t+1} Y'_{t+1} - E(Y_{t+1} Y'_{t+1})) G_2 \right]
\]

(1.4)

which implies that \( H \) is an \( H_p \times p \) real matrix and does not depend on \( \theta \). To introduce the feasible moment conditions, we replace the nuisance parameter \( \mu_z \) by its consistent estimator \( \tilde{z} = T^{-1} \sum_{t=1}^{T} z_t \), then

\[
\tilde{m}_t (\theta) = \begin{bmatrix} \tilde{\psi}_t (\theta) \\ \tilde{g}_t (\theta) \end{bmatrix} = \begin{bmatrix} (z_t - \tilde{z}) (\theta' Y_{t+1} Y'_{t+1} \theta) \\ (z_t - \tilde{z}) \otimes I_p) G'_2 Y_{t+1} Y'_{t+1} \theta \end{bmatrix}.
\]

(1.5)

By definition,

\[
H_T = T^{-1} \sum_{t=1}^{T} \frac{\partial g_t (\theta)}{\partial \theta} = T^{-1} \sum_{t=1}^{T} ((z_t - \tilde{z}) \otimes I_p) G'_2 Y_{t+1} Y'_{t+1} G_2.
\]

(1.6)

We first show the consistency of \( H_T \).

**Lemma 1.1** Under Assumption 3.4 in LL, we have \( H_T - H = O_p(T^{-\frac{1}{2}}) \).

**Proof of Lemma 1.1** It is clear that

\[
H_T = T^{-1} \sum_{t=1}^{T} ((z_t - \tilde{z}) \otimes I_p) G'_2 Y_{t+1} Y'_{t+1} G_2
\]

\[
= T^{-1} \sum_{t=1}^{T} ((z_t - \mu_z) \otimes I_p) G'_2 (Y_{t+1} Y'_{t+1} - E(Y_{t+1} Y'_{t+1})) G_2
\]

\[
- ((\tilde{z} - \mu_z) \otimes I_p) G'_2 \frac{\sum_{t=1}^{T} [Y_{t+1} Y'_{t+1} - E(Y_{t+1} Y'_{t+1})]}{T} G_2
\]

\[
= T^{-1} \sum_{t=1}^{T} ((z_t - \mu_z) \otimes I_p) G'_2 (Y_{t+1} Y'_{t+1} - E(Y_{t+1} Y'_{t+1})) G_2 + O_p(T^{-1})
\]

\[
= E \left[ ((z_t - \mu_z) \otimes I_p) G'_2 (Y_{t+1} Y'_{t+1} - E(Y_{t+1} Y'_{t+1})) G_2 \right] + O_p(T^{-\frac{1}{2}})
\]

where the third and the last equalities are by Assumption 3.4 in LL. ■

Let \( m(\theta) = E[m_t(\theta)] \) for any \( \theta \in \Theta \). The next corollary is useful for the asymptotic theory developed in the next subsections.

**Corollary 1.1** Under Assumption 3.4 in LL and the compactness of \( \Theta \), we have
(i) $\sup_{\theta \in \Theta} \left[ T^{-1} \sum_{t=1}^{T} \tilde{m}_t(\theta) - m(\theta) \right] = O_p(T^{-\frac{1}{2}})$;

(ii) $\sup_{\theta \in \Theta} \left[ T^{-1} \sum_{t=1}^{T} \frac{\partial \tilde{m}_t(\theta)}{\partial \theta} - \frac{\partial m(\theta)}{\partial \theta} \right] = O_p(T^{-\frac{1}{2}})$;

(iii) $m(\theta)$ and $\frac{\partial m(\theta)}{\partial \theta}$ are continuous in $\theta$;

(iv) $T^{-\frac{1}{2}} \sum_{t=1}^{T} \tilde{m}_t(\theta_0) \rightarrow_d N(0, \Omega_m)$, where $\Omega_m = \lim_{T \rightarrow \infty} \text{Var}\left[T^{-\frac{1}{2}} \sum_{t=1}^{T} m_t(\theta_0)\right]$.

**Proof of Corollary [1.1]** First, we note that

$$\sum_{t=1}^{T} \hat{\psi}_t(\theta) = \sum_{t=1}^{T} \left[ (z_t - \bar{z}) \theta^t Y_{t+1} Y_{t+1} \theta_s \right] \sqrt{T} \quad (1.7)$$

For $h = 1, \ldots, H$, we use $z_{h,t}$ to denote the $h$-th component in $z_t$, $\bar{z}_h$ and $\mu_{h,z}$ to denote its sample average and population mean respectively. Then we can write

$$\sum_{t=1}^{T} \left[ (z_{h,t} - \bar{z}_h) \theta^t Y_{t+1} Y_{t+1} \theta_s \right] \sqrt{T} \quad (1.8)$$

where

$$\sum_{t=1}^{T} \left[ (z_{h,t} - \bar{z}_h) Y_{t+1} Y_{t+1} \right] \sqrt{T} = \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) \left[ Y_{t+1} Y_{t+1} - E(Y_{t+1} Y_{t+1}) \right]$$

$$- (\bar{z}_h - \mu_{h,z}) \sum_{t=1}^{T} Y_{t+1} Y_{t+1} - E(Y_{t+1} Y_{t+1}) \right] \sqrt{T}$$

$$= \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) \left[ Y_{t+1} Y_{t+1} - E(Y_{t+1} Y_{t+1}) \right] + O_p(T^{-\frac{1}{2}}) \quad (1.9)$$

by Assumption 3.4 in LL. Using (1.8), (1.9) and the compactness assumption on $\Theta$, we have

$$\sum_{t=1}^{T} \left[ (z_{h,t} - \bar{z}_h) \theta^t Y_{t+1} Y_{t+1} \theta_s \right] \sqrt{T} \quad (1.10)$$

uniformly over $\Theta$ for any $h = 1, \ldots, H$, which together with (1.7), the fact that $H$ is a fixed integer and the definition of $\hat{\psi}_t(\theta)$ implies that

$$\sum_{t=1}^{T} \hat{\psi}_t(\theta) \sqrt{T} = \sum_{t=1}^{T} \psi_t(\theta) \sqrt{T} + O_p(T^{-\frac{1}{2}})$$

uniformly over $\theta \in \Theta$. By the same arguments,

$$\sum_{t=1}^{T} \hat{g}_t(\theta) \sqrt{T} = \sum_{t=1}^{T} \left[ (z_t - \bar{z}) \otimes I_p \right] G_2 Y_{t+1} Y_{t+1} \theta_s$$

$$= \sum_{t=1}^{T} \left[ (z_t - \mu_z) \otimes I_p \right] G_2 \left\{ Y_{t+1} Y_{t+1} \theta_s - E[Y_{t+1} Y_{t+1} \theta_s] \right\} + O_p(T^{-\frac{1}{2}})$$

$$= \sum_{t=1}^{T} g_t(\theta) \sqrt{T} + O_p(T^{-\frac{1}{2}})$$

(1.11)
uniformly over $\theta \in \Theta$.

(i) Using (1.10) and (1.11), we have

\[
T^{-1} \sum_{t=1}^{T} \left[ \frac{\hat{\psi}_t (\theta)}{\hat{g}_t (\theta)} \right] = \left[ \frac{T^{-1} \sum_{t=1}^{T} \hat{\psi}_t (\theta)}{T^{-1} \sum_{t=1}^{T} \hat{g}_t (\theta)} \right] + O_p \left( T^{-1} \right), \tag{1.12}
\]

uniformly over $\theta \in \Theta$. For any $h = 1, \ldots, H$,

\[
\sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) \left\{ (\theta'_t Y_{t+1})^2 - E \left[ (\theta'_t Y_{t+1})^2 \right] \right\} = T \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) \frac{Y_{t+1} Y_{t+1}' - E(Y_{t+1} Y_{t+1}')}{\theta_*} \tag{1.13}
\]

uniformly over $\theta \in \Theta$, where the second equality is by Assumption 3.4 in LL and the compactness of $\Theta$. (1.13) implies that

\[
\frac{\sum_{t=1}^{T} \hat{\psi}_t (\theta)}{T} = E \left[ \left( z_t - \mu_z \right) \left\{ (\theta'_t Y_{t+1})^2 - E \left[ (\theta'_t Y_{t+1})^2 \right] \right\} \right] + O_p \left( T^{-\frac{1}{2}} \right)
\]

uniformly over $\theta \in \Theta$. By the same arguments, we can show that

\[
\frac{\sum_{t=1}^{T} g_t (\theta)}{T} = \frac{\sum_{t=1}^{T} ((z_t - \mu_z) \otimes I_p) G_2' \left\{ Y_{t+1} Y_{t+1}' - E[Y_{t+1} Y_{t+1}'] \right\}}{T} \theta_*
\]

uniformly over $\theta \in \Theta$. Collecting the results in (1.12), (1.14) and (1.15), we get

\[
T^{-1} \sum_{t=1}^{T} \left[ \frac{\hat{\psi}_{t,T} (\theta)}{\hat{g}_{t,T} (\theta)} \right] = \begin{bmatrix} \psi (\theta) \\ g (\theta) \end{bmatrix} + O_p \left( T^{-\frac{1}{2}} \right) \tag{1.16}
\]

uniformly over $\theta \in \Theta$. This proves the claim in (i).

(ii) By definition,

\[
T^{-1} \sum_{t=1}^{T} \frac{\partial \hat{\mu}_t (\theta)}{\partial \theta'} = \begin{bmatrix} \sum_{t=1}^{T} (z_{t-\varepsilon}) \theta'_t Y_{t+1} Y_{t+1}' \\ \sum_{t=1}^{T} ((z_{t-\varepsilon}) \otimes I_p) G_2 Y_{t+1} Y_{t+1}' G_2 \end{bmatrix} = \begin{bmatrix} \sum_{t=1}^{T} (z_{t-\varepsilon}) \theta'_t Y_{t+1} Y_{t+1}' \\ \mathbb{H}_T \end{bmatrix}, \tag{1.17}
\]

where we have proved that $\mathbb{H}_T = \mathbb{H} + O_p \left( T^{-\frac{1}{2}} \right)$, and both $\mathbb{H}_T$ and $\mathbb{H}$ do not depend on $\theta$. Hence,
it is sufficient to show that
\[
\frac{1}{T} \sum_{t=1}^{T} [(z_t - \bar{z}) \theta'_s Y_{t+1} Y'_{t+1}] = E \left[(z_t - \mu_z) \theta'_s Y_{t+1} Y'_{t+1}\right] + O_p(T^{-\frac{1}{2}}) \tag{1.18}
\]
uniformly over \( \theta \in \Theta \). Note that
\[
\frac{1}{T} \sum_{t=1}^{T} (z_t - \bar{z}) \theta'_s Y_{t+1} Y'_{t+1} = \frac{1}{T} \sum_{t=1}^{T} (z_t - \mu_z) \theta'_s Y_{t+1} Y'_{t+1} - \frac{1}{T} \sum_{t=1}^{T} \theta'_s Y_{t+1} Y'_{t+1} \tag{1.19}
\]
for any \( \theta \in \Theta \). Under Assumption 3.4 in LL,
\[
\bar{z} - \mu_z = O_p(T^{-\frac{1}{2}}) \text{ and } T^{-1} \sum_{t=1}^{T} Y_{t+1} Y'_{t+1} = E [Y_{t+1} Y'_{t+1}] + O_p(T^{-\frac{1}{2}}) \tag{1.20}
\]
which combined with the compactness of \( \Theta \) implies
\[
\sup_{\theta \in \Theta} \frac{(\bar{z} - \mu_z) \theta'_s \sum_{t=1}^{T} Y_{t+1} Y'_{t+1}}{T} = O_p(T^{-\frac{1}{2}}). \tag{1.21}
\]
For any \( h = 1, \ldots, H \),
\[
\frac{1}{T} \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) \theta'_s Y_{t+1} Y'_{t+1} = \theta'_s \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) Y_{t+1} Y'_{t+1}. \tag{1.22}
\]
Under Assumption 3.4 in LL,
\[
\frac{1}{T} \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) Y_{t+1} Y'_{t+1} = E [(z_{h,t} - \mu_{h,z}) Y_{t+1} Y'_{t+1}] + O_p(T^{-\frac{1}{2}}). \tag{1.23}
\]
Using (1.22), (1.23) and the compactness of \( \Theta \), we have
\[
\frac{1}{T} \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) \theta'_s Y_{t+1} Y'_{t+1} = E [(z_{h,t} - \mu_{h,z}) \theta'_s Y_{t+1} Y'_{t+1}] + O_p(T^{-\frac{1}{2}}) \tag{1.24}
\]
uniformly over \( \theta \in \Theta \). Collecting the results in (1.19), (1.21) and (1.24), we immediately get (1.18).

(iii) By definition,
\[
\psi(\theta) = E \left[ (z_t - \mu_z) \theta'_s [Y_{t+1} Y'_{t+1} - E(Y_{t+1} Y'_{t+1})] \theta'_s \right].
\]
For any \( h = 1, \ldots, H \),
\[
E \left[ (z_{h,t} - \mu_{h,z}) \theta'_s [Y_{t+1} Y'_{t+1} - E(Y_{t+1} Y'_{t+1})] \theta'_s \right] = \theta'_s Cov(z_{h,t}, Y_{t+1} Y'_{t+1}) \theta'_s
\]
where \( Cov(z_{h,t}, Y_{t+1} Y'_{t+1}) \) is a finite real matrix under Assumption 3.4 of LL, which implies that \( \psi(\theta) \) and \( \frac{\partial \psi(\theta)}{\partial \theta^r} \) are continuous in \( \theta \). By the definition of \( m(\theta) \), we know that \( m(\theta) \) is continuous in
Moreover, \( \frac{\partial g(\theta)}{\partial \theta'} = \mathbb{H} \) which does not depend on \( \theta \). Hence we know that \( \frac{\partial m(\theta)}{\partial \theta'} \) is also continuous in \( \theta \).

(iv) Using \[1.12\], we have

\[
T^{-\frac{1}{2}} \sum_{t=1}^{T} \tilde{m}_t(\theta_0) = \begin{bmatrix}
T^{-\frac{1}{2}} \sum_{t=1}^{T} \psi_t(\theta_0) \\
T^{-\frac{1}{2}} \sum_{t=1}^{T} g_t(\theta_0)
\end{bmatrix} + O_p(T^{-\frac{1}{2}}) = T^{-\frac{1}{2}} \sum_{t=1}^{T} m_t(\theta_0) + O_p(T^{-\frac{1}{2}}) \tag{1.25}
\]

which together with \[1.3\] and Assumption 3.4 in LL implies the claim in (iv).

The asymptotic variance covariance matrix \( \Omega_m \) can be partitioned as

\[
\Omega_m = \begin{pmatrix}
\Omega_{\psi} & \Omega_{\psi g} \\
\Omega_{g\psi} & \Omega_{g}
\end{pmatrix},
\]

where

\[
\Omega_{\psi} = \lim_{T \to \infty} \text{Var} \left[ T^{-\frac{1}{2}} \sum_{t=1}^{T} \psi_t(\theta_0) \right], \quad \Omega_{g} = \lim_{T \to \infty} \text{Var} \left[ T^{-\frac{1}{2}} \sum_{t=1}^{T} g_t(\theta_0) \right],
\]

\[
\Omega_{\psi g} = \lim_{T \to \infty} \mathbb{E} \left[ T^{-1} \sum_{t=1}^{T} \sum_{s=1}^{T} \psi_t(\theta_0) g'_s(\theta_0) \right] = \Omega'_{g\psi}.
\]

### 1.1 GMM Estimation and Limit Theory

In this subsection, the weight matrix \( W_{j,T} \) for \( j = m, g \) or \( g^* \) are known random weight matrix with probability limits \( W_j \) for \( j = m, g \) or \( g^* \), respectively, where \( W_j \) are positive definite matrices. Recall that the stacked/full GMM estimator is defined as

\[
\hat{\theta}_{m,T} = \arg \min_{\theta \in \Theta} \left[ T^{-1} \sum_{t=1}^{T} \tilde{m}_t(\theta) \right]' W_m T^{-1} \sum_{t=1}^{T} \tilde{m}_t(\theta)
\]

where \( W_{m,T} \) is an \( H(p+1) \times H(p+1) \) weight matrix.

Note that Assumptions 2.1.(i) and A.1 of LL have been verified in Corollary \[1.1\]. We assume the consistency of \( W_{m,T} \), which implies that Assumption 2.1.(ii) in LL holds. Moreover, we have shown that \( \mathbb{H} \) has full rank in Lemma 3.2 of LL. Hence, the following result is directly from Proposition 2.1(i) in LL.

**Theorem 1.1** Under Assumptions 3.1-3.6 in LL, the GMM estimator \( \hat{\theta}_{m,T} \) satisfies

\[
T^{\frac{1}{2}} (\hat{\theta}_{m,T} - \theta_0) \to_d N(0, \Sigma_{\theta,m}).
\]

where \( \Sigma_{\theta,m} = (\mathbb{H}' W_{m,22} \mathbb{H})^{-1} M_\theta' W_m \Omega_m W_m M_\theta (\mathbb{H}' W_{m,22} \mathbb{H})^{-1} \), and \( W_{m,22} \) denotes the last \( Hp \times Hp \)
submatrix of $W_m$. Moreover, with the choice of $W_{m,T} = \hat{\Omega}_{m,T}^{-1} = \Omega_m^{-1} + o_p(1)$, we have

$$
\Sigma_{\theta,m}^{-1} = \mathbb{H}' (\Omega_m^{-1})_{22} \mathbb{H}
$$

where $(\Omega_m^{-1})_{22}$ denotes the last $Hp \times Hp$ submatrix of $\Omega_m^{-1}$.

We next consider GMM estimator using the Jacobian-based moment restrictions only, i.e.,

$$
\hat{\theta}_{g,T} = - (\mathbb{H}'_T W_{g,T} \mathbb{H}_T)^{-1} \mathbb{H}'_T W_{g,T} S_T
$$

where

$$
S_T \equiv T^{-1} \sum_{t=1}^T ((z_t - \bar{z}) \otimes I_p) G_y^2 Y_{t+1} Y_{t+1}' l_n.
$$

Intuitively, we can use the same arguments in showing Theorem 1.1 to derive the asymptotic normality for $\hat{\theta}_{g,T}$. However, the closed form expression of $\hat{\theta}_{g,T}$ makes the compactness assumption on $\Theta$ unnecessary. In the following, we give a different, yet more straightforward proof of the $\sqrt{T}$-normality of $\hat{\theta}_{g,T}$.

By the parametrization $\theta_\ast = G_2 \theta + l_n$, $\theta_\ast$ is automatically in $\mathcal{N}$ for any $\theta \in \mathbb{R}^{n-1}$. With the closed form expression of $\hat{\theta}_{g,T}$, we allow that $\Theta = \mathbb{R}^{n-1}$, which means that $\theta_0$ is in the interior of $\Theta$ and hence Assumption 3.6 in LL is automatically satisfied.

**Theorem 1.2** Under Assumptions 3.1-3.4 and Condition (3.3) in LL, the Jacobian-GMM estimator $\hat{\theta}_{g,T}$ satisfies

$$
T^{1/2} (\hat{\theta}_{g,T} - \theta_0) \rightarrow_d N(0, \Sigma_{\theta,g}).
$$

where $\Sigma_{\theta,g} \equiv (\mathbb{H}' W_g \mathbb{H})^{-1} \mathbb{H}' W_g \Omega_g W_g \mathbb{H} (\mathbb{H}' W_g \mathbb{H})^{-1}$. Moreover, with the choice of $W_{g,T} = \hat{\Omega}_{g,T}^{-1} = \Omega_g^{-1} + o_p(1)$, we have

$$
\Sigma_{\theta,g} = (\mathbb{H}' \Omega_g^{-1} \mathbb{H})^{-1}.
$$

**Proof of Theorem 1.2.** By definition, $l_n = \theta_\ast^0 - G_2 \theta_0$ which means that

$$
S_T = T^{-1} \sum_{t=1}^T ((z_t - \bar{z}) \otimes I_p) G_y^2 Y_{t+1} Y_{t+1}' \theta_\ast^0
$$

$$
- T^{-1} \sum_{t=1}^T ((z_t - \bar{z}) \otimes I_p) G_y^2 Y_{t+1} Y_{t+1}' G_2 \theta_0
$$

$$
= - \mathbb{H}_T \theta_0 + T^{-1} \sum_{t=1}^T \tilde{g}_t(\theta_0).
$$

(1.29)
It is clear that
\[
\frac{\sum_{t=1}^{T} \tilde{g}_t(\theta_0)}{T} = \frac{\sum_{t=1}^{T} ((z_t - \bar{z}) \otimes I_p) G'_2 Y_{t+1} Y'_{t+1} \theta_0^0}{T} - \frac{((\bar{z} - \mu_z) \otimes I_p) \sum_{t=1}^{T} G'_2 Y_{t+1} Y'_{t+1} \theta_0^0}{T} + O_p \left(\frac{1}{T^{\frac{1}{2}}}\right) \tag{1.30}
\]
where the third equality is by Assumption 3.4 in LL. \(\text{(1.30)}\) implies that
\[
\frac{\sum_{t=1}^{T} \tilde{g}_t(\theta_0)}{\sqrt{T}} = \frac{\sum_{t=1}^{T} g_t(\theta_0)}{\sqrt{T}} + O_p \left(\frac{1}{T^{\frac{1}{2}}}\right) \tag{1.31}
\]
Using Lemma 1.1, the consistency of \(W_{g,T}\), \(\text{(1.31)}\) and \(T^{-1/2} \sum_{t=1}^{T} g_t(\theta_0) = O_p(1)\), we have
\[
T^{1/2} \left(\hat{\theta}_{g,T} - \theta_0\right) = -\left(\mathbb{H}'_T W_{g,T} \mathbb{H}_T\right)^{-1} \mathbb{H}'_T W_{g,T} \left[T^{-1/2} \sum_{t=1}^{T} \tilde{g}_t(\theta_0)\right]
\]
\[
= -(\mathbb{H}'_T W_{g,T})^{-1} \mathbb{H}'_T W_{g} \left[T^{-1/2} \sum_{t=1}^{T} g_t(\theta_0)\right] + o_p(1) \tag{1.32}
\]
\[
\xrightarrow{d} N(0, \Sigma_{\theta,g})
\]
where \(\mathbb{H}'_T W_{g,T} \mathbb{H}\) is invertible as \(\mathbb{H}\) has full rank (which is proved in Lemma 3.2 of LL) and \(W_g\) is a positive definite matrix, the weak convergence is by Assumption 3.4 in LL and the Continuous Mapping Theorem (CMT). The second claim is trivial so is omitted.

The GMM estimator \(\hat{\theta}_{g^*,T}\) based on the modified moment functions is defined as
\[
\hat{\theta}_{g^*,T} = -\left(\mathbb{H}'_T W_{g^*,T} \mathbb{H}_T\right)^{-1} \mathbb{H}'_T W_{g^*,T} \left(S_T - \mathcal{F}_T\right) \tag{1.33}
\]
where \(\mathcal{F}_T = \hat{\Omega}_{g,T}^{-1} \hat{\Omega}_{g,T}^{-1} \hat{\kappa}_T\), \(\hat{\Omega}_{g,T}\) and \(\hat{\Omega}_{\psi,T}\) are consistent estimators of \(\Omega_{g,\psi}\) and \(\Omega_{\psi}\) respectively,
\[
\hat{\kappa}_T = T^{-1} \sum_{t=1}^{T} (z_t - \bar{z}) \left(\frac{G_2}{G_2 \hat{\theta}_{g,T} + l_n} \right)' Y_{t+1} Y'_{t+1} \left(\frac{G_2 \hat{\theta}_{g,T} + l_n}{G_2 \hat{\theta}_{g,T} + l_n}\right). \tag{1.34}
\]

We next present the limit theory of the GMM estimator \(\hat{\theta}_{g^*,T}\).

**Theorem 1.3** Under Assumptions 3.1-3.4 and Condition (3.3) in LL, the modified GMM estima-
tor \( \hat{\theta}_{g,T} \) satisfies
\[
T^{1/2} \left( \hat{\theta}_{g,T} - \theta_0 \right) \rightarrow_d N \left( 0, (\mathbb{H}' (\Omega_m^{-1})_{22} \mathbb{H})^{-1} \right).
\]
with the choice of \( W_{g,T} = (\Omega_m^{-1})_{22} + o_p(1) \), where \((\Omega_m^{-1})_{22}\) denotes the last \(pH \times pH\) submatrix of \(\Omega_m^{-1}\).

**Proof of Theorem 1.3.** By definition
\[
A_T = T^{-1} \sum_{t=1}^{T} (z_t - \mu_z) \left[ \left( G_2 \hat{\theta}_{g,T} + l_n \right)' \left( Y_{t+1} Y_{t+1}' + G_2 \hat{\theta}_{g,T} + l_n \right) \right] - (\bar{z} - \mu_z) \left( G_2 \hat{\theta}_{g,T} + l_n \right)' \frac{\sum_{t=1}^{T} Y_{t+1} Y_{t+1}'}{T} \left( G_2 \hat{\theta}_{g,T} + l_n \right).
\]
(1.35)

It is clear that
\[
(\bar{z} - \mu_z) \left( G_2 \hat{\theta}_{g,T} + l_n \right)' \frac{\sum_{t=1}^{T} Y_{t+1} Y_{t+1}'}{T} \left( G_2 \hat{\theta}_{g,T} + l_n \right) = (\bar{z} - \mu_z) \left( G_2 \theta_0 + l_n \right)' \left[ Y_{t+1} Y_{t+1}' + G_2 \theta_0 + l_n \right] + O_p(T^{-1}) = (\bar{z} - \mu_z) \theta_0^\prime \left[ Y_{t+1} Y_{t+1}' \right] \theta_0 + O_p(T^{-1}) = (\bar{z} - \mu_z) \theta_0^\prime \left[ Y_{t+1} Y_{t+1}' \right] \theta_0 + O_p(T^{-1})
\]
(1.36)

where the first and second equalities are by Assumption 3.4 in LL and the \( \sqrt{T} \)-consistency of \( \hat{\theta}_{g,T} \). For any \( h = 1, \ldots, H \),
\[
T^{-1} \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) \left[ \left( G_2 \hat{\theta}_{g,T} + l_n \right)' \left( Y_{t+1} Y_{t+1}' + G_2 \hat{\theta}_{g,T} + l_n \right) \right] = (G_2 \hat{\theta}_{g,T} + l_n)' \left[ \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) Y_{t+1} Y_{t+1}' \right] \left( G_2 \hat{\theta}_{g,T} + l_n \right) = (G_2 \theta_0 + l_n)' \left[ \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) Y_{t+1} Y_{t+1}' \right] \left( G_2 \theta_0 + l_n \right) + 2(\hat{\theta}_{g,T} - \theta_0) G_2' \left[ \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) Y_{t+1} Y_{t+1}' \right] \left( G_2 \theta_0 + l_n \right) + (\hat{\theta}_{g,T} - \theta_0) G_2' \left[ \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) Y_{t+1} Y_{t+1}' \right] G_2 \left( \hat{\theta}_{g,T} - \theta_0 \right) = \theta_0^\prime \left[ \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) Y_{t+1} Y_{t+1}' \right] \theta_0 + O_p(T^{-1})
\]
(1.37)

where the last equality is by the \( \sqrt{T} \)-consistency of \( \hat{\theta}_{g,T} \), Assumption 3.4 and
\[
\sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) Y_{t+1} Y_{t+1}' \left( G_2 \theta_0 + l_n \right) = \sum_{t=1}^{T} (z_{h,t} - \mu_{h,z}) Y_{t+1} Y_{t+1}' \theta_0 + O_p(T^{-\frac{1}{2}}).
\]

9
Combining the results in (1.35), (1.36) and (1.37), we get

$$
\hat{\Theta}_T = \frac{\sum_{t=1}^{T} (z_t - \mu_z) \theta_{0,*}^T \left[ Y_{t+1} Y_{t+1}' - E(Y_{t+1} Y_{t+1}') \right]}{T} \theta_{0,*} + O_p(T^{-1})
$$

$$
= \frac{1}{T} \sum_{t=1}^{T} \psi_t(\theta_0) + O_p(T^{-1}).
$$

(1.38)

As $\frac{1}{T} \sum_{t=1}^{T} \psi_t(\theta_0) = O_p(T^{-\frac{1}{2}})$ and $\hat{\Psi}_{\psi, T}, \hat{\Omega}_{\psi, T}$ are consistent, we have

$$
F_T = \hat{\Omega}_{g \psi, T}^{\psi, T} \hat{\Omega}_{\psi, T}^{-1} \hat{\Psi}_{\psi, T}^{\psi, T} \hat{\Theta}_T = \frac{\Omega_{g \psi, T} \Omega_{\psi}^{-1}}{T} \sum_{t=1}^{T} \psi_t(\theta_0) + o_p(T^{-\frac{1}{2}}),
$$

(1.39)

which together with (1.29) and (1.31) implies that

$$
S_T - F_T = -H_T \theta_0 + \frac{1}{T} \sum_{t=1}^{T} \left[ g_t(\theta_0) - \Omega_{g \psi, T} \Omega_{\psi}^{-1} \psi_t(\theta_0) \right] + o_p(T^{-\frac{1}{2}}).
$$

(1.40)

Using (1.35) and (1.40), we get

$$
\tilde{\theta}_{g^*, T} = \theta_0 - \frac{(H_T^{\prime} W_{g^*, T} H_T)^{-1} H_T^{\prime} W_{g^*, T}}{T} \sum_{t=1}^{T} \left[ g_t(\theta_0) - \Omega_{g \psi, T} \Omega_{\psi}^{-1} \psi_t(\theta_0) \right] + o_p(T^{-\frac{1}{2}})
$$

$$
= \theta_0 - \frac{(H_T^{\prime} W_{g^*} H_T)^{-1} H_T^{\prime} W_{g^*}}{T} \sum_{t=1}^{T} \left[ g_t(\theta_0) - \Omega_{g \psi, T} \Omega_{\psi}^{-1} \psi_t(\theta_0) \right] + o_p(T^{-\frac{1}{2}})
$$

(1.41)

where the second equality is by the consistency of $W_{g^*, T}$, Lemma 1.1 and

$$
T^{-1} \sum_{t=1}^{T} \left[ g_t(\theta_0) - \Omega_{g \psi, T} \Omega_{\psi}^{-1} \psi_t(\theta_0) \right] = O_p(T^{-\frac{1}{2}}).
$$

From (1.41), we deduce that

$$
\sqrt{T}(\tilde{\theta}_{g^*, T} - \theta_0) = -\frac{(H_T^{\prime} W_{g^*} H_T)^{-1} H_T^{\prime} W_{g^*}}{T} \sum_{t=1}^{T} \left[ g_t(\theta_0) - \Omega_{g \psi, T} \Omega_{\psi}^{-1} \psi_t(\theta_0) \right] + o_p(1)
$$

$$
\rightarrow_d N\left(0, \left(\Omega_{g^*}^{-1} H_T^{\prime} H_T \Omega_{g^*}^{-1}\right)^{-1}\right)
$$

where the weak convergence is by Assumption 3.4 in LL and CMT. ■
1.2 Over-identification Tests and Limit Theory

Three over-identification tests are considered in LL. The first test is based on the J-test

\[ J_{m,T} = \left[ T^{-1/2} \sum_{t=1}^{T} \hat{m}_t(\hat{\theta}_{m,T}) \right]' \hat{\Omega}_m^{-1} \left[ T^{-1/2} \sum_{t=1}^{T} \hat{m}_t(\hat{\theta}_{m,T}) \right], \]  

(1.42)

where \( \hat{\Omega}_m \) is the consistent estimator of \( \Omega_m \), which tests the validity of the stacked moment conditions in (3.9) of LL. The second test is based on the J-test

\[ J_{g,T} = \left[ T^{-1/2} \sum_{t=1}^{T} \hat{g}_t(\hat{\theta}_{g,T}) \right]' \hat{\Omega}_g^{-1} \left[ T^{-1/2} \sum_{t=1}^{T} \hat{g}_t(\hat{\theta}_{g,T}) \right], \]  

(1.43)

where \( \hat{\Omega}_g \) is the consistent estimator of \( \Omega_g \) and

\[ \hat{\theta}_{g,T}^* = - \left( \mathbb{H}_T' \hat{\Omega}_g^{-1} \mathbb{H}_T \right)^{-1} \mathbb{H}_T' \hat{\Omega}_g^{-1} \mathbf{S}_T, \]  

which tests the validity of the Jacobian moment conditions in (3.8) of LL. The test based on

\[ J_{h,T} = \left[ T^{-1/2} \sum_{t=1}^{T} m_t(\hat{\theta}_{g,T}) \right]' W_{h,T} \left[ T^{-1/2} \sum_{t=1}^{T} m_t(\hat{\theta}_{g,T}) \right], \]  

(1.44)

is proposed to test moment conditions in (3.9) of LL, where

\[ \hat{\theta}_{g,T} = - \left( \mathbb{H}_T' \hat{W}_{g,T} \mathbb{H}_T \right)^{-1} \mathbb{H}_T' \hat{W}_{g,T} \mathbf{S}_T, \]

\( W_{g,T} = W_g + o_p(1) \) and \( W_{h,T} = W_h + o_p(1) \), \( W_g \) and \( W_h \) are positive definite matrices.

We have derived the \( \sqrt{T} \)-normality for \( \theta_{m,T} \) in Theorem 1.1, which together with Corollary 1.1 and standard arguments implies the weak convergence in (2.7) of LL, i.e.

\[ J_{m,T} \xrightarrow{d} \chi^2(\nu), \]

Similarly, the \( \sqrt{T} \)-normality of \( \hat{\theta}_{g,T} \) derived in Theorem 1.2 combined with Corollary 1.1 enables us to invoke Theorem 2.2 in LL to show that

\[ J_{h,T} \xrightarrow{d} B_{H(p+1)}' \Omega_m^{1/2} \mathbb{P} W_h \mathbb{P} \Omega_m^{1/2} B_{H(p+1)} \]

where \( \mathbb{P} = I_{H(p+1)} - \text{diag}(\mathbb{H}_T \mathbb{H}(\mathbb{H}_T' \hat{W}_g \mathbb{H})^{-1} \mathbb{H}_T' \hat{W}_g) \) and \( B_{H(p+1)} \) denotes an \( H(p+1) \times 1 \) standard normal random vector.\[ ^1\]

\[ ^1\]As we have proved the \( \sqrt{T} \)-consistency of \( \hat{\theta}_{g,T} \), the uniform approximations in Corollary 1.1(i) and (ii) are only needed in the local neighborhood of \( \theta_0 \). This means that the compactness of \( \Theta \) is not needed to show the asymptotic distribution of \( J_{h,T} \) either.

11
Theorem 1.4 Under Assumptions 3.1-3.4 and Condition (3.3) in LL, we have

\[ J_{g,T} \to_d \chi^2 (H_{p} - p). \]

Proof of Theorem 1.4. For any \( \theta \), we can write

\[
\frac{\sum_{t=1}^{T} \hat{g}(\theta)}{\sqrt{T}} = \frac{\sum_{t=1}^{T} ((z_t - \bar{z}) \otimes I_p) G'_{2} Y_{t+1} Y'_{t+1} \theta_*}{\sqrt{T}} \]

\[
= \frac{\sum_{t=1}^{T} ((z_t - \mu_z) \otimes I_p) G'_{2} Y_{t+1} Y'_{t+1} G_{2}}{\sqrt{T}} (G_{2} \theta + l_n)
\]

\[
- \frac{((\bar{z} - \mu_z) \otimes I_p) \sum_{t=1}^{T} G'_{2} Y_{t+1} Y'_{t+1}}{\sqrt{T}} (G_{2} \theta + l_n). \tag{1.45}
\]

Note that

\[
\frac{\sum_{t=1}^{T} ((z_t - \mu_z) \otimes I_p) G'_{2} Y_{t+1} Y'_{t+1}}{\sqrt{T}} (G_{2} \theta + l_n)
\]

\[
= \frac{\sum_{t=1}^{T} ((z_t - \mu_z) \otimes I_p) G'_{2} Y_{t+1} Y'_{t+1}}{\sqrt{T}} (G_{2} \theta_0 + l_n)
\]

\[
+ \frac{\sum_{t=1}^{T} ((z_t - \mu_z) \otimes I_p) G'_{2} Y_{t+1} Y'_{t+1} G_{2}}{T} \sqrt{T} (G_{2} \theta_0 - \theta_0)
\]

\[
= \frac{\sum_{t=1}^{T} ((z_t - \mu_z) \otimes I_p) G'_{2} Y_{t+1} Y'_{t+1} \theta_0}{\sqrt{T}} + \sqrt{T} (G_{2} \theta_0 - \theta_0) + O_p(T^{-\frac{1}{2}}) \tag{1.46}
\]

where the last equality is by the \( \sqrt{T} \)-consistency of \( \hat{\theta}_{g,T} \) and Lemma 1.1. Next, we have

\[
\frac{((\bar{z} - \mu_z) \otimes I_p) \sum_{t=1}^{T} G'_{2} Y_{t+1} Y'_{t+1}}{\sqrt{T}} (G_{2} \theta + l_n)
\]

\[
= \frac{((\bar{z} - \mu_z) \otimes I_p) \sum_{t=1}^{T} G'_{2} Y_{t+1} Y'_{t+1}}{\sqrt{T}} (G_{2} \theta_0 + l_n)
\]

\[
+ \frac{((\bar{z} - \mu_z) \otimes I_p) \sum_{t=1}^{T} G'_{2} Y_{t+1} Y'_{t+1} G_{2}}{T} \sqrt{T} (G_{2} \theta_0 - \theta_0)
\]

\[
= \frac{((\bar{z} - \mu_z) \otimes I_p) G_{2}' E(Y_{t+1} Y'_{t+1}) \theta_{0,*}}{\sqrt{T}} + O_p(T^{-\frac{1}{2}}) \tag{1.47}
\]

where the last equality is by the \( \sqrt{T} \)-consistency of \( \hat{\theta}_{g,T} \) and

\[
\bar{z} - \mu_z = O_p(T^{-\frac{1}{2}}) \text{ and } \sum_{t=1}^{T} \frac{Y_{t+1} Y'_{t+1}}{T} = E(Y_{t+1} Y'_{t+1}) + O_p(T^{-\frac{1}{2}}).
\]

Combining the results in (1.45), (1.46) and (1.47), we have

\[
\frac{\sum_{t=1}^{T} \hat{g}(\hat{\theta}_{g,T})}{\sqrt{T}} = \frac{\sum_{t=1}^{T} g_0(\theta_0)}{\sqrt{T}} + \sqrt{T} (\hat{\theta}_{g,T} - \theta_0) + O_p(T^{-\frac{1}{2}}), \tag{1.48}
\]

12
which together with (1.32) in the proof of Theorem 1.2 implies that

$$
\frac{\sum_{t=1}^{T} \hat{g}_t(\hat{\theta}_{g,T})}{\sqrt{T}} = \left[ \Omega_g^{\frac{1}{2}} - \mathbb{H}(\mathbb{H}'\Omega_g^{-1}\mathbb{H})^{-1}\mathbb{H}'\Omega_g^{-\frac{3}{2}} \right] \left[ \Omega_g^{-\frac{1}{2}} T^{-\frac{1}{2}} \sum_{t=1}^{T} g_t (\theta_0) \right] + o_p(1). \tag{1.49}
$$

By the consistency of $\hat{\Omega}_g$ and (1.49), we deduce that

$$
J_{g,T} = \left[ T^{-1/2} \sum_{t=1}^{T} \hat{g}_t(\hat{\theta}_{g,T}) \right] \left[ \hat{\Omega}_g^{-1} \right] \left[ T^{-1/2} \sum_{t=1}^{T} \hat{g}_t(\hat{\theta}_{g,T}) \right] = \left[ \Omega_g^{-\frac{1}{2}} T^{-\frac{1}{2}} \sum_{t=1}^{T} g_t (\theta_0) \right] \left[ \Omega_g^{-\frac{1}{2}} T^{-\frac{1}{2}} \sum_{t=1}^{T} g_t (\theta_0) \right] + o_p(1) \tag{1.50}
$$

where $\mathbb{P}_g = I_{Hp} - \Omega_g^{\frac{1}{2}} \mathbb{H}(\mathbb{H}'\Omega_g^{-1}\mathbb{H})^{-1}\mathbb{H}'\Omega_g^{-\frac{3}{2}}$ is an idempotent matrix with rank $Hp - p$ as $\mathbb{H}$ has full rank $p$. Using (1.50) and Assumption 3.4 in LL, we deduce that

$$
J_{g,T} \rightarrow_d B'_{Hp}\mathbb{P}_g B_{Hp} \sim \chi^2(Hp - p)
$$

where $B_{Hp}$ denotes an $Hp \times 1$ standard normal random vector.

References
