$t$-Test in Dynamic Panel Structural Equations\textsuperscript{*}

Kentaro Akashi

\textbf{abstract}

This paper studies the asymptotic covariance estimation associated with the LIML estimator for large-$T$ dynamic panel structural equations. The proposed simple estimation method is based on the fixed-$N$ or large $K$-asymptotics relating to the number of backward-filtered instruments. We demonstrate the resulting $t$-Test which also does not need asymptotic bias correction works well with the large-$K$ asymptotics and is robust to persistent panel data rather than a plug-in estimation method.

\textbf{Keywords} : Backward Filter, Covariance Estimation, large-$K$ asymptotics.

1 Introduction

Recently estimation procedures for dynamic panel models with possibly large $T$ dimensions have been proposed, since inconsistency of some dynamic GMM estimators based on standard panel asymptotics ($N \to \infty$, $T < \infty$) had been developed. See Allvarez and Allerano (2003), and Hayakawa (2006, 2007). Although the most of previous studies have focused on GMM type estimators for reduced form models in our context, on the other hand Kunitomo and Akashi (2008) has proposed an approaching by the Limited Information Maximum Likelihood estimator (LIML) for the dynamic panel structural equations. In this study, we consider the $t$-Test based on the doubly-filtered LIML estimator which has been proposed by Kunitomo and Akashi (2008). This estimator has no asymptotic bias by using the forward-filtering for data and the backward-filtered instruments which have been proposed by Hayakawa (2006, 2007). The backward-filtering method can avoid the large heterogeneity of individual effects and growing the number of instruments. It is noteworthy to point out that the doubly-filtered LIML estimation is applicable to fixed $N$ asymptotics ($N < \infty$, $T \to \infty$), and also the approximation of its limiting distribution is reduced to the kind of large-$K$ asymptotics. The large-$K$ asymptotic theory were early developed by Kunitomo (1980) and Morimune (1983), and recently reconsidered by Bekker (1994), and Anderson, Kunitomo and Matsushita (2008) in (weak) many instruments problem. Its

\textsuperscript{*}I am grateful to Naoto Kunitomo for his thoughtful discussion and Yukitoshi Matsushita for his careful comments.
asymptotics improves the first order approximation of LIML estimator, but has the asymptotic covariance possibly depending on higher order moments.

The main purpose of this paper is to construct the \( t \)-Test by a simple asymptotic covariance estimation of the LIML estimator based on the large-\( K \) asymptotics. The simpler covariance estimation for the LIML estimator using the all available instruments seems to be difficult, since it has the asymptotic bias depending on the reduced form coefficients, so that we shall focus on the LIML estimator using the backward-filtered instrumental variables.

This paper is organized as follows. In Section 2, we explain the dynamic panel structural equation model and the covariance estimation method. Section 3 demonstrates the some finite sample properties of the \( t \)-Test. Then concluding remarks and the mathematical derivations of our result will be in Section 4 and 5, respectively.

## 2 Model and the Covariance Estimation

### 2.1 Reduced form and Instrumental variables

The reduced form of each \((i,t), \ (i = 1, ..., N; t = 1, ..., T)\) as the Panel Vector Autoregressive (Panel VARs) model (2.1) and the extended reduced form (2.2) are given by

\[
\begin{align*}
y_{it} &= (I_{1+G_2},\ O)z_{it} = \Pi z_{i(t-1)} + (\pi_i + v_{it}), \\
z_{it} &= \Pi^* z_{i(t-1)} + (\pi^*_i + v^*_{it}),
\end{align*}
\]

where \(z_{it}, z_{i(t-1)}, \pi^*_i \) and \(v^*_{it}\) are \(K' \times 1\) vectors, \(\pi\) and \(v^*_{it}\) stand for reduced form individual effects and disturbances, respectively. We stack \((1+G_2) \times 1\) endogenous variables vector \(y^'_{it} = (y_{it}^{(1)}, ..., y_{it}^{(1+G_2)})\) in the first \((1+G_2)\) rows of \(z_{it}\), then the first structural equation is obtained by

\[
\beta'y_{it} = \gamma_1' z^{(1)}_{i(t-1)} + (\pi^{(1)}_i + u_{it}),
\]

where \(\beta' = (1, -\beta'_2)\), \(\beta_2\) is the \(G_2 \times 1\) structural parameters, and \(\gamma_1\) is the \(K_1 \times 1\) coefficients parameters, which are parameters of interests \(\theta' = (\beta'_2, \gamma'_1)\). In order to make orthogonal moment conditions, we consider the forward-filtered reduced form (2.4), and define the demeaning process \(w_{it} = z_{it} - (I_{K'} - \Pi^*)^{-1}\) as (2.5),

\[
\begin{align*}
z^{(f)}_{it} &= \Pi' z^{(f)}_{i(t-1)} + v^{(f)}_{it}, \\
w_{it} &= \Pi' w_{i(t-1)} + v^*_{it}.
\end{align*}
\]
where the each element $z^{(k,f)}_{it}$ $(k = 1, \ldots, K; t = 1, \ldots, T - 1)$ of $\mathbf{z}^{(f)}_{it}$ are obtained by pre-multiplying the $(T - 1) \times T$ forward orthogonal matrix $\mathbf{A}_f$ (c.f. Kunitomo and Akashi, 2008) to $T \times 1$ vector $\mathbf{z}^{(f)}_{iT} = (z^{(1)}_{it}, \ldots, z^{(K)}_{it})'$.

Although $z^{(t-1)}_{it}$ and $\mathbf{v}^{(f)}_{it}$ are correlated with $\pi^*_i$ and $\mathbf{v}^{*(f)}_{it}$ respectively, we call (2.1), (2.2) and (2.4) as reduced form, since $z^{(t-1)}_{it}$ is predetermined variables $E[z^{(t-1)}_{it}] = E[z^{(t-1)}_{it} \mathbf{v}^{*(f)}_{it}] = \mathbf{0}_K$ and the endogeneity caused by individual effects are substantially weakened by the forward-filtering in estimation procedure.

As for the stationary assumption is conditioned on $K' \times K'$ matrix $\Pi'$ in (2.2), and the number of instruments are decomposed as

$$K' = K_1 + K_2 + K_3,$$

(2.6)

where $K_1$ and $K_2$ ($K = K_1 + K_2$) stand for the total number of instruments included or excluded variables in the first structural equation respectively, which can include lagged endogenous variables. $K_3$ stands for the total number of the variables excluded from $(1+G_2)$ reduced forms. It is included in order to hold the vector AR(1) representation of (2.2).

At period $t$ the instrumental matrix based on backward-filtering is given by

$$\mathbf{Z}^{(b)}_{t-1} = b_t[\mathbf{Z}_{t-1} - \frac{1}{t-1}(\mathbf{Z}_{t-2} + \mathbf{0} + \mathbf{Z}_0)],$$

(2.7)

where $\mathbf{Z}^{(b)}_{t-1} = (z^{(t-1)}_{1}, \ldots, z^{(t-1)}_{N})$ is the $K \times N$ matrix and $0 < b_t \leq 1$. Then the ratio of the total number of instruments to the total sample size is given by

$$\frac{K(T-2)}{N_0(T-2)} \xrightarrow{T \to \infty} c = \frac{K}{N_0}.$$  

(2.8)

Therefore this positive ratio composes the kind of large-$K$ asymptotics under panel structural equation models provided $K < \infty$, $N_0 < \infty$, while the original large-$K$ theory consists of that $K \to \infty$, $K/N < \infty$ as $N \to \infty$. Moreover, when the order of instruments is reduced to $O(T)$, the LIML estimator does not need the double asymptotics $N, T \to \infty$, and can regards the individual number as fixed $N_0 < \infty$. The double asymptotics may worsen the approximation of the limiting distribution, since it is further approximation of fixed $T$ or fixed $N$ asymptotics.

2.2 Estimation result and Covariance estimations

The LIML estimator was originally developed by Anderson and Rubin (1949), and Kunitomo and Akashi (2008) applied it to dynamic panel structural models. The LIML estimator using the backward filtered instruments $(1, -\hat{\theta}'_{L1}) =
The components of $T_{\text{LIML}}$ result for the doubly-filtered LIML is as follow. As $M$ where

$$\begin{align*}
\Phi^* & = D'\mathcal{E}[w_{it}w_{it}']D, \\
\Xi_3' & = \left(\frac{1}{1-c}\mathcal{E}[u_{it}u_{it}']\lim_{n_0\to\infty}\frac{1}{n_0}\sum_{t=2}^{T-1}\mathcal{E}[d'tW_{t-1}]D\right), \\
\Xi_4 & = \left(\frac{1}{1-c}\right)^2\mathcal{E}[(u_{it}^2 - \sigma^2)u_{it}u_{it}']\left(\lim_{n_0\to\infty}\frac{1}{n_0}\sum_{t=2}^{T-1}\mathcal{E}[d'td_t] - c^2\right), \\
u_{it}^\perp & = (0, I_{G_2})[I_{1+G_2} - \frac{\Omega\beta\beta'}{\beta'\Omega\beta}]v_{it},
\end{align*}$$

are defined by

$$\begin{align*}
(1, -\hat{\beta}_{2L}, -\hat{\gamma}_{1L})' \text{ is defined by}
\begin{align*}
\left(\frac{1}{n_0}G^{(f)} - \lambda_{n_0}\frac{1}{q_{n_0}}H^{(f)}\right)\left(\frac{1}{-\hat{\theta}_{L}}\right) = 0,
\end{align*} 
(2.9)
\end{align*}$$

where $n_0 = N_0(T - 2)$, $q_{n_0} = n_0 - K(T - 2)$, and $\lambda_{n_0}$ is the smallest root of

$$\begin{align*}
\left|\frac{1}{n_0}G^{(f)} - \lambda_{n_0}\frac{1}{q_{n_0}}H^{(f)}\right| = 0. 
(2.10)
\end{align*}$$

The $(1 + G_2 + K_1) \times (1 + G_2 + K_1)$ matrices are given by

$$\begin{align*}
G^{(f)} & = \sum_{t=2}^{T-1} \begin{pmatrix} Y_t^{(f)} \\ Z_t^{(1,f')} \end{pmatrix} M_t(Y_t^{(f)}, Z_{t-1}^{(1,f)}), 
(2.11)
\end{align*}$$

and

$$\begin{align*}
H^{(f)} & = \sum_{t=2}^{T-1} \begin{pmatrix} Y_t^{(f)}' \\ Z_t^{(1,f)} \end{pmatrix} [I_{N_0} - M_t](Y_t^{(f)}, Z_{t-1}^{(1,f)}'), 
(2.12)
\end{align*}$$

where $M_t = Z_{t-1}^{(b)}(Z_{t-1}^{(b)}Z_{t-1}^{(b)})^{-1}Z_{t-1}^{(b)}$, $Y_t^{(f)}' = (y_{it}^{(f)}, ..., y_{N_0}^{(f)})$ and $Z_t^{(1,f)} = (z_{1(t-1)}^{(1,f)}, ..., z_{N_0(t-1)}^{(1,f)})$ are $(1 + G_2) \times N_0$ and $K_1 \times N_0$ matrices respectively, whose component vectors $(y_{it}^{(1,f)}, z_{it}^{(1,f)})$ are forward filtered.

Under the Assumptions of Kunitomo and Akashi (2008), the estimation step result for the doubly-filtered LIML is as follow. As $T \to \infty$, regardless of whether $N_0 \to \infty$ or is fixed, then

$$\begin{align*}
\sqrt{N_0(T - 2)}(\hat{\theta}_{L2} - \begin{pmatrix} \hat{\beta}_2 \\ \hat{\gamma}_1 \end{pmatrix}) \overset{d}{\to} \mathcal{N}(0, \Psi^*), 
(2.13)
\end{align*}$$

where

$$\begin{align*}
\Psi^* & = \Phi^* \left[\sigma^2\Phi^* + \begin{pmatrix} I_{G_2} \\ 0 \end{pmatrix}(c, \Omega\sigma^2 - \Omega\beta\beta'\Omega)_{22} + \Xi_4)(I_{G_2}, O) + \Xi_3 + \Xi_3'\right]\Phi^*^{-1}.
\end{align*}$$

The components of $\Psi^*$ are defined by

$$\begin{align*}
\Phi^* & = D'\mathcal{E}[w_{it}w_{it}']D, \\
\Xi_3' & = \left(\frac{1}{1-c}\mathcal{E}[u_{it}u_{it}']\lim_{n_0\to\infty}\frac{1}{n_0}\sum_{t=2}^{T-1}\mathcal{E}[d'tW_{t-1}]D\right), \\
\Xi_4 & = \left(\frac{1}{1-c}\right)^2\mathcal{E}[(u_{it}^2 - \sigma^2)u_{it}u_{it}']\left(\lim_{n_0\to\infty}\frac{1}{n_0}\sum_{t=2}^{T-1}\mathcal{E}[d'td_t] - c^2\right), \\
u_{it}^\perp & = (0, I_{G_2})[I_{1+G_2} - \frac{\Omega\beta\beta'}{\beta'\Omega\beta}]v_{it},
\end{align*}$$

as (2.14), (2.15), (2.16), and (2.17).
\[ \sigma^2 = \beta' \Omega \beta, \quad c_* = c/(1 - c), \quad \mathbf{d}_t = \text{diag}(\mathbf{M}_t) \cdot \mathbf{u}_{N_0}. \]

In the double asymptotics \( N_0, T \to \infty \), \( \Psi^* \) is simplified as \( \sigma^2 \hat{\Phi}^{-1} \) due to \( c = 0 \), and it is identical to the lower bound of a cross-sectional structural model as if we used the standard asymptotics, which means \( N_0 \) goes to infinity and \( T = 1 \), \( K < \infty \), and non existence of individual effects \( \pi_t = 0 \) in (2.1). When we take more instrumental \( K'' \) variables including the \( K \) variables \((K < K'')\), the first term \( \sigma^2 \hat{\Phi}^{-1} \) does not change since the \( K \) variables are the optimal instruments in terms of double asymptotics. But the redundant instruments makes the second term \( c_\ast \Omega = \Omega \sigma^2 - \Omega \beta' \hat{\Omega} \) be larger in the large-\( K \)-asymptotics due to \( c = K/N_0 < K''/N_0 \).

As for the consistent estimators of \( \Phi^*, \Omega, \sigma^2 \) without explicitly estimation for \( \Pi^* \) are given by

\[
\hat{\Phi}^* = (0, I_{G_2 + K_1}) \begin{bmatrix} 1/n_0 \mathbf{G}^{(f)} - c_0 \mathbf{H}^{(f)} \\ q_{n_0} \end{bmatrix} \begin{pmatrix} 0' \\ I_{G_2 + K_1} \end{pmatrix}, \quad (2.18)
\]

\[
\hat{\Omega} = (I_{1 + G_2}, O) \begin{bmatrix} 1/q_{n_0} \mathbf{H}^{(f)} \\ O \end{bmatrix} \begin{pmatrix} I_{1 + G_2} \\ O \end{pmatrix}, \quad (2.19)
\]

and \( \hat{\sigma}^2 = \hat{\beta}' \hat{\Omega} \hat{\beta} \), \( \hat{\beta} = (1, -\hat{\beta}'_2) \), where \( c_0 = K/N_0 \). The essentially same construction are used by Bekker (1994) and Matsushita (2007).

Hence, given the consistent estimators for \( \Xi_3, \Xi_4 \), we can consider the consistent estimator of \( \Psi^* \) as follow

\[
\tilde{\Psi}^* = \hat{\Phi}^* \left[ \hat{\sigma}^2 \hat{\Phi}^* + \begin{pmatrix} I_{G_2} \\ O \end{pmatrix} (c_{a_0} \hat{\Omega} \hat{\sigma}^2 - \hat{\Omega} \hat{\beta} \hat{\beta}' \hat{\Omega})_{22} + \hat{\Xi}_4) \right] (I_{G_2}, O) + \hat{\Xi}_3 + \hat{\Xi}_3' \hat{\Phi}^* \left[ \hat{\Phi}^* \right]^{-1},
\]

where \( c_{a_0} = c_0/(1 - c_0) \), then we may call this the plug-in method.

It is important to check the conditions for disappearance of the complicated components \( \Xi_3 \) and \( \Xi_4 \), the normality assumption of \( v_{it} \) is an example and the other case are discussed by Anderson et.al. (2008). Although we could consider the consistent estimator of \( \Xi_3 \) and \( \Xi_4 \), it is also meaningful to develop alternative estimator for \( \Psi^* \) estimating the higher order components automatically. This is so because it is necessary for \( \hat{\Xi}_3 \) to estimate the reduced form coefficient \( \Pi^* \) and some difficulty caused by unobservability of \( w_{it} \). For some persistent panel data, which has large autoregressive coefficients, the plug-in method can be poor, so this is the another reason why we consider the alternative.

The alternative consistent estimator of \( \Psi^* \) is defined by

\[
\hat{\Psi}^* = \hat{\Phi}^* \left[ \hat{\Phi}^* \right]^{-1} \hat{m}_0 \hat{m}' \hat{\Phi}^* \left[ \hat{\Phi}^* \right]^{-1}, \quad (2.20)
\]
\[ \hat{\Gamma}^* = (0, I_{G_2 + K_1})[I_{1+G_2+K_1} - \frac{1}{\hat{\beta}' \hat{\Omega} \hat{\beta}} \begin{pmatrix} \hat{\Omega} \hat{\beta} \\ 0 \end{pmatrix} (1, -\hat{\theta}')] \] 
\[ \hat{F}_{no} = \frac{1}{N_0(T-2)} \sum_{t=2}^{T-1} \left[ G_t^{(f)} - c_0 H_t^{(f)} \right] \left( \begin{array}{c} 1 \\ -\hat{\theta} \end{array} \right) (1, -\hat{\theta}) (G_t^{(f)} - c_0 H_t^{(f)}) \] 
\[ G_t^{(f)} = (Y_t^{(f)}, Z_t^{(1,f)})' M_t (Y_t^{(f)}, Z_t^{(1,f)}) , \] 
\[ H_t^{(f)} = (Y_t^{(f)}, Z_t^{(1,f)})' [I_{N_0} - M_t] (Y_t^{(f)}, Z_t^{(1,f)}) , \]
and we may call (2.20) the \textit{sample variance form method}, since it corresponds to the sum of squares of the sampling errors as usual way.

Although \( \lambda_{no} \) is a consistent estimator for \( c \), the parameter sequence \( c_0 = K/N_0 \) is utilized, because we have the knowledge of the limit which is \( c_0 = c \) itself, or \( c_0 \to c = 0 \) in double asymptotics. It can be considered that the correction term \( -c_0 H_t^{(f)} \) corresponds to the estimation of \( \Xi_3 \) and \( \Xi_4 \), so that the second term \( c_0 [\Omega \sigma^2 - \Omega \beta' \Omega] \) is extracted from \( G_t^{(f)} \). Also if we knew the higher order terms were negligible, we would not need to set \( c_0 \).

The assumptions of the following theorem is identical to the estimation step’s. The derivation will be given in Section 5.

\textbf{Theorem :} Let the Assumptions (A1) to (A3) of Kunitomo and Akashi (2008) hold, and \( c_0 < 1 \) under \( N_0 < \infty \), or \( 0 < \lim N_0/T < \infty \) as \( N_0 \to \infty \). Then, as \( T \to \infty \), regardless whther \( N_0 \) is fixed or tends to infinity
\[ \hat{\Psi}^* \overset{p}{\to} \Psi^* . \] 
Therefore,
\[ t_{Samp.} = \sqrt{\frac{N_0 T}{e_j' \hat{\Psi} e_j}} e_j' (\hat{\theta} - \theta) \overset{d}{\to} N(0, 1) , \] 
where \( e_j (j = 1, ..., G_2 + K_1) \) is the \( j \)-th unit vector.

From this result we may not need to regards \( N_0 \) as fixed or tending to infinity in practical.

\section{Finite Sample Properties of \( t \)-Tests}

This section presents some Monte Carlo experiments to investigate the finite sample properties of several \( t \)-statistics. We have the concern about the following comparison, (i) the \( t \)-Test based on GMM with those based on LIML estimator,
(ii) the t-Test based on Plug-in method with those based on the Sample variance form method under the persistent dynamic panel data, (iii) the number of instrumental variables with the power of test.

We use the following data generation process

\[
y^{(1)}_{it} = \beta_2 y^{(2)}_{it} + \gamma_1 y^{(1)}_{it-1} + \gamma_3 x_{it} + u_{it},
\]

\[y^{(1)}_{it} = \beta_2 y^{(2)}_{it} + \gamma_1 y^{(1)}_{it-1} + \gamma_3 x_{it} + u_{it}, \quad \text{(3.27)}\]

and the extended reduced form is given by

\[
z_{it} = \begin{pmatrix}
\gamma_1 & \beta_2 \gamma_{21} & \beta_2 \gamma_{22} & \gamma_3 & 0 \\
0 & \gamma_{21} & \gamma_{22} & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & \gamma_{31} & \gamma_{32} \\
0 & 0 & 0 & 1 & 0
\end{pmatrix}
\begin{pmatrix}
y^{(1)}_{it-1} \\
y^{(2)}_{it-1} \\
y^{(2)}_{it-2} \\
x_{it} \\
x_{it-1}
\end{pmatrix}
+ \begin{pmatrix}
v^{(1)}_{it} \\
v^{(2)}_{it} \\
v^{(1)}_{it} \\
\epsilon_{it} \\
\end{pmatrix}, \quad \text{(3.28)}
\]

where \( \beta_2 = 0, \ \gamma_{21} = \gamma_{31} = 0.25, \ \gamma_{22} = \gamma_{32} = 0.5 \) and we will control \( \gamma_1 \). The disturbances are generated by zero mean normal distribution with \( \text{Var}[v^{(g)}_{it}] = \text{Var}[\epsilon_{it}] = 1, \ \text{Corr}[v^{(1)}_{it}, v^{(2)}_{it}] = 0.25 \) for \( g = 1, 2 \).

Note that (i) the individual effects are ignored, since we use the backward filter for instruments with the forward filtering which removes its effects exactly from an estimator, (ii) \( x_{it} \) is interpreted as an independent variable due to \( \text{Corr}[\epsilon_{it}, v^{(1)}_{it}] = 0 \) and \( x_{it-1} \) is the \( K_3 \) variable \( (K = K_1 + K_2 = 2 + 2) \), (iii) by the normality we do not consider the estimation of \( \Xi_3 \) and \( \Xi_4 \) when we use the Plug-in method. The number of repetitions is 3,000 and we discard the initial 15 periods data.

Figure 1 in Appendix shows the null distributions of \( \beta_2 = 0 \) under \( \gamma_1 = 0, c_0 = 4/100 \), where \( t_{\text{Plug}} \) stands for the t-statistic using \( \hat{\Psi}^* \) and \( t_{\text{GMM}} \) is based on GMM estimator whose denominator is the same of \( t_{\text{Plug}, c=0} \), that is \( \sigma^2 \hat{\Phi}^{-1} \). The figure and Table 1 indicate the fixed \( N_0 \) or large-\( K \) asymptotics also improves the approximation of null distributions under the usage of the backward-filtered instruments. On the other hand the distributions of \( t_{\text{GMM}} \) are not centered. The main reason is considered by its asymptotic bias

\[
\text{Asy. bias of } \sqrt{N_0 T} (\hat{\theta}_{\text{GMM}} - \theta) = \sqrt{N_0 T} O(c_0) = \sqrt{\frac{T}{N_0}} O(K_1 + K_2). \quad \text{(3.29)}
\]

Therefore, it diverges as \( T \to \infty \) under \( N_0 < \infty \) and converges possibly non zero under \( 0 \leq \lim T/N_0 < \infty \).

As for the persistent data which means the large autoregressive coefficient \( \gamma_1 \), the distribution of \( t_{\text{Plug}} \) in Figure 2 shrinks forward to the origin. It seems to be due to the upward bias of the \( \hat{\Omega} \), so that the term \( c_{\epsilon_0} [\hat{\Omega} \hat{\sigma}^2 - \hat{\Omega} \hat{\beta} \hat{\beta}' \hat{\Omega}] \) normalizes \( t_{\text{Plug}} \) extremely, and Table 2 indicates the convergence of \( \hat{\Omega} \) is slow. On the other
hand the distribution of $t_{Samp.}$ converges to the standard normal distribution at the moderate $T$ periods.

Figure 3 shows the empirical power for $H_0 : \gamma_1 = 0$ based on $t_{Samp.}$, then the number of instrumental variables used for the estimation are changed as follows

$$K = 3 : \{y_{it-1}^{(1)}, y_{it-1}^{(2)}, x_{it}\}, \ K = 5 : \{y_{it-1}^{(1)}, y_{it-1}^{(2)}, y_{it-2}^{(2)}, x_{it}, x_{it-1}\}. \quad (3.30)$$

From the figure we consider that the lack of instruments which belong to the true lag structure of the reduced form ($K = 4$) can loses the power heavily and the additional instruments loses it through the second term $c_*[\Omega\sigma^2 - \Omega\beta\beta'\Omega]$.

4 Conclusions

In this paper, we have developed a simple asymptotic covariance estimation of the doubly filtered LIML estimator for dynamic panel structural equations. The $t$-statistic based on the covariance estimation does not need asymptotic bias corrections and embodies the fixed $N$ or large-$K$ asymptotics which improves the approximation of null distributions. In practical point of view the LIML estimation and testing approach is effective for the large econometric model. For example, the long cross-countries or regional data whose order of $N$ is around 50, 100 with several explanatory variables seems to be suitable to our approaching.

5 Mathematical Derivations

Proof of Theorem : Firstly, we show that

$$\frac{1}{n} G^{(f)} \overset{p}{\rightarrow} G_0 = \left( \begin{array}{c} \theta' \\ (I_{G_2+K_1}) \end{array} \right) D'[w_{i(t-1)}'w_{i(t-1)'}]D(\theta, I_{G_2+K_1}),$$

$$+ c \left( \begin{array}{cc} \Omega & \Omega \\ \Omega & \Omega \end{array} \right),$$

$$\frac{1}{q_n} H^{(f)} \overset{p}{\rightarrow} H_0 = \left( \begin{array}{cc} \Omega & \Omega \\ \Omega & \Omega \end{array} \right). \quad (5.32)$$
By the relation of
\[
\begin{pmatrix}
Y_{t}^{(f)'y'} \\
Z_{t-1}^{(f)'y'}
\end{pmatrix} = \begin{pmatrix}
\theta' \\
I_{G_2+K_1}
\end{pmatrix} D'Z_{t-1}^{(f)'y'} + \begin{pmatrix}
V_{t}^{(f)'y'} \\
O'
\end{pmatrix}
= D'Z_{t-1}^{(f)'y'} + \begin{pmatrix}
V_{t}^{(f)'y'} \\
O'
\end{pmatrix}, \quad (5.33)
\]
we have the decomposition
\[
G^{(f)} = G^{(f,1)} + G^{(f,2)} + G^{(f,2)'} + G^{(f,3)}. \quad (5.35)
\]
Then we can show that
\[
\frac{1}{n} G^{(f,2)} = \frac{1}{NT} D' \sum_{t=1}^{T-1} Z_{t-1}^{(f)'y'} M_{t} (V_{t}^{(f)}, O) \rightarrow O_{G+K_1}. \quad (5.36)
\]
This is so because \((1/\sqrt{NT}) \sum_{t=1}^{T-1} Z_{t-1}^{(f)'y'} M_{t} \psi_{t}^{(f)} \rightarrow O_p(1) + O(1)\) by the same arguments as used for \((1/\sqrt{NT}) \sum_{t=1}^{T-1} Z_{t-1}^{(f)'y'} \mu_{t}^{(f)} \rightarrow O_p(1) + O(1)\) in Kunitomo and Akashi (2008).

Using the following relation
\[
Z_{t-1}^{(f)'y'} = c_t[I_K - \frac{1}{T-t} (\sum_{j=1}^{T-t} \psi_j)] W_{t-1} - c_t \tilde{V}_{tT}
= \psi_t W_{t-1} - c_t \tilde{V}_{tT}, \quad (5.37)
\]
we further decompose \((1/n)G^{(f,1)} = (1/n)D' \sum_{t=1}^{T-1} Z_{t-1}^{(f)'y'} Z_{t-1}^{(f)} D'\) as follow
\[
\frac{1}{n} \sum_{t=1}^{T-1} Z_{t-1}^{(f)'y'} Z_{t-1}^{(f)} = \frac{1}{n} \sum_{t=1}^{T-1} \psi_t W_{t-1} M_t W_{t-1} \psi_t' - \frac{1}{n} \sum_{t=1}^{T-1} c_t \psi_t W_{t-1} M_t \tilde{V}_{tT}
- \frac{1}{n} \sum_{t=1}^{T-1} c_t \tilde{V}_{tT} M_t W_{t-1} \psi_t' + \frac{1}{n} \sum_{t=1}^{T-1} c_t^2 \tilde{V}_{tT} M_t \tilde{V}_{tT}. \quad (5.38)
\]
Moreover, in view of Lemma 2b of Kunitomo and Akashi (2008), and \(c_t^2 = 1 - 1/(T - t + 1)\), we can show that
\[
\frac{1}{n} \sum_{t=1}^{T-1} \psi_t W_{t-1} M_t W_{t-1} - c_t^2 \psi_t W_{t-1} M_t W_{t-1}
- \frac{1}{n} \sum_{t=1}^{T-1} \frac{c_t^2}{T-t} W_{t-1} M_t W_{t-1} (\sum_{j=1}^{T-t} \psi_j)' - \frac{1}{n} \sum_{t=1}^{T-1} \frac{c_t^2}{T-t} (\sum_{j=1}^{T-t} \psi_j) W_{t-1} M_t W_{t-1}
+ \frac{1}{n} \sum_{t=1}^{T-1} (\frac{c_t}{T-t})^2 (\sum_{j=1}^{T-t} \psi_j W_{t-1} M_t W_{t-1} (\sum_{j=1}^{T-t} \psi_j)' \rightarrow E[\psi_{t(t-1)} W_{t(t-1)}]. \quad (5.39)
\]
The second and third terms of (5.38) have zero means and its variances are shown to tend to zeros. In effect,

\[
Var\left[\frac{1}{n}\sum_{t=1}^{T-1} c_t e'_j \Psi_t w'_{t-1} M_t \tilde{V}_{tT} \text{e}_k\right] = \frac{1}{N^2 T^2} \sum_{t=1}^{T-1} \sum_{s=1}^{T-1} c_t c_s \mathcal{E}[(e'_j \Psi_t w'_{t-1} M_t \tilde{V}_{tT} \text{e}_k)(e'_k \tilde{V}_{sT} M_s w_{s-1} \Psi'_s \text{e}_j)]
\]

\[
\leq \frac{1}{N^2 T^2} \sum_{t=1}^{T-1} \sum_{s=1}^{T-1} \sqrt{c_t^2 \mathcal{E}[(e'_j \Psi_t w'_{t-1} M_t \tilde{V}_{tT} \text{e}_k)]^2} \sqrt{c_s^2 \mathcal{E}[(e'_k \tilde{V}_{sT} M_s w_{s-1} \Psi'_s \text{e}_j)]^2}.
\]

where \(e_j (j, k = 1, \ldots, K)\) are \(j-th\) unit vector. And we have

\[
c_t^2 \mathcal{E}[(e'_j \Psi_t w'_{t-1} M_t \tilde{V}_{tT} \text{e}_k)]^2 = c_t^2 \mathcal{E}[(\text{e}_h \Psi_t \text{w}'_{t-1} M_t \tilde{V}_{tT} \text{e}_k)]^2 \mathcal{E}[\text{e}_j \Psi_t \text{w}'_{t-1} M_t \tilde{V}_{tT} \text{e}_k]
\]

\[
\leq c_t^2 \left[\frac{1}{(T - t)^2} e_k \sum_{h=1}^{T-t} \Phi_h \mathcal{E}[(\text{w}^{*}_{i0} \text{w}'_{i0}) \Phi'_h \text{e}_k)] \mathcal{E}[(e'_j \Psi_t \text{w}'_{t-1} M_t \tilde{V}_{tT} \text{e}_k)]
\]

\[
= N(\frac{c_t^2}{T - t})^2 (e_k \sum_{h=1}^{T-t} \Phi_h \mathcal{E}[(\text{w}^{*}_{i0} \text{w}'_{i0}) \Phi'_h \text{e}_k]) \times (e'_j \text{I}_k - \frac{1}{T - t} (\sum_{h=1}^{T-t} \text{I}^{*h}) \mathcal{E}[\text{w}^{*}_{i0} \text{w}'_{i0}] \text{I}_k - \frac{1}{T - t} (\sum_{h=1}^{T-t} \text{I}^{*h}) \text{e}_j) = O(\frac{N}{T - t}),
\]

since \(\sum_{h=1}^{T-t} e_k \Phi_h \mathcal{E}[(\text{w}^{*}_{i0} \text{w}'_{i0}) \Phi'_h \text{e}_k] = O(T - t)\). Thus,

\[
Var\left[\frac{1}{n_0} \sum_{t=1}^{T-1} c_t e'_j \Psi_t w'_{t-1} M_t \tilde{V}_{tT} \text{e}_k\right] \leq \frac{1}{N_0^2 T^2} \sum_{t=1}^{T-1} \sqrt{O(\frac{N_0}{T - t})} \sum_{s=1}^{T-1} \sqrt{O(\frac{N_0}{T - s})} = O(\frac{\sqrt{T}}{N_0 T^2}) \rightarrow 0.
\] (5.42)

For the fourth term of (5.38), its mean is given by

\[
\mathcal{E}\left[\frac{1}{n} \sum_{t=1}^{T-1} c_t^2 e'_j \tilde{V}_{tT} M_t \tilde{V}_{tT} \text{e}_k\right] = \frac{1}{NT} \sum_{t=1}^{T-1} c_t^2 \mathcal{E}[tr(M_t \mathcal{E}[(\tilde{V}_{tT} \text{e}_k e'_j \tilde{V}_{tT}])]]
\]

\[
= \frac{1}{NT} \sum_{t=1}^{T-1} c_t^2 tr(M_t) \mathcal{E}[\text{e}'_j \tilde{V}_{tT} \text{V}'_{tT} \text{e}_k]
\]

\[
= O\left(\frac{1}{NT} \sum_t \frac{tr(M_t)}{T - t + 1}\right) \rightarrow 0.
\] (5.43)

Its variance is shown to tend to zero in the same way as used for \(\Upsilon_2^{(k)}\) and \(\Upsilon_2^{(k)}\), defined in Kunitomo and Akashi (2008).
Finally, we consider about \((1/n)\mathbf{G}^{(f,3)}\). Using the fact that \(\mathcal{E}_t[\mathbf{v}_{it}^{(f)}\mathbf{v}_{it}^{(f)\prime}] = \Omega\),

\[
\mathcal{E}_t\left[\frac{1}{n} \sum_{t=1}^{T-1} \mathbf{e}_g' \mathbf{V}_t^{(f)\prime} \mathbf{M}_t \mathbf{V}_t^{(f)} \mathbf{e}_h\right] = \frac{1}{NT} \mathcal{E}[tr(\mathbf{M}_t \mathbf{V}_t^{(f)} \mathbf{e}_h \mathbf{e}_g' \mathbf{V}_t^{(f)\prime})]
= \frac{\mathbf{e}_g' \Omega \mathbf{e}_h}{NT} \sum_{t=1}^{T-1} tr(\mathbf{M}_t) \rightarrow c(\mathbf{e}_g' \Omega \mathbf{e}_h). \tag{5.44}
\]

Moreover, using \(\mathbf{V}_t^{(f)} = (\mathbf{V}_t - \overline{\mathbf{V}}_t)/c_t\),

\[
\frac{1}{NT} \sum_{t=1}^{T-1} \mathbf{V}_t^{(f)\prime} \mathbf{M}_t \mathbf{V}_t^{(f)} = \frac{1}{NT} \sum_{t=1}^{T-1} c_t^{-2} \mathbf{V}_t' \mathbf{M}_t \mathbf{V}_t - \frac{1}{NT} \sum_{t=1}^{T-1} c_t^{-2} \mathbf{V}_t' \mathbf{M}_t \mathbf{V}_t \tag{5.45}
- \frac{1}{NT} \sum_{t=1}^{T-1} c_t^{-2} \mathbf{V}_t' \mathbf{M}_t \mathbf{V}_t + \frac{1}{NT} \sum_{t=1}^{T-1} c_t^{-2} \mathbf{V}_t' \mathbf{M}_t \mathbf{V}_t.
\]

In view of Lemma 1 of Kunitomo and Akashi (2008) \(\text{Var}[\mathbf{v}_t^{(g)} \mathbf{M}_t \mathbf{v}_t^{(h)}] = O(t)\) and \(\text{Cov}[\mathbf{v}_t^{(g)} \mathbf{M}_t \mathbf{v}_t^{(h)}, \mathbf{v}_s^{(g)} \mathbf{M}_t \mathbf{v}_s^{(h)}] = 0\) for \(t \neq s\). Hence, the variance of the first term satisfies

\[
\text{Var}\left[\frac{1}{NT} \sum_{t=1}^{T-1} \mathbf{e}_g' \mathbf{V}_t^{(f)\prime} \mathbf{M}_t \mathbf{V}_t^{(f)} \mathbf{e}_h\right] = \frac{1}{N^2T^2} \sum_{t=1}^{T-1} (1 + \frac{1}{T-t})^2 O(t) \rightarrow 0. \tag{5.46}
\]

The second and third terms of the right-hand side of (5.45) are analogous to \(\gamma_{21n}^{(k)}\) and \(\gamma_{22n}^{(k)}\), and their variances are shown to tend to zero using similar arguments as used for those terms.

Turning to consider that \(1/q_0 \mathbf{H}^{(f)} \overset{p}{\rightarrow} \mathbf{H}_0\),

\[
\frac{1}{n} \sum_{t=1}^{T-1} \left( \begin{array}{c} \mathbf{Y}_t^{(f)\prime} \\ \mathbf{Z}_{t-1}^{(1,f)\prime} \end{array} \right) \left( \begin{array}{c} \mathbf{Y}_t^{(f)} \\ \mathbf{Z}_{t-1}^{(1,f)} \end{array} \right) = \mathbf{D}^* \frac{1}{n} \sum_{t=1}^{T-1} \mathbf{Z}_{t-1}^{(f)\prime} \mathbf{Z}_{t-1}^{(f)} \mathbf{D}^* + \mathbf{D}^* \frac{1}{n} \sum_{t=1}^{T-1} \mathbf{Z}_{t-1}^{(f)\prime} (\mathbf{V}_t^{(f)}, \mathbf{O}) \tag{5.47}
+ \frac{1}{n} \sum_{t=1}^{T-1} \left( \begin{array}{c} \mathbf{V}_t^{(f)\prime} \\ \mathbf{O}' \end{array} \right) \mathbf{Z}_{t-1}^{(f)} \mathbf{D}^* + \frac{1}{n} \sum_{t=1}^{T-1} \left( \begin{array}{c} \mathbf{V}_t^{(f)\prime} \\ \mathbf{O}' \end{array} \right) (\mathbf{V}_t^{(f)}, \mathbf{O}).
\]

The mean of the second and third terms \(1/(\mathbf{N}_0^T) \sum_t E[\mathbf{Z}_{t-1}^{(f)\prime} \mathbf{V}_t^{(f)}] = 1/T(\mathbf{I}_K - \mathbf{P}^*)^{-1} E[\mathbf{v}_i^{(f)\prime} \mathbf{v}_i^{(f)}] + O(1/\mathbf{N}_0^T)\) converge zero as \(T \rightarrow \infty\), and the mean squared convergence are shown. Moreover,

\[
\frac{1}{n} \sum_{t=1}^{T-1} \mathbf{e}_g' \mathbf{Z}_{t-1}^{(f)\prime} \mathbf{Z}_{t-1}^{(f)} \mathbf{e}_k = \frac{1}{N_0T} \sum_{i=1}^{N_0} \sum_{t=1}^{T} \mathbf{w}_{i(t-1)}^{(j)} \mathbf{w}_{i(t-1)}^{(k)} - \frac{1}{N_0T} \sum_{i=1}^{N_0} \frac{1}{T} (\mathbf{L}^{(j)\prime} \mathbf{w}_{i(t-1)}^{(j)} \mathbf{w}_{i(t-1)}^{(k)} \mathbf{L}^{(j)}).
\tag{5.48}
\]
since $1/T \sum_{t=1}^{T} w_{i(t-1)}^{(j)} w_{i(t-1)}^{(k)} \overset{p}{\to} \mathcal{E}[w_{i(t-1)}^{(j)} w_{i(t-1)}^{(k)}]$ by the ergodic theorem, and the second term converges to $1/N_0 \sum_{i=1}^{N_0} (1 + \alpha_p(1))^2 = \alpha_p(1)$ by the fact that $(1/T) \mathcal{L}_T w_{i(t-1)}^{(j)} \overset{p}{\to} 0$. Using the similar argument, it is shown that $1/n \sum_{t=1}^{T-1} V_t^{(f)} V_t^{(f)} \overset{p}{\to} \Omega$.

Therefore,

$$\frac{1}{q_n} H^{(f)} \overset{p}{\to} \frac{1}{1-c} \left( \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{T-1} \left( \begin{array}{c} Y_t^{(f)} \\ Z_t^{(1,f)} \end{array} \right) \left( \begin{array}{c} Y_t^{(f)} \\ Z_t^{(1,f)} \end{array} \right) - G_0 \right) = H_0. \quad (5.49)$$

Thus, $\hat{\Phi}^* \overset{p}{\to} \Phi^*$, $\hat{\Omega} \overset{p}{\to} \Omega$ and $\hat{\sigma}^2 \overset{p}{\to} \sigma^2$.

Next, we shall show that $\Gamma^* F_n \Gamma^* \overset{p}{\to} \Psi_0^*$, where $\Psi^* = \Phi^*^{-1} \Psi_0 \Phi^*^{-1}$. The following equality exactly holds

$$\Gamma^* F_n \Gamma^* = \frac{1}{N_0 T} \sum_{t=1}^{T-1} A_t^{(f)} A_t^{(f)'} \quad (5.50)$$

$$= \frac{1}{N_0 T} \sum_{t=1}^{T-1} (A_t^{(N)} + D' \sum_{j=1}^{4} [Y_{jt}^{(M)} - Y_{jt}^{(cN)}] + Y_{jt}^{(u)}) \quad (5.51)$$

$$\times (A_t^{(N)} + D' \sum_{j=1}^{4} [Y_{jt}^{(M)} - Y_{jt}^{(cN)}] + Y_{jt}^{(u)}),$$

where

$$A_t^{(f)} = D' Z_{t-1}^{(f)} N_t u_t^{(f)} + \left( \begin{array}{c} U_t^{(f)} \\ O' \end{array} \right) N_t u_t^{(f)}, \quad (5.52)$$

$$A_t^{(N)} = D' W_{t-1} N_t u_t + \left( \begin{array}{c} U_t^{(c)} \\ O' \end{array} \right) N_t u_t, \quad (5.53)$$

$$N_t = M_t - c[I_{N_0} - M_t] = \frac{1}{1-c} [M_t - cI_{N_0}], \quad (5.54)$$

$$(U_t^{(f)}, U_t^{(c)}) = (0, I_{G_2})[I_{1+G_2} - \frac{\Omega \beta \beta'}{\beta' \Omega \beta}] (V_t^{(f)}, V_t), \quad (5.55)$$

$$Y_{1t}^{(P)} = \frac{-1}{1-c} W_{t-1} P_{u_t}, \quad (5.56)$$

$$Y_{2t}^{(P)} = \frac{-1}{1-c} \frac{c_t}{T} W_{t-1} P_{u_t}, \quad (5.57)$$

$$Y_{3t}^{(P)} = \frac{-1}{1-c} V_{t} P_{u_t}, \quad (5.58)$$

$$Y_{4t}^{(P)} = \frac{1}{1-c} V_{t} P_{u_t}, \quad (5.59)$$

$$Y_t^{(u)} = \left( \begin{array}{c} U_t^{(c)} \\ O' \end{array} \right) N_t u_t^{(f)} - \left( \begin{array}{c} U_t^{(c)} \\ O' \end{array} \right) N_t u_t. \quad (5.60)$$
and the argument of \((P)\) takes \(P = M_t\) or \(P = cI_{N_0}\).

We first show that \((1/N_0T)\sum_t(D' \sum_{j=1}^{4}[\gamma_{jt}^{(M_t)} - \gamma_{jt}^{(cI_{N_t})}] + \gamma_{it}^{(u)})\sum_{j=1}^{4}[\gamma_{jt}^{(M_t)} - \gamma_{jt}^{(cI_{N_t})}] + \gamma_{it}^{(u)}(t)\) converges to zero in probability, and it suffices to check that the diagonal terms of the each \((1/N_0T)\sum_t \gamma_{jt}^{(P)} \gamma_{jt}^{(P')}\) converges to zero.

Using the result of Kunitomo and Akashi (2008), for any \(k (k = 1, \ldots, K)\),

\[
\mathbb{E}\left[\frac{1}{N_0T} \sum_{t=1}^{T-1} e_k' \gamma_{1t}^{(M_t)} \gamma_{1t}^{(M_t)'} e_k\right] = \left(\frac{1}{1-c}\right)^2 \frac{1}{N_0T} \sum_{t=1}^{T-1} \mathbb{E}[\omega_{t-1}^{(k)} M_t u_t u_t' M_t \omega_{t-1}^{(k)}] = \left(\frac{1}{1-c}\right)^2 \frac{1}{N_0T} \frac{\sigma^2}{T-t+1} \sum_{t=1}^{T-1} \mathbb{E}[\omega_{t-1}^{(k)} M_t \omega_{t-1}^{(k)}] = O\left(\frac{N_0 \log T}{N_0 T}\right) \to 0, \tag{5.61}
\]

so that for large \(T\) there is a positive constant \(c_k\) such that \(\mathbb{E}\left[\frac{1}{N_0T} \sum_{t=1}^{T-1} e_k' \gamma_{1t}^{(M_t)} \gamma_{1t}^{(M_t)'} e_k\right] \leq c_k(\log T/T).\) Using the Markov inequality and taking \(\epsilon_T = c_k/\sqrt{T},\)

\[
\Pr\left(\left|\frac{1}{N_0T} \sum_{t=1}^{T-1} e_k' \gamma_{1t}^{(M_t)} \gamma_{1t}^{(M_t)'} e_k\right| \leq \epsilon_T\right) \geq 1 - c_k \frac{\log T}{T \epsilon_T} \to 1. \tag{5.62}
\]

Similarly we can show that \((1/N_0T)\sum_{t=1}^{T-1} e_k' \gamma_{2t}^{(M_t)} \gamma_{2t}^{(M_t)'} e_k \to 0.\)

For the term \((1/N_0T)\sum_{t=1}^{T-1} e_k' \gamma_{3t}^{(M_t)} \gamma_{3t}^{(M_t)'} e_k,\) we have

\[
\frac{1}{N_0T} \sum_{t=1}^{T-1} e_k' \gamma_{3t}^{(M_t)} \gamma_{3t}^{(M_t)'} e_k = \left(\frac{1}{1-c}\right)^2 K \sum_{j=1}^{K} \frac{1}{N_0T} \sum_{t=1}^{T-1} \frac{1}{T-t} \sum_{h=1}^{T-t} e_k' \Phi_h e_j e_j' V_{T-h}^{(M_t)} M_t u_t)^2
\]

\[
+ \sum_{i,j=1}^{K} \frac{1}{N_0T} \sum_{t=1}^{T-1} \frac{1}{T-t} \sum_{h=1}^{T-t} e_k' \Phi_h e_i e_j' V_{T-h}^{(M_t)} M_t u_t) \left(\frac{1}{T-t} \sum_{h=1}^{T-t} e_k' \Phi_h e_j e_j' V_{T-h}^{(M_t)} M_t u_t)\right].
\]

By the Cauchy-Schwarz and the Markov inequalities, it suffices to consider

\[
\mathbb{E}\left[\frac{1}{N_0T} \sum_{t=1}^{T-1} \left(\frac{1}{T-t} \sum_{h=1}^{T-t} e_k' \Phi_h e_j e_j' V_{T-h}^{(M_t)} M_t u_t)^2\right]\right] \tag{5.63}
\]

\[
= \frac{1}{N_0T} \sum_{t=1}^{T-1} \left(\frac{1}{T-t}\right)^2 \sum_{h=1}^{T-t} (e_k' \Phi_h e_j)^2 \mathbb{E}[(e_j' V_{T-h}^{(M_t)} M_t u_t)^2] = \frac{1}{N_0T} O(\log T) \to 0,
\]

so that the cross terms are negligible

\[
(1/N_0T)\sum_{t=1}^{T-1} e_k' \gamma_{4t}^{(M_t)} \gamma_{4t}^{(M_t)'} e_k \to 0.
\]

Similarly for the term \((1/N_0T)\sum_{t=1}^{T-1} e_k' \gamma_{4t}^{(M_t)} \gamma_{4t}^{(M_t)} e_k,\) it suffices for any \(k, j\) to consider

\[
\frac{1}{N_0T} \sum_{t=1}^{T-1} \mathbb{E}\left[(\frac{1}{T-t} \sum_{h=1}^{T-t} e_k' \Phi_h e_j e_j' V_{T-h}^{(M_t)} M_t u_t)^2\right] \to 0. \tag{5.64}
\]
Using the result of Lemma 1 in Kunitomo and Akashi (2008),

$$E\left[ \frac{1}{T - t} \sum_{h=1}^{T-t} e_k' \Phi_h e_j e_j' V_{T-h}^s M_t \bar{u}_{iT} \right]^2 
(5.65)$$

$$= Var\left[ \frac{1}{T - t} \sum_{h=1}^{T-t} e_k' \Phi_h e_j e_j' V_{T-h}^s M_t \bar{u}_{iT} \right] + (E\left[ \frac{1}{T - t} \sum_{h=1}^{T-t} e_k' \Phi_h e_j e_j' V_{T-h}^s M_t \bar{u}_{iT} \right] tr(M_t))^2$$

$$= O\left( \frac{1}{(T - t)^2} \right)$$,

so that (5.64) is valid and $(1/N_0 T) \sum_{t=1}^{T-1} e_k' Y_{4t}^{(M_t)} Y_{4t}^{(M_t)'} e_k \overset{p}{\to} 0$.

As for the terms $(1/N_0 T) \sum_{t=1}^{T-1} e_k' Y_{jt}^{(c_{1N})} Y_{jt}^{(c_{1N})'} e_k$ $(j = 1, 2)$ are analogous to $(1/N_0 T) \sum_{t=1}^{T-1} e_k' Y_{jt}^{(M_t)} Y_{jt}^{(M_t)'} e_k$ $(j = 1, 2)$ respectively, so that the both term converge to zero in probability.

For the term $(1/N_0 T) \sum_{t=1}^{T-1} e_k' Y_{3t}^{(c_{1N})} Y_{3t}^{(c_{1N})'} e_k$, it suffices to consider

$$E[\left( \frac{c}{1-c} \right)^2 \frac{1}{N_0 T} \sum_{t=1}^{T-1} \left( \frac{1}{T - t} \sum_{h=1}^{T-t} e_k' \Phi_h e_j e_j' V_{T-h}^s u_t \right)^2] 
(5.66)$$

$$= \frac{c}{(1-c)^2 N_0^2 T} \sum_{t=1}^{T-1} \left( \frac{1}{T - t} \right)^2 \sum_{h=1}^{T-t} (e_k' \Phi_h e_j)^2 E[(e_j' V_{T-h}^s u_t)^2] = \frac{K}{T} O(\log T) \to 0,$$

since $c = K/N_0$ or zero, and $\sum_{h=1}^{T-t}(\Phi_{k(j,h)})^2 E[(V_{T-h}^s u_t)^2] = O((T - t) N_0^2)$. Using the similar arguments we have $E\left[ (c/1-c)^2 (1/N_0 T) \sum_{t=1}^{T-1} e_k' \phi_{jt}^{(c_{1N})} \phi_{jt}^{(c_{1N})'} e_k \right] = O(K/T)$, so that this term is also negligible.

It can be shown that $(1/N_0 T) \sum_{t=1}^{T-1} e_k' Y_{jt}^{(w)} Y_{jt}^{(w)'} e_g$, $(g = 1, ..., 1 + G_2)$ converge to zero in probability by the arguments as used for the terms $(1/N_0 T) \sum_{t=1}^{T-1} e_k' \phi_{jt}^{(p)} \phi_{jt}^{(p)'} e_k$, $(j = 3, 4)$ using the decomposition based on $V_{jt}^{(p)} = (V_t - \bar{V}_{iT})/c_t$ and $u_{jt}^{(f)} = (u_t - \bar{u}_{iT})/c_t$.

Secondly, we shall show the following equalities

$$\frac{1}{N_0 T} \sum_{t=1}^{T-1} A_t^{(N)} A_t^{(N)'} = \frac{1}{N_0 T} \sum_{t=1}^{T-1} A_t A_t' + o_p(1) \quad (5.67)$$

$$= \frac{1}{N_0 T} \sum_{t=1}^{T-1} E_t[A_t A_t'] + o_p(1), \quad (5.68)$$

and we can show that $(1/N_0 T) \sum_{t=1}^{T-1} E_t[A_t A_t'] = \Psi_0 + o_p(1)$, where $A_t = D' W_{t-1} u_t + (U_t^p, O)' N_t u_t$. Note that $N_t = I_{N_0} - (1 + c_*) (I_{N_0} - M_t)$ and for any $j, k (j, k =
1, ..., \( K \) we have to consider

\[
\frac{(1 + c_s)^2}{N_0T} \sum_{t=1}^{T-1} e_j' D' W_{t-1}' [I_{N_0} - M_t] u_j u'_t [I_{N_0} - M_t] W_{t-1} D e_k' \xrightarrow{p} 0, \quad (5.69)
\]

\[
-\frac{(1 + c_s)}{N_0T} \sum_{t=1}^{T-1} e_j' D' W_{t-1}' [I_{N_0} - M_t] u_j u'_t W_{t-1} D e_k' \xrightarrow{p} 0, \quad (5.70)
\]

\[
-\frac{(1 + c_s)}{N_0T} \sum_{t=1}^{T-1} e_j' D' W_{t-1}' [I_{N_0} - M_t] u_j u'_t N_t (U_t^\perp, O) e_k' \xrightarrow{p} 0, \quad (5.71)
\]

but the result of Kunitomo and Akashi (2008) implies \((1/N_0T) \sum_{t=1}^{T-1} (e_j' D' W_{t-1}' [I_{N_0} - M_t] u_j)^2 \xrightarrow{p} 0\) and it can be shown that \((1/N_0T) \sum_{t=1}^{T-1} (u_j^T W_{t-1} D e_k')^2\) and \((1/N_0T) \sum_{t=1}^{T-1} (u_j^T N_t u_j e_k)^2\) converge to a constant. So that using the Cauchy-Schwarz inequality, (5.67) is established.

To verify (5.68) by mean squared convergence, we consider the variances of the martingale difference sequence, for \( j, k \),

\[
Var\left[ \frac{1}{T} \sum_{t=1}^{T-1} \frac{1}{N_0} e_{j, t} (A_t A'_t - \mathcal{E}[A_t A'_t]) e_k \right] = \frac{1}{T^2} \sum_{t=1}^{T-1} \frac{1}{N_0^2} \mathcal{E}[(e_{j, t} A_t A'_t e_k)^2] - \mathcal{E}[(e_{j, t} \mathcal{E}[A_t A'_t e_k)^2])]. \quad (5.72)
\]

Then we consider that \((1/T^2) \sum_{t} (1/N_0^2) \mathcal{E}[(e_{j, t} D' W_{t-1} u_t)^4]\) and \((1/T^2) \sum_{t} (1/N_0^2) \mathcal{E}[(e_{k, t} U_{t}^\perp N_t u_t)^4]\) converge to zero, since \(\mathcal{E}[(e_{j, t} A_t A'_t e_k)^2] \geq \mathcal{E}[(e_{j, t} \mathcal{E}[A_t A'_t e_k)^2]\) and that \(\sum_t \mathcal{E}[(e_{j, t} A_t A'_t e_k)^2]\) is bounded by the terms \(\sum_t \mathcal{E}[(e_{k, t} D' W_{t-1} u_t)^4]\) and \(\sum_t \mathcal{E}[(e_{k, t} U_{t}^\perp N_t u_t)^4]\). For any \( t, N_0 \),

\[
\frac{1}{N_0^2} \mathcal{E}[(e_{k, t} D' W_{t-1} u_t)^4] = \frac{1}{N_0^2} \mathcal{E}[(e_{k, t} D' \sum_{i=1}^{N_0} w_{i(t-1)u_{it}})^4] = \mathcal{E}[(e_{k, t} D' w_{i(t-1)u_{it}})^4] + \mathcal{E}[(e_{k, t} D' w_{i(t-1)u_{it}})^2] \mathcal{E}[(e_{k, t} D' w_{j(t-1)u_{jt}})^2] = O(1),
\]

since \(\mathcal{E}[(e_{k, t} D' w_{i(t-1)u_{it}})^4] = 0\) and the independence of \(i\)'s. For any \( t \) and fixed \( N_0 \), \(O(\mathcal{E}[(e_{k, t} U_{t}^\perp N_t u_t)^4]) = O(\mathcal{E}[(e_{k, t} U_{t}^\perp M_t u_t)^4]))\), and from the result of Lemma 3 in Kunitomo and Akashi (2008), we have \(\mathcal{E}[(e_{k, t} U_{t}^\perp M_t u_t)^4] = O(1)\). Hence (5.68) is valid.

Finally we show that \( \hat{\Gamma}^* \hat{F}_n \hat{\Gamma}^* \xrightarrow{p} \Gamma^* F_n \Gamma^* \). For this purpose, begin by considering that

\[
\frac{\sqrt{N_0T}}{\log T} \left( \frac{1}{N_0T} G^{(f)} - G_0 \right) = O_p(\sqrt{\frac{N_0}{T}}). \quad (5.75)
\]
This is so because $\text{Var}[(1/\sqrt{N_0T})e_g^t G^{(f)} e_h] = O(1)$, $\mathcal{E}[(\sqrt{N_0 T} e_g^t ((1/N_0T) G^{(f)} - G_0) e_h^t)] = O(\sqrt{(N_0/T)^2} \sum_t (1/t))$ for $g, h = 1, \ldots, (1+G+K_1)$ under $\lim(N_0/T) > 0$, and the evaluation by the Chebyshev inequality. Then, using the relations of $G$-totic normality and also those of $\hat{\theta}$ inequality, for any $\vartheta$ variance trem. The order of $\hat{\theta}$ is $O_p(c_0 - c) + O_p(\sqrt{\frac{\log T}{T}})$,

$$\hat{\Omega} - \Omega = O_p(\sqrt{\frac{\log T}{T}} \sqrt{T})$$

$$= \frac{1}{\sqrt{N_0}} [O_p\left(\sqrt{\frac{K}{\sqrt{N_0}}} - \sqrt{N_0c} + O_p\left(\frac{\log T}{\sqrt{T}}\right)\right)]$$

and note that $d_{n_0} \to 0$ as $T \to \infty$ regardless whether $N_0$ is fixed or tends to infinity. We consider the following decomposition and the validity of the second equality

$$\hat{\Gamma}^* \hat{F}_n \hat{\Gamma}^*$$

$$= \frac{1}{N_0 T} \sum_{t=1}^{T-1} (\Gamma^* + \hat{\Gamma}_\epsilon^*)(F_t^{(f)} + \hat{F}_t^{(f)})(\theta_1 + \hat{\theta}_1)(\theta_1' + \hat{\theta}_1')(F_t^{(f)} + \hat{F}_t^{(f)})(\Gamma^* + \hat{\Gamma}_\epsilon^*)$$

$$= \Gamma^* F_n \Gamma^* + O_p(d_{n_0}),$$

where

$$F_t^{(f)} = G_t^{(f)} - c_\epsilon H_t^{(f)}, \quad \hat{F}_t^{(f)} = (c_0 - c) H_t^{(f)}, \quad \hat{\Gamma}_\epsilon^* = \hat{\Gamma}^* - \Gamma^*,$$ (5.78)

$$\theta_1' = (1, \theta'), \quad \theta_1' = (1, \hat{\theta}) - (1, \theta').$$ (5.79)

Pick up the term $(1/N_0T) \sum_t \Gamma^* F_t^{(f)} \theta_1 \hat{\theta}_1' F_t^{(f)} \hat{\Gamma}_\epsilon^* e_k$, and using the Cauchy-Schwarz inequality, for any $j, k = 1, \ldots, (G_2 + K_1)$,

$$\frac{1}{N_0 T} \sum_{t=1}^{T-1} \left| e_j^t \Gamma^* F_t^{(f)} \theta_1 \hat{\theta}_1' F_t^{(f)} \hat{\Gamma}_\epsilon^* e_k \right|$$

$$\leq \left[ \frac{1}{N_0 T} \sum_{t=1}^{T-1} (e_j^t \Gamma^* F_t^{(f)} \theta_1)^2 \right]^{1/2} \left[ \frac{1}{N_0 T} \sum_{t=1}^{T-1} (\hat{\theta}_1' F_t^{(f)} \hat{\Gamma}_\epsilon^* e_k)^2 \right]^{1/2},$$ (5.80)

where we have shown that the first term of (5.80) converges to a square root of a variance trem. The order of $\hat{\theta}_1$ of the second term is $O_p(1/\sqrt{N_0T})$ by its asymptotic normality and also those of $\hat{\Gamma}_\epsilon^*$ are $O_p(d_{n_0}/\sqrt{N_0})$, since $\Gamma^*$ is a continuous function of $\theta, \Omega$ and (5.76). Therefore we can rewrite the terms which are multiplied
by at least the one error term \( \hat{\theta}_{1e} \) or \( \hat{\Gamma}_{*} \) as follow. For \( l, m = 1, ..., (1 + G_2 + K_1) \),

\[
\frac{1}{N_0 T} \sum_{t=1}^{T-1} (\hat{\theta}_{1e}^t F_t^{(f)} \hat{\Gamma}_e e_k)^2
\]

\[
= \sum_{l, m=1}^{1+G_2+K_1} (e_k^l \hat{\Gamma}_e^* e_l) \left[ \frac{1}{N_0 T} \sum_{t=1}^{T-1} e_l^t F_t^{(f)} \hat{\theta}_{1e}^t \hat{\theta}_{1e}^t F_t^{(f)} e_m \right] (e_m^l \hat{\Gamma}_e^* e_k),
\]

\[
= \sum_{l, m=1}^{1+G_2+K_1} (l_2^t e_l) \left[ \frac{1}{N_0 T} \sum_{t=1}^{T-1} e_l^t F_t^{(f)} e_1 e_1 F_t^{(f)} e_m \right] (e_m^l \nu_{2n}), \tag{5.81}
\]

where \( e_{2n} = (d_{n0}/\sqrt{N_0}) \nu_{1n} \) and \((e_{2n}, e_{1n})\) are \((1 + G_2 + K) \times 1\) vectors whose element’s are \(O_p(1)\) respectively. Thus we focus on the middle term of (5.81),

\[
\frac{1}{N_0 T} \sum_{t=1}^{T-1} |e_l^t F_t^{(f)} e_1 e_1 F_t^{(f)} e_m|
\]

\[
\leq \sqrt{\frac{1}{N_0^2 T} \sum_{t=1}^{T-1} (e_l^t F_t^{(f)} \nu_{1n} d_n)^2} \sqrt{\frac{1}{N_0^2 T} \sum_{t=1}^{T-1} (e_m^t F_t^{(f)} \nu_{1n} d_n)^2} \tag{5.82}
\]

Moreover, consider the first term of (5.82) without loss of generality, for \(g, g', h, h' = 1, ..., (1 + G_2 + K_1)\),

\[
\frac{1}{N_0^2 T} \sum_{t=1}^{T-1} (e_l^t F_t^{(f)} \nu_{1n} d_n)^2 \leq \sum_{g, h=1}^{1+G_2+K_1} \sum_{g', h'=1}^{1+G_2+K_1} |e_g^t e_h||d_{n0} \nu_{1n} e_h||e_g^t e_{g'}||d_{n0} \nu_{1n} e_{h'}|
\]

\[
\times \left( \frac{1}{N_0^2 T} \sum_{t=1}^{T-1} |e_g^t F_t^{(f)} e_h||e_g^t F_t^{(f)} e_{h'}| \right). \tag{5.83}
\]

By the relation of that \( |e_g^t F_t^{(f)} e_h| \leq (1 + c_*) |(e_g^t Z_r^{(f)} Z_r^{(f)} e_g)(e_h^t Z_r^{(f)} Z_r^{(f)} e_h)|^{1/2} \), the last term of (5.83) is evaluated as

\[
\mathcal{E} \left[ \frac{1}{T} \sum_{t=1}^{T-1} \frac{e_g^t F_t^{(f)} e_h}{N_0} \frac{e_g^t F_t^{(f)} e_{h'}}{N_0} \right]
\]

\[
\leq \frac{(1 + c_*)^2}{T} \sum_{t=1}^{T-1} \left( \mathcal{E} \left[ \frac{e_g \sum_{i=1}^{N_0} z_{it}^{(f)} \cdot z_{it}^{(f)} e_g}{N_0} \frac{e_h \sum_{i=1}^{N_0} z_{it}^{(f)} \cdot z_{it}^{(f)} e_h}{N_0} \right] \right)^{1/2}
\]

\[
\times \left( \mathcal{E} \left[ \frac{e_{g'} \sum_{i=1}^{N_0} z_{it}^{(f)} \cdot z_{it}^{(f)} e_{g'}}{N_0} \frac{e_{h'} \sum_{i=1}^{N_0} z_{it}^{(f)} \cdot z_{it}^{(f)} e_{h'}}{N_0} \right] \right)^{1/2} = O(1), \tag{5.84}
\]

so that for any \(T, N_0\), we conclude this term as \(O_p(1)\) by the Markov inequality, and the order of (5.83) is \(O_p(d_{n0}^2)\) since \((1 + G_2 + K_1) < \infty\). Therefore (5.80) is evaluated as \(O_p(d_{n0})\), and also the other terms of (5.77) are \(O_p(d_{n0})\) at most. Thus
we obtain the desired result. Q.D.E.

References


ments,” Discussion Paper CIRJE-F-467, Graduate School of Economics, University of Tokyo.

Figure 1: Null Distributions: $H_0 : \beta_2 = 0$ with $N = 100, T = 15$

Table 1: Empirical sizes of statistics that test $H_0 : \beta_2 = 0$ with $K = 4, \gamma_1 = 0$

<table>
<thead>
<tr>
<th>$(N, T)$</th>
<th>(100,15)</th>
<th></th>
<th>(100,30)</th>
<th></th>
<th>(200,30)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>size</td>
<td>0.01</td>
<td>0.05</td>
<td>0.10</td>
<td></td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>$t_{GMM}$</td>
<td>0.011</td>
<td>0.049</td>
<td>0.106</td>
<td></td>
<td>0.026</td>
<td>0.104</td>
</tr>
<tr>
<td>$t_{Plug,c=0}$</td>
<td>0.022</td>
<td>0.085</td>
<td>0.146</td>
<td></td>
<td>0.022</td>
<td>0.087</td>
</tr>
<tr>
<td>$t_{Plug}$</td>
<td>0.007</td>
<td>0.037</td>
<td>0.079</td>
<td></td>
<td>0.007</td>
<td>0.048</td>
</tr>
</tbody>
</table>
Figure 2: Null Distributions: \( H_0 : \beta_2 = 0 \) with \( \gamma_1 = 0.7, N = 15, T = 30 \)

Table 2: Empirical sizes of statistics that test \( H_0 : \beta_2 = 0 \) with \( N = 15, K = 4 \)

<table>
<thead>
<tr>
<th>( (\gamma_1, T) )</th>
<th>(0.7,30)</th>
<th>(0.7,200)</th>
<th>(0.9,30)</th>
<th>(0.9,200)</th>
</tr>
</thead>
<tbody>
<tr>
<td>size</td>
<td>0.05</td>
<td>0.10</td>
<td>0.05</td>
<td>0.10</td>
</tr>
<tr>
<td>( t_{Plug} )</td>
<td>0.013</td>
<td>0.034</td>
<td>0.017</td>
<td>0.047</td>
</tr>
<tr>
<td>( t_{Samp} )</td>
<td>0.062</td>
<td>0.103</td>
<td>0.056</td>
<td>0.109</td>
</tr>
</tbody>
</table>

Figure 3: Power of Tests: \( H_0 : \gamma_1 = 0 \) with \( N = 15, T = 30 \)