Instrumental Variables Estimation of a Nearly Nonstationary Error Component Model*

In Choi†
Kookmin University, Seoul

October, 1999

Abstract

This paper studies instrumental variables (IV) estimation for an error component model with stationary and nearly nonstationary regressors. It is assumed that the numbers of cross section and time series observations are infinite. Furthermore, autoregressive disturbances are assumed for the error component model, the structure of which may vary with individuals. The estimators considered are the Within-IV, Within-IV-OLS, Within-IV-GLS and IV-GLS estimators. The GLS estimators use Gohberg's formula which is particularly useful when autoregressive structures are imposed on the disturbance terms. Both sequential and joint limit theories for the estimators are derived. It is shown that all of the estimators have normal distributions in the limit. Additionally, Wald tests for coefficient vectors and Durbin-Wu-Hausman tests for exogeneity are studied. Simulation results regarding the estimator efficiency and the size and power of the Wald and Durbin-Wu-Hausman tests are also reported.

---

*Part of this paper was written while the author was visiting the Department of Economics, University of Southern California. I thank the faculty and staff of the department, especially Quang Vuong, Isabelle Perrigne, Cheng Hsiao and Young Miller, for their support and hospitality. In addition, I thank the participants at the USC Econometrics Seminar, especially Cheng Hsiao and Hashem Pesaran, who provided insightful comments on this paper. My thanks also go to Jinwoo Jun and Hyoseok Lee who provided research assistance. Last, I gratefully acknowledge financial support from the Korea Research Foundation in the form of a nondirected research fund.

†**Address correspondence to:** In Choi, Department of Economics, Kookmin University, 861–1, Chongnung-dong, Songbuk-gu, Seoul, 136–702 KOREA. E-mail: ichoi@kmu.kookmin.ac.kr

1 Introduction

Recently, there has been a surge of interest in the nonstationary panel data analysis where both the numbers of time series and cross section observations are assumed to be infinite. This reflects the facts that many time series data in economics appear to be nonstationary and that panel data covering a relatively long time period is now available in such areas as macroeconomics and international finance. Studies which develop econometric tools for nonstationary panel data are diverse, but may be categorized by two broad themes: unit root and cointegration tests for panel data and regressions methods for nonstationary panel data. Articles which belong to the first category are Levin and Lin (1992), Im, Pesaran and Shin(1995), Maddala and Wu (1996) and Choi (1997), and those which belong to the second category include, among others, Pedroni (1996, 1997), Kao and Chiang (1997), Phillips and Moon (1998) and Choi (1998).

In addition to the econometric tools studied in these articles, this paper purports to develop the instrumental variables (IV) estimation methods for an error component model which involves both stationary and nearly nonstationary variables. An error component model which involves both stationary and difference-stationary variables is studied in Choi (1998). The model considered in this paper, though, is more general than the one in Choi (1998) in that the regressor-error dependence is allowed for both types of regressors and that nearly nonstationary variables instead of difference-stationary variables are considered. In the time series literature, it is well documented that the presence of nearly nonstationary variables poses serious problems for inferential procedures (cf. Cavanagh, Elliott and Stock, 1995). Furthermore, the model considered in this paper has autoregressive (AR) disturbances as in Choi (1998), the structure of which may vary with individuals. In this respect and in the respect that the model contains both stationary and nearly nonstationary variables, the model is general enough to include some earlier error component models as special cases.

The estimators which are considered in this paper are Within-IV, Within-IV-OLS, IV-GLS and Within-IV-GLS estimators. The Within-IV, Within-IV-OLS and Within-IV-GLS are obtained by running IV, IV-OLS and IV-GLS regressions, respectively, on the error component model where individual effects are subtracted before running regressions. These estimators and the IV-GLS estimator are shown to have normal distributions in the limit under proper moment conditions either when $T \to \infty$ and then $N \to \infty$ or when $T, N \to \infty$ simultaneously, where T denotes the number of time series observations and N the number of cross section observations. The joint limit results obtained under the condition $T, N \to \infty$ require more restrictive moment conditions.

There may seem to be nothing remarkable about the asymptotic normality results obtained in this paper. However, the time series IV estimation for the linear model involving either nearly nonstationary or difference-stationary regressors does not yield even asymptotic mixture normality results required for standard testing procedures to be applied to the model unless the regressors are totally exogenous. Moreover, standard methods which are used to correct the second order bias in nonstationary time series regressions (e.g., Phillips and Hansen, 1990, Saikkonnen, 1991) do not provide mixture normality results in the limit when the regressors are nearly nonstationary, as discussed in Cavanagh, Elliott and Stock (1995). Therefore, the asymptotic normality results in this paper show that additional data collection which makes panel regression feasible enables us to overcome these well known difficulties in nonstationary time series regressions.

Besides the properties of the IV estimators, asymptotic properties of Wald tests are reported in this paper. In addition, the Durbin-Wu-Hausman tests (cf. Durbin, 1954, Wu, 1973 and Hausman, 1978; hereafter, DWH tests) for exogneity which can be used under more general conditions than in the original papers are proposed and studied. The finite sample properties of the estimators and the Wald and DWH tests are also studied by simulation.

The plan of this paper is as follows. Section 2 introduces the model and assumptions. Section 3 derives the asymptotic distributions of the Within-IV, Within-IV-OLS, IV-GLS and Within-IV-GLS estimators for the error component model. Section 4 considers Wald tests for coefficient vectors. Section 5 proposes the DWH tests for the error component model and studies their asymptotic distributions. Section 6 reports some simulation results regarding the estimator efficiency and the empirical size and power of the Wald and DWH tests. Section 7 contains summary and further remarks.

The following notation will be used throughout the paper.

Notation	Meaning
$\Rightarrow$	weak convergence
$\xrightarrow{p}$	convergence in probability
L	lag operator
$\triangleq$	definitional relation
$\parallel A \parallel$	$\sqrt{tr(AA')}$
$[A]_{(h,j)}$	(h,j)-th element of matrix A
$r \wedge t$	$min(r,t)$
P_a	$A(A'A)^{-1}A'$
$B(r)$	Brownian motion

2 The Model and Preliminary Assumptions

Consider the one-way error component model

$$y_{it} = \beta + \alpha_1' x_{1it} + \alpha_2' x_{2it} + u_{it}, \quad i = 1, ..., N;\ t = 1, ..., T, \tag{1}$$

where the $k_1 \times 1$ vector x_{1it} is $I(0)$, the $k_2 \times 1$ vector x_{2it} is nearly nonstationary in the sense that $\left(I - e^{C_{x2i}} L\right) x_{2it} \triangleq \Delta_{C_{x_{2i}}} x_{2it}$ is $I(0)$ ($C_{x_{2i}}$ is a constant matrix) and u_{it} is the $I(0)$ disturbance term. In model (1), the index i denotes households, individuals, countries, etc., and the index t time. Here, the coefficient vectors α_1 and α_2 do not change with individuals and time, but the coefficients associated with the disturbance terms are assumed to vary with individuals.

It is sometimes convenient to write model (1) in matrix notation:

$$\begin{bmatrix} y_1 \\ . \\ . \\ y_N \end{bmatrix} = i_{NT}\beta + \begin{bmatrix} X_1 \\ . \\ . \\ X_N \end{bmatrix} \alpha + \begin{bmatrix} u_1 \\ . \\ . \\ u_N \end{bmatrix} \tag{2}$$

or

$$y = i_{NT}\beta + X\alpha + u, \tag{3}$$

where $y_i \triangleq [y_{i1}, ..., y_{iT}]'$, i_{NT} is a unit vector of size NT, $X_i \triangleq \begin{bmatrix} X_{1i} & X_{2i} \end{bmatrix} = \begin{bmatrix} x'_{1i1} & x'_{2i1} \\ . & . \\ . & . \\ x'_{1iT} & x'_{2iT} \end{bmatrix}$ and $\alpha \triangleq [\alpha'_1, \alpha'_2]'$.

The regressor x_{1it} may be assumed to have a non-zero mean such that

$$x_{1it} = c + m_{it},$$

where $c \neq 0$ and m_{it} is zero-mean stationary. But the value of the parameter c does not affect the asymptotic distributions of the IV estimators of α_1 and α_2, so it is assumed that $c = 0$.

The disturbance term u_{it} for model (1) is decomposed as

$$u_{it} = \mu_i + v_{it}, \tag{4}$$

where μ_i is an unobservable random variable which signifies individual effects and v_{it} denotes the remained disturbance. For v_{it}, assume an $AR(p_i)$ structure

$$v_{it} + \rho_{i1}v_{i(t-1)} + ... + \rho_{ip_i}v_{i(t-p_i)} = w_{it}, \tag{5}$$

where w_{it} is a white noise process with variance σ_w^2. We assume that all the roots of the characteristic equation $1 + \rho_{i1}z + ... + \rho_{ip_i}z^{p_i} = 0$ lie outside the unit circle for all i. This stationarity condition implies that

$$\mathfrak{s}_{p_i} = \sum_{k=0}^{p_i} \rho_{ik} > 0, \ (\rho_{i0} = 1).$$

Note that the autoregressive coefficients and orders are allowed to be heterogenous across individuals in specification (5).

In this paper, we are concerned with the case where

$$E(x_{1it}w_{is}) \neq 0 \text{ and } E(\Delta_{C_{x_{2i}}} x_{2it}w_{is}) \neq 0 \text{ for all } t \text{ and } s,$$

which necessitates the use of IV estimation. In addition, assume

Assumption 1 *μ_i is independent of x_{1jt},x_{2jt} and v_{jt} for all i, j and t.*

Assumption 2 $\mu_i \sim iid\ (0, \sigma_\mu^2)$.

Assumption 3 *When* $i \neq j$, $q_{it} \triangleq [x'_{1it}, \Delta_{C_{x_{2i}}} x'_{2it}, w_{it}]$ *is independent of* q_{js} *for all* t *and* s.

Assumptions 1 and 2 imply independence of the individual effect variable μ_i, which is usually assumed in the panel data analysis. But these assumptions are not required for the Within estimators, because the individual effect variable μ_i is eliminated for these estimators prior to estimation. Assumption 3 is a usual assumption in panel data analysis. Under this assumption, we may apply the central limit theorem and the law of large numbers by sending N to infinity.

3 Asymptotic Properties of IV Estimators

There are various IV estimators which can be considered for the error component model (1). Natural candidates are the IV, IV-OLS, IV-GLS, Within-IV, Within-IV-OLS and Within-IV-GLS estimators. The IV-OLS and IV-GLS estimators are IV estimators for correlated disturbances as discussed in Bowden and Turkington (1984). The Within-IV-OLS and Within-IV-GLS estimators are similar to these estimators, respectively, except that the model (1) is transformed prior to estimation. Among the various estimators, this section will focus on the Within-IV, Within-IV-OLS, IV-GLS and Within-IV-GLS estimators, because it is apparently deduced from the results in Choi (1998) that the IV and IV-OLS estimators are less efficient than the four estimators. The Within and GLS estimators have been regarded as standard tools for estimating stationary error component models (cf. Baltagi, 1995). The estimators considered here are extensions of these estimators to the case where regressors and disturbances are correlated and where nearly nonstationary regressors are present.

3.1 Within-IV Regression

The Within-IV regression subtracts individual effects from a given model and runs IV regression on this transformed model. The transformed model for model (1) is

$$y_{it} - \bar{y}_{i.} = \alpha_1'(x_{1it} - \bar{x}_{1i.}) + \alpha_2'(x_{2it} - \bar{x}_{2i.}) + u_{it} - \bar{u}_{i.}, \tag{6}$$

where $\bar{q}_{i.} \triangleq \frac{1}{T}\sum_{t=1}^{T} q_{it}$. Because $u_{it} - \bar{u}_{i.} = v_{it} - \bar{v}_{i.}$, the individual effect variable μ_i is eliminated for each individual in model (6).

Suppose that $I(0)$ vector of size l_1, z_{1it}, and nearly nonstationary vector of size l_2, z_{2it}, (i.e., $\Delta_{C_{z_{2i}}} z_{2it} = I(0)$) are available as instruments, where $l_1 \geq k_1$ and $l_2 \geq k_2$ so that there are sufficient number of instruments. Then, the IV estimator for the coefficient vector $\alpha = [\alpha_1, \alpha_2]$ is defined as

$$\hat{\alpha}_{WIV} \triangleq \left(\bar{X}' P_{\bar{z}} \bar{X}\right)^{-1} \bar{X}' P_{\bar{z}} \bar{y},$$

where the matrix $\bar{X}$ and $\bar{Z}$ are formed by stacking the row vectors $[x'_{1it} - \bar{x}'_{1i.}, x'_{2it} - \bar{x}'_{2i.}]$ and $[z'_{1it} - \bar{z}'_{1i.}, z'_{2it} - \bar{z}'_{2i.}]$, respectively, with the sequence of indices $(i,t) = (1,1), (1,2), ..., (N, T-1), (N,T)$. We also let

$$\bar{Z} \triangleq \left[\begin{array}{ccc} \bar{Z}'_1 & \cdots & \bar{Z}'_N \end{array} \right]' \text{ and } \bar{Z}_i \triangleq \left[\begin{array}{cc} \bar{Z}_{1i} & \bar{Z}_{2i} \end{array} \right].$$

Similar notation will be used for X, y and $v \triangleq [v_{11}, v_{12}, ..., v_{N(T-1)}, v_{NT}]'$.

Using $z_{1it} - \bar{z}_{1i.}$ rather than z_{1it} as instruments allows us to assume that $E(z_{1it}) = 0$ for all i without loss of generality. Moreover, using $z_{2it} - \bar{z}_{2i.}$ as instruments makes the Within-IV estimator invariant to the initial variables z_{2i0}.

To derive the asymptotic distribution of $\hat{\alpha}_{WIV}$, we make the following assumption regarding the regressors, instruments and disturbance terms.

Assumption 4 $\left[x'_{1it},\ z'_{1it},\ \Delta_{C_{x_{2i}}} x'_{2it},\ \Delta_{C_{z_{2i}}} z'_{2it} \right]'$ *is a vector stationary sequence which satisfies the following sample moment conditions for all i as $T \to \infty$.*

$$(a)\ \frac{1}{T} \Sigma_{t=1}^T (x_{1it} - \bar{x}_{1i.})(z_{1it} - \bar{z}_{1i.})' \xrightarrow{p} E\left(x_{1i0} z'_{1i0} \right). \quad (b)\ \frac{1}{T^{3/2}} \Sigma_{t=1}^T (x_{1it} - \bar{x}_{1i.})(z_{2it} - \bar{z}_{2i.})' \xrightarrow{p} 0.$$

$$(c)\ \frac{1}{T^{3/2}} \Sigma_{t=1}^T (x_{2it} - \bar{x}_{2i.})(z_{1it} - \bar{z}_{1i.})' \xrightarrow{p} 0.$$

$$(d) \frac{1}{T^2} \Sigma_{t=1}^T (x_{2it} - \bar{x}_{2i.})(z_{2it} - \bar{z}_{2i.})' \Rightarrow \int_0^1 \bar{K}_{x_{2i}}(r) \bar{K}'_{z_{2i}}(r) dr,$$

where $\bar{K}_{a_{2i}}(r) = K_{a_{2i}}(r) - \int_0^r K_{a_{2i}}(s) ds$ $(a = x, z)$, $K_{a_{2i}}(r)$ is a vector diffusion process satisfying the stochastic differential equation system $dK_{a_{2i}}(r) = C_{a_{2i}} K_{a_{2i}}(r) dr + dB_{a_{2i}}(r)$ and $\frac{1}{\sqrt{T}} \Sigma_{t=1}^{[Tr]} \Delta_{C_{a_{2i}}} a_{it} \Rightarrow B_{a_{2i}}(r)$.

$$(e)\ \frac{1}{T} \Sigma_{t=1}^T (z_{1it} - \bar{z}_{1i.})(z_{1it} - \bar{z}_{1i.})' \xrightarrow{p} E\left(z_{1i0} z'_{1i0} \right). \quad (f)\ \frac{1}{T^{3/2}} \Sigma_{t=1}^T (z_{1it} - \bar{z}_{1i.})(z_{2it} - \bar{z}_{2i.})' \xrightarrow{p} 0.$$

$$(g)\ \frac{1}{T^2} \Sigma_{t=1}^T (z_{2it} - \bar{z}_{2i.})(z_{2it} - \bar{z}_{2i.})' \Rightarrow \int_0^1 \bar{K}_{z_{2i}}(r) \bar{K}'_{z_{2i}}(r) dr.$$

$$(h)\ \frac{1}{\sqrt{T}} \Sigma_{t=1}^T (z_{1it} - \bar{z}_{1i.}) v_{it} \Rightarrow N\left(0, \Sigma_{h=-\infty}^{\infty} E\left(z_{1i0} z'_{1ih} v_{i0} v_{ih} \right) \right),$$

where the infinite series on the right-hand-side converges.

$$(i)\ \frac{1}{T} \Sigma_{t=1}^T (z_{2it} - \bar{z}_{2i.}) v_{it} \Rightarrow \int_0^1 \bar{K}_{z_{2i}}(r) dB_{v_i}(r) dr,$$

where $\frac{1}{\sqrt{T}} \Sigma_{t=1}^{[Tr]} v_{it} \Rightarrow B_{v_i}(r)$ and $\bar{K}_{z_{2i}}(r)$ is independent of $B_{v_i}(r)$.

Remarks:

(a) Assumption 4 is a general set of assumptions for Theorem 1 below to hold. Assumptions at this level of generality seem to be useful in that a broad class of random sequences can be shown to satisfy these assumptions and, therefore, conditions for Theorem 1. For example, linear processes and mixing processes under proper conditions satisfy Assumption 4.

(b) From the law of large numbers for stationary sequences[1] and the weak convergence results in Phillips (1988), it is obvious that parts (a)-(g) of Assumption 4 hold as $T \to \infty$ under proper moment conditions on x_{1it}, z_{1it}, $\Delta_{C_{z_{2i}}} x_{2it}$ and $\Delta_{C_{z_{2i}}} z_{2it}$.

(c) Parts (h) and (i) require that for all i

$$E(z_{1it}v_{it}) = 0 \text{ for all } t, \tag{7}$$

and

$$E(\Delta_{C_{z_{2i}}} z_{2it} v_{is}) = 0 \text{ for all } t \text{ and } s \tag{8}$$

respectively, and that z_{1it}, v_{it} and $\Delta_{C_{z_{2i}}} z_{2it}$ satisfy proper moment conditions.[2] Condition (7) indicates that lags of z_{1t} can be used as instruments. However, condition (8) requires that z_{2it} be strictly exogenous. Moreover, under condition (8), $\bar{K}_{z_{2i}}(r)$ is independent of $dB_{v_i}(r)$.

For sequential limit results, Assumption 4 suffices. However, for joint limit results, additional moment conditions in the following assumption are required to meet the conditions for Corollary 1 and Theorem 3 in Phillips and Moon (1998).

Assumption 5 *Let $\varepsilon > 0$. For all h, j and i,*

(a) $\sup_T \sup_{1 \le t,s \le T} E \mid [x_{1it} z'_{1is}]_{(h,j)} \mid^{1+\varepsilon} < \infty$;

(b) $\sup_T \sup_{1 \le t,s \le T} E \mid [z_{1it} z'_{1is}]_{(h,j)} \mid^{1+\varepsilon} < \infty$;

(c) $\sup_T \sup_{1 \le t,s \le T} E \mid \left[\frac{1}{\sqrt{s}} x_{1it} z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty$;

(d) $\sup_T \sup_{1 \le t,s \le T} E \mid \left[\frac{1}{\sqrt{t}} x_{2it} z'_{1is}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty$;

(e) $\sup_T \sup_{1 \le t,s \le T} E \mid \left[\frac{1}{\sqrt{s}} z_{1it} z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty$;

(f) $\sup_T \sup_{1 \le t,s \le T} E \mid \left[\frac{1}{\sqrt{ts}} z_{2it} z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty$.

(g) $\sup_T \sup_{1 \le t,s \le T} E \mid \left[\frac{1}{\sqrt{ts}} x_{2it} z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty$;

(h)-(1)

$$E(z_{1it}v_{is}) = 0 \textit{ for all } t \textit{ and } s;$$

(h)-(2) As $T \to \infty$,

$$E\left(\frac{1}{\sqrt{T}}\Sigma_{t=1}^{T}(z_{1it} - \bar{z}_{1i.})v_{it}\right)\left(\frac{1}{\sqrt{T}}\Sigma_{t=1}^{T}(z_{1it} - \bar{z}_{1i.})v_{it}\right)' \to \Sigma_{h=-\infty}^{\infty} E\left(z_{1i0} z'_{1ih} v_{i0} v_{ih}\right),$$

where the infinite series on the right-hand-side converges and is positive definite;

---

[1] See, for example, Chapter 4 of Hannan (1970) for the laws of large numbers for stationary sequences.

[2] See, for example, Chapter 4 of Hannan (1970) and Chapter 6 of Fuller (1976) for the central limit theorems which give the results similar to part (h) in Assumption 4 . Lemma 1 in Choi (1998) presents a set of conditions for part (h) in Assumption 4 when z_{1it} and v_{it} are stationary linear processes.

(h)-(3)

$$\lim_{N,T\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}E\left(\frac{1}{\sqrt{T}}\Sigma_{t=1}^{T}(z_{1it}-\bar{z}_{1i.})v_{it}\right)\left(\frac{1}{\sqrt{T}}\Sigma_{t=1}^{T}(z_{1it}-\bar{z}_{1i.})v_{it}\right)'=\Xi_1>0;$$

(i)-(1)

$$E(\Delta_{C_{z_{2i}}}z_{2it}v_{is})=0 \textit{ for all } t \textit{ and } s;$$

(i)-(2) As $T\to\infty$,

$$\begin{aligned}E\left(\frac{1}{T}\Sigma_{t=1}^{T}(z_{2it}-\bar{z}_{2i.})v_{it}\right)\left(\frac{1}{T}\Sigma_{t=1}^{T}(z_{2it}-\bar{z}_{2i.})v_{it}\right)' &\to \sigma_{v_i}^2E\left(\int_0^1\bar{K}_{z_{2i}}(r)\bar{K}'_{z_{2i}}(r)dr\right)\\ &= \sigma_{v_i}^2\int_0^1\bar{\Gamma}_i(z_2,z_2,r,r)dr>0,\end{aligned}$$

where $\sigma_{v_i}^2=\Sigma_{h=-\infty}^{\infty}E\left(v_{i0}v_{ih}\right)$ *and* $\bar{\Gamma}_i(z_2,z_2,r,r)$ *is defined in Lemma 2;*

(i)-(3)

$$\lim_{N,T\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}E\left(\frac{1}{T}\Sigma_{t=1}^{T}(z_{2it}-\bar{z}_{2i.})v_{it}\right)\left(\frac{1}{T}\Sigma_{t=1}^{T}(z_{2it}-\bar{z}_{2i.})v_{it}\right)'=\Xi_2>0.$$

Remarks: (a) Parts (a)-(g), (h)-(2) and (i)-(2) are for the uniform integrability condition in Corollary 1 and Theorem 3 in Phillips and Moon (1998), which is required to show that sequential and joint limits are identical.

(b) Under (h)-(1) and (i)-(1), $E\left(\frac{1}{\sqrt{T}}\Sigma_{t=1}^{T}(z_{1it}-\bar{z}_{1i.})v_{it}\right)=0$ and $E\left(\frac{1}{T}\Sigma_{t=1}^{T}(z_{2it}-\bar{z}_{2i.})v_{it}\right)=0$, respectively, so that the central limit theorem can be applied to these quantities across i.

(c) The convergence part in (h)-(2) can be proven along the line of Fuller (1976, p. 232) when more specific conditions on z_{1it} and v_{it} are given. Moreover, the convergence part in (i)-(2) can also be proven, for example, when $\Delta_{C_{z_{2i}}}z_{2it}$ and v_{it} are assumed to be linear processes.

(d) Conditions (h)-(3) and (i)-(3) are made to meet condition (iv) of Theorem 3 in Phillips and Moon (1998). These conditions essentially amount to the positive definiteness of limiting covariance matrices and, therefore, do not seem to be restrictive.

Additionally, assume

Assumption 6 *(a) The limits in parts (b) and (c) exist.*

(b)

$$rank\left(\begin{bmatrix}\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}E\left(x_{1i0}z'_{1i0}\right) & 0\\ 0 & \lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\int_0^1\bar{\Gamma}_i(x_2,z_2,r,r)dr\end{bmatrix}\right)=k_1+k_2,$$

where $\bar{\Gamma}_i(\cdot,\cdot,\cdot,\cdot)$ *is defined in Lemma 2 below.*

(c)

$$rank\left(\begin{bmatrix}\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}E\left(z_{1i0}z'_{1i0}\right) & 0\\ 0 & \lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\int_0^1\bar{\Gamma}_i(z_2,z_2,r,r)dr\end{bmatrix}\right)=l_1+l_2.$$

Under this assumption, the limiting distributions in Theorem 1 are well defined.

The following two lemmas will be used to evaluate the means and variances of the weak limits in Assumption 4.

Lemma 1 *Let* $E\left(B_{x_{2i}}(r)B'_{z_{2i}}(t)\right) = \Omega_{xz_i}(r \wedge t)$. *Then,*

$$\begin{aligned} E\left(K_{x_{2i}}(r)K'_{z_{2i}}(t)\right) &= \Omega_{xz_i}(r \wedge t) + \Omega_{xz_i}\int_0^t (r \wedge s)e^{(t-s)C'_{z2i}}dsC'_{z_{2i}} \\ &+C_{x_{2i}}\int_0^r e^{(r-s)C_{x2i}}(t \wedge s)ds\Omega_{xz_i} \\ &+C_{x_{2i}}\int_0^r\int_0^t e^{(r-s)C_{x2i}}\Omega_{xz_i}(s \wedge u)e^{(t-u)C'_{z2i}}dsduC'_{z_{2i}} \\ &\triangleq \Gamma_i(x_2, z_2, r, t). \end{aligned}$$

Lemma 2 *(a)*

$$\begin{aligned} E\left(\bar{K}_{x_{2i}}(r)\bar{K}'_{z_{2i}}(r)dr\right) &= \Gamma_i(x_2, z_2, r, r) - \int_0^r \Gamma_i(x_2, z_2, s, r)ds \\ &-\int_0^r \Gamma_i(x_2, z_2, r, s)ds + \int_0^r\int_0^r \Gamma_i(x_2, z_2, s, t)dsdt \\ &\triangleq \bar{\Gamma}_i(x_2, z_2, r, r). \end{aligned}$$

(b)

$$E\left(\bar{K}_{z_{2i}}(r)\bar{K}'_{z_{2i}}(r)dr\right) = \bar{\Gamma}_i(z_2, z_2, r, r).$$

Remarks: Obviously, every element of the matrices $\Gamma_i(x_2, z_2, r, t)$, $\bar{\Gamma}_i(x_2, z_2, r, r)$ and $\bar{\Gamma}_i(z_2, z_2, r, r)$ is finite. This fact will be used when the law of large numbers and the central limit theorem are applied to quantities involving $\bar{K}_{x_{2i}}(r)$ and $\bar{K}_{z_{2i}}(r)$.

Now, we report asymptotic distributions of the Within-IV in the following theorem.

Theorem 1 *(a) Suppose that Assumptions 3, 4, 6 hold and that the limits in parts (i) and (ii) exist. Then, as* $T \to \infty$ *and* $N \to \infty$ *sequentially,*

(i)

$$\sqrt{NT}(\hat{\alpha}_{1WIV} - \alpha_1) \Rightarrow N\left(0, A\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^N\Sigma_{h=-\infty}^{\infty}E\left(z_{1i0}z'_{1ih}v_{i0}v_{ih}\right)A'\right),$$

where

$$A \triangleq \left(\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^N E\left(x_{1i0}z'_{1i0}\right)\left(\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^N E\left(z_{1i0}z'_{1i0}\right)\right)^{-1}\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^N E\left(z_{1i0}x'_{1i0}\right)\right)^{-1}$$
$$\times\left(\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^N E\left(x_{1i0}z'_{1i0}\right)\right)\left(\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^N E\left(z_{1i0}z'_{1i0}\right)\right)^{-1};$$

(ii)

$$\sqrt{N}T(\hat{\alpha}_{2WIV}-\alpha_2) \Rightarrow N\left(0, B\left(\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\sigma_{v_i}^2\int_0^1 \bar{\Gamma}_i(z_2,z_2,r,r)dr\right)B'\right)$$

where $\sigma_{v_i}^2=\Sigma_{h=-\infty}^{\infty}E\left(v_{i0}v_{ih}\right)$ and, letting $\Psi(a,b)\triangleq\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\int_0^1\bar{\Gamma}_i(a,b,r,r)dr$,

$$B\triangleq\left(\Psi(x_2,z_2)\Psi(z_2,z_2)^{-1}\Psi(z_2,x_2)\right)^{-1}\Psi(x_2,z_2)\Psi(z_2,z_2)^{-1}.$$

(b) The results in part (a) hold as $N,T\to\infty$ simultaneously under Assumptions 3, 4, 5 and 6.

Remarks

(a) This theorem shows that the Within-IV estimator is consistent and has a normal distribution in the limit.

(b) The estimator $\hat{\alpha}_{2WIV}$ is $\sqrt{N}T$-consistent, while $\hat{\alpha}_{1WIV}$ is $\sqrt{NT}$-consistent. The $\sqrt{N}T$-consistency of course reflects the presence of nearly $I(1)$ regressors.

(c) For the joint limit results in part (b), we do not require a condition controlling the expansion rates of N and T unlike in Phillips and Moon (1998). This is because $E\left(\frac{1}{\sqrt{T}}\Sigma_{t=1}^{T}(z_{1it}-\bar{z}_{1i.})v_{it}\right)=0$ and $E\left(\frac{1}{T}\Sigma_{t=1}^{T}(z_{2it}-\bar{z}_{2i.})v_{it}\right)=0$ under Assumption 5 (h)-(1) and (i)-(1), so that there are no biases which have to be controlled by a condition on the expansion rates of N and T.

3.2 Within-IV-OLS Regression

The Within-IV-OLS estimator is obtained by applying the IV-OLS regression to model (6). Note that the Within-IV-OLS estimator is equivalent to the Within-IV estimator when the number of instruments is equal to that of regressors. Let the variance-covariance matrix of $[v_{i1}-\bar{v}_{i.},...,v_{iT}-\bar{v}_{i.}]'$ be

$$\breve{\sigma}_w^2\bar{V}_i\triangleq\breve{\sigma}_w^2\begin{bmatrix}1 & \gamma_{i1} & \cdot & \cdot & \gamma_{i(T-1)}\\ \gamma_{i1} & 1 & & & \\ \cdot & & \cdot & & \\ \cdot & & & \cdot & \gamma_{i1}\\ \gamma_{i(T-1)} & \cdot & \cdot & \gamma_{i1} & 1\end{bmatrix},\ \breve{\sigma}_w^2\triangleq\sigma_w^2-\frac{1}{T}\sigma_w^2$$

and let $\bar{V}$ be a matrix whose block-diagonal components are composed of $\bar{V}_1$, ... , $\bar{V}_N$ with the rest of elements being equal to zero. Then, the Within-IV-OLS estimator is defined as

$$\hat{\alpha}_{WIVO}\triangleq\left(\bar{X}'\bar{Z}(\bar{Z}'\bar{V}\bar{Z})^{-1}\bar{Z}'\bar{X}\right)^{-1}\bar{X}'\bar{Z}(\bar{Z}'\bar{V}\bar{Z})^{-1}\bar{Z}'\bar{y}. \tag{9}$$

To derive the asymptotic distribution of the Within-IV-OLS estimator, the following assumption is required.

Assumption 7 *(a) Parts (a), (b), (c), (d), (h) and (i) of Assumption 4 hold.*

For all i, as $T \to \infty$,

(b)

$$\frac{1}{T}\bar{Z}'_{1i}\bar{V}_i\bar{Z}_{1i} \xrightarrow{p} E\left(z_{1i0}z'_{1i0}\right) + 2\Sigma_{h=1}^{\infty}E\left(z_{1i0}z'_{1ih}\right)\gamma_{ih},$$

where the infinite series on the right-hand-side is assumed to converge;

(c)

$$\frac{1}{T^{3/2}}\bar{Z}'_{1i}\bar{V}_i\bar{Z}_{2i} \xrightarrow{p} 0;$$

and

(d)

$$\frac{1}{T^2}\bar{Z}'_{2i}\bar{V}_i\bar{Z}_{2i} \Rightarrow \left(1 + 2\Sigma_{h=1}^{\infty}\gamma_{ih}\right)\int_0^1 \bar{K}_{z_{2i}}(r)\bar{K}'_{z_{2i}}(r)dr.$$

Under proper moment conditions, part (b) of Assumption 7 holds by the law of large numbers for stationary sequences and parts (c) and (d) by the weak convergence results in Phillips (1988). In part (d), note that $\Sigma_{h=1}^{\infty}\gamma_{ih}$ converges due to the given autoregressive structure.

Additionally, assume

Assumption 8 *(a) The limits in part (b) of Assumption 6 and in part (c) below exist.*

(b) Part (b) of Assumption 6 hold.

(c) Let $\pi_i \triangleq E\left(z_{1i0}z'_{1i0}\right) + 2\Sigma_{h=1}^{\infty}E\left(z_{1i0}z'_{1ih}\right)\gamma_{ih}$. Then,

$$\begin{aligned} & rank\left(\begin{bmatrix} \lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\pi_i & 0 \\ 0 & \lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\left(1+2\Sigma_{h=1}^{\infty}\gamma_{ih}\right)\int_0^1\bar{\Gamma}_i(z_2,z_2,r,r)dr \end{bmatrix}\right) \\ = \;& l_1 + l_2. \end{aligned}$$

This assumption is required for the limiting distributions in Theorem 2 to be well defined.

The asymptotic distribution of the Within-IV-OLS estimator is reported in the following theorem.

Theorem 2 *(a) Suppose that Assumptions 3, 7 and 8 hold and the limits in parts (i) and (ii) exist. Then, as $T \to \infty$ and $N \to \infty$ sequentially,*

(i)

$$\sqrt{NT}(\hat{\alpha}_{1WIVO} - \alpha_1) \Rightarrow N\left(0, C\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\Sigma_{h=-\infty}^{\infty}E\left(z_{1i0}z'_{1ih}v_{i0}v_{ih}\right)C'\right),$$

where

$$C \triangleq \left(\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}E\left(x_{1i0}z_{1i0}'\right)\left(\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\pi_i\right)^{-1}\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}E\left(z_{1i0}x_{1i0}'\right)\right)^{-1}$$
$$\times\left(\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}E\left(x_{1i0}z_{1i0}'\right)\right)\left(\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\pi_i\right)^{-1},$$
$$\pi_i \triangleq E\left(z_{1i0}z_{1i0}'\right)+2\Sigma_{h=1}^{\infty}E\left(z_{1i0}z_{1ih}'\right)\gamma_{ih};$$

(ii)

$$\sqrt{N}T(\hat{\alpha}_{2WIVO}-\alpha_2)$$
$$\Rightarrow\ N\left(0,D\left(\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\sigma_{v_i}^2\int_0^1\bar{\Gamma}_i(z_2,z_2,r,r)dr\right)D'\right)$$

where $\sigma_{v_i}^2 \triangleq \Sigma_{h=-\infty}^{\infty}E\left(v_{i0}v_{ih}\right)$ and, letting $\Theta(a,b)\triangleq\lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}(1+2\Sigma_{h=1}^{\infty}\gamma_{ih})\int_0^1\bar{\Gamma}_i(a,b,r,r)dr$,

$$D\triangleq\left(\Psi(x_2,z_2)\Theta(z_2,z_2)^{-1}\Psi(z_2,x_2)\right)^{-1}\Psi(x_2,z_2)\Theta(z_2,z_2)^{-1}.$$

(b) The results in part (a) hold as $N,T\to\infty$ simultaneously under Assumptions 3, 5, 7 and 8.

Remarks: (a) The Within-IV-OLS estimator has essentially the same properties as the Within-IV estimator except that their variance-covariance matrices are different.

(b) When $\Sigma_{h=1}^{\infty}\gamma_{ih}$ are the same across i[3], part (a)-(ii) shows that $\hat{\alpha}_{2WIV}$ and $\hat{\alpha}_{2WIVO}$ have the same asymptotic distribution.

3.3 IV-GLS Regression

This subsection considers the IV-GLS regression. The structure of the variance-covariance matrix of the disturbance term u_{it} has added complexity due to the presence of serial correlation in v_{it}. In the literature of panel data analysis, it is usually assumed that v_{it} has the same autoregressive coefficients and order across all individuals. However, it is assumed here that each individual has different autoregressive coefficients and orders as specified in equation (5).

To calculate the GLS estimator, the inverse of the variance-covariance matrix of the disturbance term u_{it} is required. However, letting $u_i\triangleq[u_{i1},...,u_{iT}]'$, $v_i\triangleq[v_{i1},...,v_{iT}]'$, $V_i\triangleq E(v_iv_i')/\sigma_w^2$ and $\Phi_i\triangleq E(u_iu_i')=\sigma_\mu^2 i_Ti_T'+\sigma_w^2V_i$, we have as in Choi (1998)

$$\Phi_i^{-1}=\frac{V_i^{-1}}{\sigma_w^2}-\frac{\sigma_\mu^2V_i^{-1}i_Ti_T'V_i^{-1}}{\sigma_w^4+\sigma_\mu^2\sigma_w^2i_T'V_i^{-1}i_T},\tag{10}$$

where

$$V_i^{-1}=A_iA_i'-B_iB_i'=A_i'A_i-B_i'B_i,\tag{11}$$

[3] This holds when $\mathfrak{s}_{p_i}$ are the same across i, because $1+2\Sigma_{h=1}^{\infty}\gamma_{ih}=\sigma_{v_i}^2=\frac{\sigma_w^2}{\mathfrak{s}_{p_i}^2}$.

$$A_i \triangleq \begin{bmatrix} \rho_{i0} & 0 & . & . & 0 \\ . & & . & . & \\ . & & . & . & \\ \rho_{i(T-1)} & \rho_{i(T-2)} & . & . & \rho_{i0} \end{bmatrix} \text{ and } B_i \triangleq \begin{bmatrix} \rho_{iT} & 0 & . & . & 0 \\ . & & . & . & \\ . & & . & . & \\ \rho_{i1} & \rho_{i2} & . & . & \rho_{iT} \end{bmatrix}$$

with the convention that $\rho_{i0} = 1$ and $\rho_{ij} = 0$ for $j > p_i$. Equation (10), which uses Gohberg's formula (cf. Godolphin and Unwin, 1983), enables us to avoid a brute force inversion of the variance-covariance matrix. The conventional approach to serially correlated errors in panel data analysis has been to combine the Prais-Winsten method and the Wansbeek-Kapteyn method (cf. Wansbeek and Kapteyn, 1982, 1983). But the use of the Prais-Winsten method for $AR(p)$ processes requires p initial variables, which are complicated to obtain. Using equation (10), we can bypass such complications.

Equation (10) gives the inverse of the variance-covariance matrix for $u = [u_1', ..., u_N']'$ as

$$\Phi^{-1} = [E(uu')]^{-1} = diag[\Phi_1^{-1}, ..., \Phi_N^{-1}]. \tag{12}$$

Then, using equation (12) and the notation for equations (2) and (3), we write the IV-GLS estimator for model (1) as

$$\begin{bmatrix} \hat{\beta}_{IVG} \\ \hat{\alpha}_{IVG} \end{bmatrix} \triangleq \left(F(X,Z)F(Z,Z)^{-1}F(Z,X)\right)^{-1} F(X,Z)F(Z,Z)^{-1}G(Z,y), \tag{13}$$

where

$$F(X,Z) \triangleq \begin{bmatrix} \sum_{i=1}^N i_T'\Phi_i^{-1}i_T & \sum_{i=1}^N i_T'\Phi_i^{-1}Z_i \\ \sum_{i=1}^N X_i'\Phi_i^{-1}i_T & \sum_{i=1}^N X_i'\Phi_i^{-1}Z_i \end{bmatrix}, G(Z,y) \triangleq \begin{bmatrix} \sum_{i=1}^N i_T'\Phi_i^{-1}y_i \\ \sum_{i=1}^N Z_i'\Phi_i^{-1}y_i \end{bmatrix}$$

and Z_i is formed by stacking the row vectors $\begin{bmatrix} z_{1it}' & z_{2it}' \end{bmatrix}$ $(t = 1, ..., T)$.

Formula (13) is cumbersome for both calculation and asymptotic analysis. Therefore, the following lemma considers simplifying the expressions in formula (13).

Lemma 3 *Let* $\mathfrak{l}_i' \triangleq \mathfrak{s}_{p_i} \overbrace{[\varsigma_0, ..., \varsigma_{p_i-1}, 1, ..., 1,]}^{T}$, $\mathfrak{m}_i' \triangleq \overbrace{[0, ..., 0, \mathfrak{t}_{p_i}, ..., \mathfrak{t}_1]}^{T}$, $\tilde{W}_i \triangleq W_i + \rho_{i1}W_i^{[1]} + ... + \rho_{ip_i}W_i^{[p_i]}$ *and* $\ddot{W}_i \triangleq \rho_{i1}W_i^{[T-1]} + ... + \rho_{ip_i}W_i^{[T-p_i]}$, *where* $\mathfrak{s}_x \triangleq \sum_{k=0}^x \rho_{ik}$, $\varsigma_x \triangleq \mathfrak{s}_x/\mathfrak{s}_{p_i}$, $\mathfrak{t}_x \triangleq \sum_{k=x}^{p_i} \rho_{ik}$, $W^{[x]} \triangleq \begin{bmatrix} 0 & ... & 0 & w_1' & ... & w_{T-x}' \end{bmatrix}'$ *when* $W = \begin{bmatrix} w_1' & ... & w_T' \end{bmatrix}'$ *and* $W_i = X_i, Z_i$ *or* y_i. *Then,*

(*a*) $i_T'\Phi_i^{-1}i_T = (\mathfrak{l}_i'\mathfrak{l}_i - \mathfrak{m}_i'\mathfrak{m}_i)/\sigma_w^2 - \tau_i(\mathfrak{l}_i'\mathfrak{l}_i - \mathfrak{m}_i'\mathfrak{m}_i)^2$

(*b*) $i_T'\Phi_i^{-1}W_i = (\mathfrak{l}_i'\tilde{W}_i - \mathfrak{m}_i'\ddot{W}_i)/\sigma_w^2 - \tau_i(\mathfrak{l}_i'\mathfrak{l}_i - \mathfrak{m}_i'\mathfrak{m}_i)(\mathfrak{l}_i'\tilde{W}_i - \mathfrak{m}_i'\ddot{W}_i)$

(*c*) $W_i'\Phi_i^{-1}Z_i = [\tilde{W}_i'\tilde{Z}_i - \ddot{W}_i'\ddot{Z}_i]/\sigma_w^2 - \tau_i(\mathfrak{l}_i'\tilde{W}_i - \mathfrak{m}_i'\ddot{W}_i)'(\mathfrak{l}_i'\tilde{Z}_i - \mathfrak{m}_i'\ddot{Z}_i)$

where $\tau_i = \sigma_\mu^2/\{\sigma_w^4 + \sigma_\mu^2\sigma_w^2(\mathfrak{l}_i'\mathfrak{l}_i - \mathfrak{m}_i'\mathfrak{m}_i)\}$.

To derive the asymptotic distributions of $\hat{\beta}_{IVG}$ and $\hat{\alpha}_{IVG}$, we make the following assumption for all i and for $a, b = 1, ..., p_i$.

Assumption 9 *(a)* $z_{i1}, ..., z_{ip_i} = O_p(1)$; $x_{i1}, ..., x_{ip_i} = O_p(1)$; $u_{i1}, ..., u_{ip_i} = O_p(1)$.

(b) $\sum_{t=1}^{T} x_{1it} = O_p(\sqrt{T})$; $\sum_{t=1}^{T} z_{1it} = O_p(\sqrt{T})$; $\sum_{t=1}^{T} w_{it} = O_p(\sqrt{T})$.

(c) $\frac{1}{T^{3/2}} \sum_{t=1}^{T} z_{2it} \Rightarrow \int_0^1 K_{z_{2i}}(r)dr$; $\frac{1}{T^{3/2}} \sum_{t=1}^{T} x_{2it} \Rightarrow \int_0^1 K_{x_{2i}}(r)dr$.

(d) $\frac{1}{T} \sum_{t=p_i+1}^{T} z_{1i(t-a)} x'_{1i(t-b)} \xrightarrow{p} \gamma_{xz_i}(| a-b |)$, *where* $\gamma_{xz_i}(| a-b |) \triangleq E(z_{1i(1-a)} x'_{1i(1-b)})$; $\frac{1}{T^{3/2}} \sum_{t=p_i+1}^{T} z_{1i(t-a)} x'_{2i(t-b)} \xrightarrow{p} 0$; $\frac{1}{T^{3/2}} \sum_{t=p_i+1}^{T} z_{2i(t-a)} x'_{1i(t-b)} \xrightarrow{p} 0$; $\frac{1}{T^2} \sum_{t=p_i+1}^{T} z_{2i(t-a)} x'_{2i(t-b)} \Rightarrow \int_0^1 K_{z_{2i}}(r) K'_{x_{2i}}(r)dr$.

(e) $\frac{1}{T} \sum_{t=p_i+1}^{T} z_{1i(t-a)} z'_{1i(t-b)} \xrightarrow{p} \gamma_{zz_i}(| a-b |)$, *where* $\gamma_{zz_i}(| a-b |) \triangleq E(z_{1i(1-a)} z'_{1i(1-b)})$; $\frac{1}{T^{3/2}} \sum_{t=p_i+1}^{T} z_{1i(t-a)} z'_{2i(t-b)} \xrightarrow{p} 0$; $\frac{1}{T^2} \sum_{t=p_i+1}^{T} z_{2i(t-a)} z'_{2i(t-b)} \Rightarrow \int_0^1 K_{z_{2i}}(r) K'_{z_{2i}}(r)dr$.

(f) $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left(z_{1it} + \rho_{i1} z_{1i(t-1)} + ... + \rho_{ip_i} z_{1i(t-p_i)} \right) w_{it} \Rightarrow N\left(0, \sigma_w^2 \delta_i(z_1, z_1)\right)$, *where* $\delta_i(z_1, z_1) \triangleq E\left(z_{1i1} + \rho_{i1} z_{1i0} + ... + \rho_{ip_i} z_{1i(1-p_i)}\right)\left(z_{1i1} + \rho_{i1} z_{1i0} + ... + \rho_{ip_i} z_{1i(1-p_i)}\right)'$;
$\frac{1}{T} \sum_{t=1}^{T} z_{2i(t-a)} w_{it} \Rightarrow \int_0^1 K_{z_{2i}}(r) dB_{w_i}(r)$, *where* $\frac{1}{\sqrt{T}} \Sigma_{t=1}^{[Tr]} w_{it} \Rightarrow B_{w_i}(r)$ *and* $K_{z_{2i}}(r)$ *is independent of* $B_{w_i}(r)$.

Additionally, the following assumption is made for joint limit results.

Assumption 10 *(a) Parts (a)-(g) of Assumption 5 hold.*

Let $\varepsilon > 0$. *For all* h, j *and* i,

(b) $\sup_T \sup_{1 \leq t \leq T} E \mid [s'_{1it}]_{(h)} \mid^{1+\varepsilon} < \infty$ $(s = x, z)$;

(c) $\sup_T \sup_{1 \leq t \leq T} E \mid \left[\frac{1}{\sqrt{t}} s'_{2it}\right]_{(h)} \mid^{1+\varepsilon} < \infty$ $(s = x, z)$;

(d)

$$\lim_{N,T \to \infty} \frac{1}{N} \Sigma_{i=1}^{N} E\left(i'_T \Phi_i^{-1} u_i\right)\left(i'_T \Phi_i^{-1} u_i\right)' = \xi > 0;$$

(e)-(1)

$$E(z_{1it} w_{is}) = 0 \text{ for all } t \text{ and } s;$$

(e)-(2) As $T \to \infty$,

$$E\left(\frac{1}{\sqrt{T}} Z'_{1iT} \Phi_i^{-1} u_i\right)\left(\frac{1}{\sqrt{T}} Z'_{1iT} \Phi_i^{-1} u_i\right)' \to \frac{1}{\sigma_w^2} \delta_i(z_1, z_1),$$

where $\delta_i(z_1, z_1)$, *defined in Assumption 9 (f), is positive definite;*

(e)-(3)

$$\lim_{N,T} \frac{1}{N} \Sigma_{i=1}^{N} E\left(\frac{1}{\sqrt{T}} Z'_{1iT} \Phi_i^{-1} u_i\right)\left(\frac{1}{\sqrt{T}} Z'_{1iT} \Phi_i^{-1} u_i\right)' = \Upsilon_1 > 0;$$

(f)-(1)

$$E(\Delta_{C_{z_{2i}}} z_{2it} w_{is}) = 0 \text{ for all } t \text{ and } s;$$

(f)-(2) As $T \to \infty$,

$$E\left(\frac{1}{T} Z'_{2iT} \Phi_i^{-1} u_i\right)\left(\frac{1}{T} Z'_{2iT} \Phi_i^{-1} u_i\right)' \to \frac{\mathfrak{s}_{p_i}^2}{\sigma_w^2} \int_0^1 \bar{\Gamma}_i(z_2, z_2, r, r)dr > 0,$$

where $\bar{\Gamma}_i(z_2, z_2, r, r)$ *is defined in Lemma 2;*

(f)-(3)

$$\lim_{N,T\to\infty} \frac{1}{N}\Sigma_{i=1}^{N} E\left(\frac{1}{T}Z'_{2iT}\Phi_i^{-1}u_i\right)\left(\frac{1}{T}Z'_{2iT}\Phi_i^{-1}u_i\right)' = \Upsilon_2 > 0.$$

Note that remarks similar to those after Assumption 5 also apply to this assumption.

For the asymptotic distributions in Theorem 3 to be well defined, we assume

Assumption 11 *(a) The limits in parts (b) and (c) exist.*

(b)

$$rank\left(\begin{bmatrix} \lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\delta_i(x_1,z_1) & 0 \\ 0 & \lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\mathfrak{s}_{p_i}^2\int_0^1\bar{\Gamma}_i(x_2,z_2,r,r)dr \end{bmatrix}\right) = k_1+k_2,$$

where $\bar{\Gamma}_i(\cdot,\cdot,\cdot,\cdot)$ *is defined in Lemma 2 and* $\delta_i(x_1, z_1)$ *in Theorem 3.*

(c)

$$\begin{aligned} & rank\left(\begin{bmatrix} \lim_{N\to\infty}\frac{1}{N}\Sigma_{i=1}^{N}\delta_i(z_1,z_1) & 0 \\ 0 & \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathfrak{s}_{p_i}^2\int_0^1\bar{\Gamma}_i(z_2,z_2,r,r)dr \end{bmatrix}\right) \\ = \; & l_1+l_2. \end{aligned}$$

Now, the asymptotic distribution of the GLS estimator is reported in the following theorem.

Theorem 3 *(a) Suppose that Assumptions 1, 2, 3, 9 and 11 hold. Then, as* $T\to\infty$ *and* $N\to\infty$ *sequentially,*

(i)

$$\sqrt{N}(\hat{\beta}_{IVG}-\beta) \Rightarrow N(0,\sigma_\mu^2);$$

(ii)

$$\sqrt{NT}(\hat{\alpha}_{1IVG}-\alpha_1) \Rightarrow N(0,\sigma_w^2 E^{-1}),$$

where

$$E \triangleq \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\delta_i(x_1,z_1)\left(\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\delta_i(z_1,z_1)\right)^{-1}\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\delta_i(z_1,x_1)$$

and $\delta_i(\cdot,\cdot)$ is defined in Assumption 9;

(iii)

$$\sqrt{N}T(\hat{\alpha}_{2IVG}-\alpha_2) \Rightarrow N(0,\sigma_w^2 G^{-1}),$$

where

$$\begin{aligned} G \triangleq & \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathfrak{s}_{p_i}^2\int_0^1\bar{\Gamma}_i(x_2,z_2,r,r)dr\left(\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathfrak{s}_{p_i}^2\int_0^1\bar{\Gamma}_i(z_2,z_2,r,r)dr\right)^{-1} \\ & \times\lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N}\mathfrak{s}_{p_i}^2\int_0^1\bar{\Gamma}_i(z_2,x_2,r,r)dr. \end{aligned}$$

(b) The results in part (a) hold as $N, T \to \infty$ simultaneously under Assumptions 1, 2, 3, 9, 10 and 11.

Remark: This theorem shows that the GLS estimator has essentially the same properties as the Within-IV and Within-IV-OLS estimators except that their variance-covariance matrices are different.

3.4 Within-IV-GLS Regression

This section considers the IV-GLS estimation of model (6) which was used for the Within-IV regression . In model (6), $u_{it} - \bar{u}_{i.} = v_{it} - \bar{v}_{i.}$, and hence

$$\begin{aligned} & (u_{it} - \bar{u}_{i.}) + \rho_{i1}(u_{i(t-1)} - \bar{u}_{i.}) + ... + \rho_{ip_i}(u_{i(t-p_i)} - \bar{u}_{i.}) \\ = \; & (v_{it} - \bar{v}_{i.}) + \rho_{i1}(v_{i(t-1)} - \bar{v}_{i.}) + ... + \rho_{ip_i}(v_{i(t-p_i)} - \bar{v}_{i.}) \\ = \; & w_{it} - \bar{w}_{i.}. \end{aligned}$$

Thus, the inverse of the variance-covariance matrix for $u_{it} - \bar{u}_{i.}$ is

$$(\breve{\sigma}_w^2 \bar{V}_i)^{-1} = \frac{1}{\breve{\sigma}_w^2}(A_i A_i' - B_i B_i'), \tag{14}$$

where $\breve{\sigma}_w^2 \triangleq \sigma_w^2 - \frac{1}{T}\sigma_w^2$ due to relation (11). Using equation (14), we write the Within-IV-GLS estimator as

$$\hat{\alpha}_{WIVG} \triangleq \left(\bar{F}(X,Z)\bar{F}(Z,Z)^{-1}\bar{F}(Z,X)\right)^{-1} \bar{F}(X,Z)\bar{F}(Z,Z)^{-1}\bar{G}(Z,y), \tag{15}$$

where $\bar{F}(X,Z) \triangleq \sum_{i=1}^N \bar{X}_i'\bar{V}_i^{-1}\bar{Z}_i$, $\bar{G}(Z,y) \triangleq \sum_{i=1}^N \bar{Z}_i'\bar{V}_i^{-1}\bar{y}_i$. This estimator can be calculated without inverting the matrix V_i by using relation (14).

As in previous sections, we assume the following for joint limit results.

Assumption 12 *(a) Parts (a)-(g) of Assumption 5 hold.*

For all i,

(b)-(1)

$$E(z_{1it}w_{is}) = 0 \text{ for all } t \text{ and } s;$$

(b)-(2) As $T \to \infty$,

$$E\left(\frac{1}{\sqrt{T}}\sum_{i=1}^N \bar{Z}_{1i}'\bar{V}_i^{-1}\bar{v}_i)\right)\left(\frac{1}{\sqrt{T}}\sum_{i=1}^N \bar{Z}_{1i}'\bar{V}_i^{-1}\bar{v}_i\right)' \to \delta_i(z_1, z_1),$$

where $\delta_i(z_1, z_1)$, defined in Assumption 9 (f), is positive definite;

(b)-(3)

$$\lim_{N,T\to\infty} \frac{1}{N}\Sigma_{i=1}^N E\left(\frac{1}{\sqrt{T}}\bar{Z}_{1i}'\bar{V}_i^{-1}\bar{v}_i\right)\left(\frac{1}{\sqrt{T}}\bar{Z}_{1i}'\bar{V}_i^{-1}\bar{v}_i\right)' = \Pi_1 > 0;$$

(c)-(1)

$$E(\Delta_{C_{z_{2i}}} z_{2it}w_{is}) = 0 \text{ for all } t \text{ and } s;$$

(c)-(2) As $T \to \infty$,

$$E\left(\frac{1}{T}\bar{Z}_{2i}'\bar{V}_i^{-1}\bar{v}_i\right)\left(\frac{1}{T}\bar{Z}_{2i}'\bar{V}_i^{-1}\bar{v}_i\right)' \to \mathfrak{s}_{p_i}^2\int_0^1 \bar{\Gamma}_i(z_2, z_2, r, r)dr > 0,$$

where $\bar{\Gamma}_i(z_2, z_2, r, r)$ *is defined in Lemma 2;*

(c)-(3)

$$\lim_{N,T\to\infty}\frac{1}{N}\Sigma_{i=1}^N E\left(\frac{1}{T}\bar{Z}_{2iT}'\bar{V}_i^{-1}\bar{v}_i\right)\left(\frac{1}{T}\bar{Z}_{2iT}'\bar{V}_i^{-1}\bar{v}_i\right)' = \Pi_2 > 0.$$

The following theorem reports the asymptotic distribution of the Within-GLS estimator.

Theorem 4 *(a) Suppose that Assumptions 3, 9 and 11 hold. Then, as* $T \to \infty$ *and* $N \to \infty$ *sequentially,* $\sqrt{NT}(\hat{\alpha}_{1WIVG} - \alpha_1)$ *and* $\sqrt{N}T(\hat{\alpha}_{2WIVG} - \alpha_2)$ *have the same asymptotic distributions as* $\sqrt{NT}(\hat{\alpha}_{1IVG} - \alpha_1)$ *and* $\sqrt{N}T(\hat{\alpha}_{2IVG} - \alpha_2)$*, respectively.*

(b) Suppose that Assumptions 3, 9, 11 and 12 hold. Then, as $N, T \to \infty$ *simultaneously,*$\sqrt{NT}(\hat{\alpha}_{1WIVG} - \alpha_1)$ *and* $\sqrt{N}T(\hat{\alpha}_{2WIVG} - \alpha_2)$ *have the same asymptotic distributions as* $\sqrt{NT}(\hat{\alpha}_{1IVG} - \alpha_1)$ *and* $\sqrt{N}T(\hat{\alpha}_{2IVG} - \alpha_2)$*, respectively.*

Remark: This theorem shows that the Within–IV-GLS estimator has the same asymptotic distribution as the IV-GLS estimator, which implies that using the Within-IV-GLS estimator does not entail any asymptotic efficiency loss relative to IV-GLS.

3.5 Efficiency Comparison

This subsection compares the asymptotic efficiency of the four estimators studied so far. Because $\hat{\alpha}_{WIVG}$ is asymptotically equivalent to $\hat{\alpha}_{IVG}$, only the Within-IV, Within-IV-OLS and IV-GLS estimators will be compared here.

First, $\hat{\alpha}_{1WIV}$ is asymptotically less efficient than $\hat{\alpha}_{1WIVO}$, because the Within-IV estimator disregards the correlated structure of disturbances and is known to be asymptotically less efficient than the corresponding IV-OLS estimator in stationary regression.[4]

Second, it is not obvious how to compare $\hat{\alpha}_{1IVG}$ with $\hat{\alpha}_{1WIV}$ and $\hat{\alpha}_{1WIVO}$, as is well known in stationary regression (cf. Bowden and Turkington, 1984). But these estimators will be compared through simulation in Section 6.

Third, when $\mathfrak{s}_{p_i}$ are the same for all i, Theorems 1,2 and 3 indicate that $\hat{\alpha}_{2IVG}$, $\hat{\alpha}_{2WIV}$ and $\hat{\alpha}_{2WIVO}$ are asymptotically equivalent.[5] However, in general, these estimators are not asymptotically equivalent, and it is not obvious from Theorem 1, 2 and 3 which estimator is most efficient. Section 6 will compare these estimators by simulation.

[4] See Chapter 3 of Bowden and Turkington (1984) and Section 5 of Newey and McFadden (1994).

[5] Note that $\sigma_{v_i}^2 = \frac{\sigma_w^2}{s_{p_i}^2}$.

3.6 Feasible Estimation

This subsection discusses methods to make the Within-IV-OLS, IV-GLS and Within-IV-GLS estimators feasible. For the feasible estimators, parameters $\sigma_\mu^2, \sigma_w^2, \rho_{i1}, ..., \rho_{ip_i}$ should be estimated. As discussed in Choi (1998), the method-of-moments estimator of σ_μ^2 is defined as

$$\hat{\sigma}_\mu^2 \triangleq \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\hat{u}_{it}^2 - \hat{v}_{it}^2), \tag{16}$$

where $\hat{u}_{it}$ denote residuals obtained from the IV regression on equation (1)[6] and $\hat{v}_{it}$ are residuals from the With-IV regression. Because $\hat{u}_{it}$ and $\hat{v}_{it}$ are consistent estimators of u_{it} and v_{it}, this estimator is also consistent. When $\hat{\sigma}_\mu^2$ is negative, replacing negative $\hat{\sigma}_\mu^2$ by zero has been a common practice (cf. Baltagi, p. 17, 1995).

To estimate parameters $\sigma_w^2, \rho_{i1}, ..., \rho_{ip_i}$ consistently, we take the following steps.
(a) Run the Within-IV regression and retrieve residuals $\hat{v}_{it}$.
(b) Run autoregressions of order p_i for each individual by using the residuals $\hat{v}_{it}$. This gives estimators $\hat{\rho}_{i1}, ..., \hat{\rho}_{ip_i}$.
(c) From the autoregressions in step (b), get the residuals $\hat{w}_{it}$ and let the estimator of σ_w^2 as

$$\hat{\sigma}_w^2 = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{w}_{it}^2.$$

It can be deduced form standard theory in time series analysis that

$$\hat{\rho}_{ij} \xrightarrow{p} \rho_{ij}(j = 1, ..., p_i); \hat{\sigma}_w^2 \xrightarrow{p} \sigma_w^2.$$

The feasible IV-OLS, IV-GLS and Within-IV-GLS estimators are obtained by plugging the estimators of $\sigma_\mu^2, \sigma_w^2, \rho_{i1}, ..., \rho_{ip_i}$ into formulae (9), (13) and (15), respectively.

It can be shown by using the methods in Choi (1998) that the GLS and feasible GLS estimators have the same sequential asymptotic distribution at no extra conditions. The details for this do not seem to deserve separate space here, because they are very similar to the proof of Theorem 5 in Choi (1998).[7]

4 Hypothesis Tests on Coefficient Vectors

This section considers Wald tests to test the null hypotheses

$$H_0 : R\alpha = r \tag{17}$$

and

$$H_0 : S\beta = s, \tag{18}$$

[6] The IV estimator of α using z_{1it} and z_{2it} as instruments is obviously consistent under proper conditions.

[7] Showing that the GLS and feasible GLS estimators have the same joint limit distribution seems to be more involved and complicated. But because it is not the main focus of this paper, it is relegated to future work.

where R is a $J \times k$ matrix of rank J , $\alpha = [\alpha_1', \alpha_2']'$, S is a $J \times (k+1)$ matrix of rank J and $\beta = [\beta_0, \alpha_1', \alpha_2']'$. The Wald tests for the null hypothesis (17) are formulated by using the Within-IV, feasible Within-IV-OLS and feasible Within-IV-GLS estimators, and the Wald test for (18) by using the feasible IV-GLS estimator.

The Wald test using the Within-IV estimator $\hat{\alpha}_{WIV}$ is defined as

$$W_{WIV} \triangleq (R\hat{\alpha}_{WIV} - r)' \left(RK \sum_{i=1}^{N} \left[\bar{Z}_i' \hat{U}_{iWIV} \right] \left[\bar{Z}_i' \hat{U}_{iWIV} \right]' K'R' \right)^{-1} (R\hat{\alpha}_{WIV} - r), \tag{19}$$

where

$$K \triangleq \left(\bar{X}' P_{\bar{z}} \bar{X} \right)^{-1} \bar{X}' \bar{Z} (\bar{Z}' \bar{Z})^{-1}$$

and $\hat{U}_{iWIV}$ denotes the i-th block residual vector from the Within-IV regression.

Likewise, the Wald test using the feasible Within-IV-OLS is defined as

$$W_{WIVO} \triangleq (R\hat{\alpha}_{FWIVO} - r)' \left(RQ_{\hat{V}} \sum_{i=1}^{N} \left[\bar{Z}_i' \hat{U}_{iFWIVO} \right] \left[\bar{Z}_i' \hat{U}_{iFWIVO} \right]' Q_{\hat{V}}' R' \right)^{-1} (R\hat{\alpha}_{FWIVO} - r),$$

where $\hat{V}$ is a consistent estimator of the variance-covariance matrix $\bar{V}$,

$$Q_{\hat{V}} \triangleq \left(\bar{X}' \bar{Z} (\bar{Z}' \hat{V} \bar{Z})^{-1} \bar{Z}' \bar{X} \right)^{-1} \bar{X}' \bar{Z} (\bar{Z}' \hat{V} \bar{Z})^{-1} \tag{20}$$

and $\hat{U}_{iFWIVO}$ denotes the i-th block residual vector from the feasible Within-IV-OLS regression.

Similarly, the Wald test using the feasible Within-IV-GLS estimator is

$$\begin{aligned} W_{WIVG} &\triangleq (R\hat{\alpha}_{FWIVG} - r)' \\ &\times \left(RM_{\hat{V}} \sum_{i=1}^{N} \left[\bar{Z}_i' \hat{V}_i^{-1} \hat{U}_{iFWIVG} \right] \left[\bar{Z}_i' \hat{V}_i^{-1} \hat{U}_{iFWIVG} \right]' M_{\hat{V}}' R' \right)^{-1} \\ &\times (R\hat{\alpha}_{FWIVG} - r), \end{aligned} \tag{21}$$

where, letting $\bar{F}_{\hat{V}}(\cdot, \cdot)$ denote $\bar{F}(\cdot, \cdot)$ with $\bar{V}_i$ replaced by $\hat{V}_i$ (a consistent estimator of $\bar{V}_i$ using $\hat{\alpha}_{FWIVG}$),

$$M_{\hat{V}} \triangleq \left(\bar{F}_{\hat{V}}(X, Z) \bar{F}_{\hat{V}}(Z, Z)^{-1} \bar{F}_{\hat{V}}(Z, X) \right)^{-1} \bar{F}_{\hat{V}}(X, Z) \bar{F}_{\hat{V}}(Z, Z)^{-1} \tag{22}$$

and $\hat{U}_{iFWIVG}$ is the i-th block residual vector from the feasible Within–IV-GLS regression.

In the same manner, the Wald test for hypothesis (18) which uses the feasible IV-GLS estimator is defined as

$$W_{IVG} = (S\hat{\beta}_{FIVG} - s)' \left(SM_{\hat{\Phi}} \sum_{i=1}^{N} \left[\begin{array}{c} i_T' \hat{\Phi}_i^{-1} \hat{U}_{iFIVG} \\ X_i' \hat{\Phi}_i^{-1} \hat{U}_{iFIVG} \end{array} \right] \left[\begin{array}{c} i_T' \hat{\Phi}_i^{-1} \hat{U}_{iFIVG} \\ X_i' \hat{\Phi}_i^{-1} \hat{U}_{iFIVG} \end{array} \right]' M_{\hat{\Phi}} S' \right)^{-1} (S\hat{\beta}_{FIVG} - s),$$

where, letting $F_{\hat{\Phi}}(\cdot,\cdot)$ denote $F(\cdot,\cdot)$ with Φ_i replaced by $\hat{\Phi}_i$,

$$M_{\hat{\Phi}} \triangleq \left(F_{\hat{\Phi}}(X,Z)F_{\hat{\Phi}}(Z,Z)^{-1}F_{\hat{\Phi}}(Z,X)\right)^{-1} F_{\hat{\Phi}}(X,Z)F_{\hat{\Phi}}(Z,Z)^{-1}$$

and $\hat{U}_{iFIVG}$ is the i-th block residual vector from the feasible IV-GLS regression.

In the following theorem, sequential asymptotic distributions of the Wald tests are reported.

Theorem 5 *Under the same assumptions as for parts (a) of Theorems 1, 2, 3, and 4, respectively,*

$$W_{WIV}, W_{WIVO}, W_{IVG}, W_{WIVG} \Rightarrow \chi^2_J.$$

5 Tests of Exogeneity

The IV estimation methods studied so far presume that the regressors and the regression errors are correlated. But in practice, this needs to be tested. Therefore, this section considers exogeneity tests which test the null hypothesis that x_{1it} and Δx_{2it} are uncorrelated with v_{it}[8] for all i. For this, we may use the DWH tests. The alternative hypothesis is that x_{1it} and Δx_{2it} are correlated with v_{it} for infinitely many i so that the Within and GLS estimators studied in Choi (1998) are inconsistent.

The DWH tests can be formulated by using the feasible Within-GLS estimator $\hat{\alpha}_{FWG}$ (cf. Choi, 1998) and the feasible Within-IV-GLS estimator $\hat{\alpha}_{FWIVG}$. The feasible Within-GLS and Within-IV-GLS estimators are both consistent under the null[9], but under the alternative only the feasible Within-IV-GLS estimator is consistent. Thus, the DWH tests using the difference of these two estimators can tell the null from the alternative. The DWH test is formally defined as

$$DWH_G \triangleq (\hat{\alpha}_{FWIVG} - \hat{\alpha}_{FWG})' \left(VC_{WIVG} - VC_{WG}\right)^{-1} (\hat{\alpha}_{FWIVG} - \hat{\alpha}_{FWG}), \quad (23)$$

where

$$VC_{WIVG} \triangleq M_{\hat{V}} \sum_{i=1}^{N} \left[\bar{Z}_i' \hat{V}_{iFWIVG}^{-1} \hat{U}_{iFWIVG}\right] \left[\bar{Z}_i' \hat{V}_{iFWIVG}^{-1} \hat{U}_{iFWIVG}\right]' M_{\hat{V}}',$$

$$VC_{WG} \triangleq M_{FWG} Q_{FWG} M_{FWG},$$

$$M_{FWG} \triangleq \left[\sum_{i=1}^{N} \bar{X}_i' \hat{V}_{iFWG}^{-1} \bar{X}_i\right]^{-1},$$

[8] One may also be interested in testing the null hypothesis $E(\mu_i \mid x_{1it}, \Delta x_{2it}) = 0$ as in Hausman and Taylor (1981). But when $T \to \infty$, $\hat{\alpha}_{WIV}$, $\hat{\alpha}_{WIVO}$, $\hat{\alpha}_{IVG}$ and $\hat{\alpha}_{WIVG}$ are all consistent regardless of $E(\mu_i \mid x_{1it}, \Delta x_{2it}) = 0$. Thus, the null hypothesis does not deserve our attention when $T \to \infty$.

[9] Strictly speaking, the Within-GLS estimator corresponding to nonstationary variables is consistent, but its asymptotic distribution diverges under the alternative which makes the DWH test consistent.

$$Q_{FWG} \triangleq \sum_{i=1}^{N} \left[\bar{X}_i' \hat{V}_{iFWG}^{-1} \hat{U}_{iFWG}\right] \left[\bar{X}_i' \hat{V}_{iFWG}^{-1} \hat{U}_{iFWG}\right]',$$

$M_{\hat{V}}$ is defined in equation (22), $\hat{V}_{iA}$ is a consistent estimator of $\bar{V}_i$ using $\hat{\alpha}_A$, and $\hat{U}_{iA}$ is the residual vector using $\hat{\alpha}_A$. We may also consider formulating the DWH tests by using the GLS and IV-GLS estimators. But because these estimators have the same asymptotic distributions as the Within-GLS and Within-IV-GLS estimators, respectively, the DWH test using the GLS and IV-GLS estimators has the same asymptotic local power as the DWH test defined above.

Alternately, one may combine $\hat{\alpha}_{WIV}$ with the Within estimator $\hat{\alpha}_W$ (cf. Choi, 1998) to formulate the DWH test. However, due to serially correlated errors, the asymptotic variance-covariance matrix of $\hat{\alpha}_{WIV} - \hat{\alpha}_W$ is not the difference of those for $\hat{\alpha}_{WIV}$ and $\hat{\alpha}_W$. Therefore, as studied in Choi and Yu (1999), the DWH test using $\hat{\alpha}_{WIV}$ and $\hat{\alpha}_W$ should take a form different from test (23). Choi and Yu (1999) term the DWH test for serially correlated errors a generalized DWH test. The generalized DWH test using $\hat{\alpha}_{WIV}$ and $\hat{\alpha}_W$ is defined as

$$GDWH_{WIV} \triangleq (\hat{\alpha}_{WIV} - \hat{\alpha}_W)' VC_{WIV}^{-1} (\hat{\alpha}_{WIV} - \hat{\alpha}_W),$$

where

$$VC_{WIV} \triangleq K \sum_{i=1}^{N} \left[\bar{Z}_i' \hat{U}_{iWIV}\right] \left[\bar{Z}_i' \hat{U}_{iWIV}\right]' K' - K \sum_{i=1}^{N} \left[\bar{Z}_i' \hat{U}_{iWIV}\right] \left[\bar{X}_i' \hat{U}_{iW}\right]' \left(\bar{X}'\bar{X}\right)^{-1}$$
$$- \left(\bar{X}'\bar{X}\right)^{-1} \sum_{i=1}^{N} \left[\bar{X}_i' \hat{U}_{iW}\right] \left[\bar{Z}_i' \hat{U}_{iWIV}\right]' K' + \left(\bar{X}'\bar{X}\right)^{-1} \sum_{i=1}^{N} \left[\bar{X}_i' \hat{U}_{iW}\right] \left[\bar{X}_i' \hat{U}_{iW}\right]' \left(\bar{X}'\bar{X}\right)^{-1},$$

matrix K is defined in equation (19), and $\hat{U}_{iW}$ denotes the i-th block residual vector from the Within regression. The generalized DWH test using $\hat{\alpha}_{FWIVO}$ and $\hat{\alpha}_W$ is similarly defined once matrix K is replaced by $Q_{\bar{V}}$ defined in (20), and is denoted as $GDWH_{WIVO}$.

Sequential asymptotic distributions of the DWH tests are reported in the following theorem.

Theorem 6 *Suppose that Assumption 9 (f) with z_1 and z_2 replaced by $x_1$and x_2, respectively, holds. Under the same assumptions as for parts (a) of Theorems 4, 1 and 2, respectively,*

$$DWH_G, GDWH_{WIV}, GDWH_{WIVO} \Rightarrow \chi^2_{k_1+k_2}.$$

6 Simulation

This section reports simulation results regarding the finite sample efficiency comparison of the estimators we have studied and the empirical size and power of the Wald and DWH tests. There are so many factors that should be considered in these experiments. Broadly speaking, the degree of instrument-regressor correlation, the degree of serial correlation in instruments, the number of instruments, the degree of regressor-error dependence, the degree of serial correlation and heterogeneity of errors, the structure of errors and the number of sample sizes should matter for the results in this section. But because large-scale simulation seems to be beyond the scope of this paper, we chose the degree of instrument-regressor correlation, the degree of serial correlation in instruments, degree of regressor-error dependence and the structure of errors to be fixed for the finite sample efficiency comparison and the Wald tests, though some variations of the degree of regressor-error dependence were consider for the DWH tests. For the number of instruments, only the cases of two instruments for each regressor and one instrument for each regressor were considered. Also, only two cases for the degree of serial correlation and heterogeneity of errors were examined. Moreover, errors were assumed to have $AR(1)$ structure, the order of which is assumed to be known. Last, numbers of sample sizes considered were $(N,T) = (30,50)$, $(30,100)$, $(30,200)$, $(80,50)$, $(80,100)$, $(80,200)$, $(150,50)$, $(150,100)$ and $(30,200)$.

The data for simulations in the case of two instruments for each regressor were generated by

$$y_{it} = x_{1it} + x_{2it} + u_{it};\ (i = 1, ..., N;\ t = 1, ..., T);\ N = 30, 80, 150;\ T = 50, 100, 200) \tag{24}$$

$$e_{it} \sim iid\ N(0,\Psi), \Psi = \begin{bmatrix} 1 & .7 & .7 & .7 & .7 & .7 & \tau \\ .7 & 1 & .7 & .7 & .7 & .7 & .0 \\ .7 & .7 & 1 & .7 & .7 & .7 & .0 \\ .7 & .7 & .7 & 1 & .7 & .7 & \tau \\ .7 & .7 & .7 & .7 & 1 & .7 & .0 \\ .7 & .7 & .7 & .7 & .7 & 1 & .0 \\ \tau & .0 & .0 & \tau & .0 & .0 & 1 \end{bmatrix}, \ \tau = .2; \tag{25}$$

$$x_{1it} = .5a_{1i}x_{1i(t-1)} + [e_{it}]_{(1)}\ ;\ z_{1it} = \begin{bmatrix} .5b_{11i} & 0 \\ .2b_{12i} & .5b_{11i} \end{bmatrix} z_{1i(t-1)} + \begin{bmatrix} [e_{it}]_{(2)} \\ [e_{it}]_{(3)} \end{bmatrix}; \tag{26}$$

$$\Delta x_{2it} = .5a_{2i}\Delta x_{2i(t-1)} + [e_{it}]_{(4)}\ ;\ \Delta z_{1it} = \begin{bmatrix} .5b_{21i} & 0 \\ .2b_{22i} & .5b_{21i} \end{bmatrix} \Delta z_{1i(t-1)} + \begin{bmatrix} [e_{it}]_{(5)} \\ [e_{it}]_{(6)} \end{bmatrix}; \tag{27}$$

$$\begin{aligned} v_{it} &= \rho a_{3i} v_{i(t-1)} + [e_{it}]_{(7)}\ ; \\ u_{it} &= \mu_i + v_{it}\ ;\ \mu_i \sim iid\ N(0,1), \end{aligned} \tag{28}$$

where $b_{j,ki} \sim iid\ U[0,1]$ $(j,k = 1,2)$ and $a_{ki} \sim iid\ U[0,1]$ $(k = 1,2,3)$. In this data generation, there are two instruments for each regressor and the instruments satisfy conditions for being instruments. Moreover, the regressors are correlated with the regression errors, which is signified by the value of the parameter τ. For the regressors, instruments and errors, we generated autoregressive series of time span $T + 30$ with initial value 0 and took the last T elements of the series. Parameter ρ signifies the degree of serial correlation in the errors and is chosen to either .5 or .7. Because the same seed number was used for each individual to generate the uniformly distributed numbers a_{3i}, v_{it} has definitely higher serial correlation when $\rho = .7$. Furthermore, when $\rho = .7$, there is more heterogeneity in the errors. The number of iterations was set at 3,000 when $N \leq 80$, and 1,000 when $N = 150$. Last, the covariance matrices for the regressors, instruments and errors are set to be heterogenous across i in experimental design (25).

The data for simulations in the case of one instruments for each regressor were generated similarly by replacing (25), (26), (27) and (28), respectively, with

$$e_{it} \sim iid\ N(0,\Psi), \Psi = \begin{bmatrix} 1 & .7 & .7 & .7 & \tau \\ .7 & 1 & .7 & .7 & .0 \\ .7 & .7 & 1 & .7 & \tau \\ .7 & .7 & .7 & 1 & .0 \\ \tau & .0 & \tau & .0 & 1 \end{bmatrix}, \tau = .2. \quad (29)$$

$$x_{1it} = .5a_{1i}x_{1i(t-1)} + [e_{it}]_{(1)}\ ;\ z_{1it} = .5b_{11i}z_{1i(t-1)} + [e_{it}]_{(2)}\ ;$$
$$\Delta x_{2it} = .5a_{2i}\Delta x_{2i(t-1)} + [e_{it}]_{(3)}\ ;\ \Delta z_{1it} = .5b_{21i}\Delta z_{1i(t-1)} + [e_{it}]_{(4)}\ ;$$
$$v_{it} = \rho a_{3i}v_{i(t-1)} + [e_{it}]_{(5)}\ .$$

6.1 Estimator Efficiency

Table 1 reports the simulated mean squared errors of the Within-IV, Within-IV-OLS, Within-IV-GLS, feasible Within-IV-OLS, feasible IV-GLS and feasible Within-IV-GLS estimators for the case of two instruments for each regressor when the mean squared errors of the IV-GLS estimator are normalized to be 1. The data for the results in Table 1 were generated by equations (24), (25), (26), (27) and (28). The first number in each parenthesis denotes the relative mean squared error corresponding to the coefficient for the stationary regressor and the second to the coefficient for the nonstationary regressor.

Table 1: Relative Mean Squared Errors: The Case of Two Instruments

(i) $\rho = .5$

(N,T)	WIV	WIVOLS	WIVGLS	FWIVOLS	FIVGLS	FWIVGLS
(30,50)	(1.15,.45)	(1.15,.45)	(1.00,1.03)	(1.15,.45)	(1.05,1.06)	(1.05,1.09)
(30,100)	(1.17,.35)	(1.16,.35)	(1.00,1.02)	(1.17,.35)	(1.02,1.05)	(1.02,1.07)
(30,200)	(1.21,.24)	(1.21,.24)	(1.00,1.00)	(1.21,.24)	(1.02,1.05)	(1.02,1.05)
(80,50)	(1.19,.59)	(1.19,.59)	(1.00,1.04)	(1.19,.59)	(1.04,1.08)	(1.04,1.12)
(80,100)	(1.20,.46)	(1.19,.46)	(1.00,1.02)	(1.19,.46)	(1.02,1.05)	(1.02,1.07)
(80,200)	(1.22,.34)	(1.22,.34)	(1.00,1.00)	(1.22,.34)	(1.02,1.02)	(1.02,1.02)
(150,50)	(1.18,.75)	(1.18,.75)	(1.00,1.05)	(1.18,.75)	(1.05,1.08)	(1.04,1.13)
(150,100)	(1.18,.56)	(1.18,.56)	(1.00,1.01)	(1.18,.56)	(1.02,1.07)	(1.02,1.07)
(150,200)	(1.25,.37)	(1.24,.37)	(1.00,1.00)	(1.24,.38)	(1.02,1.05)	(1.02,1.05)

(ii) $\rho = .7$

(N,T)	WIV	WIVOLS	WIVGLS	FWIVOLS	FIVGLS	FWIVGLS
(30,50)	(1.39,.65)	(1.39,.65)	(1.00,1.04)	(1.39,.65)	(1.05,1.07)	(1.05,1.11)
(30,100)	(1.42,.55)	(1.42,.55)	(1.00,1.03)	(1.42,.55)	(1.02,1.05)	(1.02,1.08)
(30,200)	(1.48,.42)	(1.48,.41)	(1.00,1.01)	(1.48,.41)	(1.02,1.05)	(1.02,1.05)
(80,50)	(1.46,.82)	(1.46,.82)	(1.00,1.06)	(1.46,.82)	(1.04,1.08)	(1.04,1.15)
(80,100)	(1.49,.68)	(1.48,.68)	(1.00,1.03)	(1.48,.68)	(1.02,1.05)	(1.02,1.08)
(80,200)	(1.53,.55)	(1.52,.55)	(1.00,1.01)	(1.52,.55)	(1.02,1.02)	(1.02,1.02)
(150,50)	(1.45,1.03)	(1.45,1.03)	(1.00,1.08)	(1.45,1.03)	(1.05,1.07)	(1.05,1.14)
(150,100)	(1.46,.84)	(1.46,.84)	(1.00,1.02)	(1.46,.84)	(1.02,1.07)	(1.02,1.09)
(150,200)	(1.57,.65)	(1.57,.65)	(1.00,1.01)	(1.57,.65)	(1.02,1.05)	(1.02,1.06)

Notes: 1. The first number in each parenthesis denotes the mean squared error corresponding to the coefficient for the stationary regressor and the second to the coefficient for the nonstationary regressor, when the mean squared errors of the GLS estimator are normalized to be 1.
2. The number of iterations was set at 3,000 when $N \leq 80$, and 1,000 when $N = 150$.
3. The autoregressive coefficients for regression errors were randomly chosen from the uniform distribution U[0,ρ].
4. These notes also apply to Table 2.

We may draw several conclusions from Table 1. First, the IV-GLS and Within-IV-GLS estimators are most efficient for the coefficient corresponding to the stationary regressor, and the feasible GLS estimators have slightly higher mean squared errors than the IV-GLS and Within-IV-GLS estimators. The relative advantage of the GLS estimators over the Within-IV and Within-IV-OLS estimators becomes more apparent as time series span increases and as errors have higher serial correlations and heterogeneity (i.e., when $\rho = 0.7$).

Second, for the coefficient corresponding to the nonstationary regressor, the Within-IV, Within-IV-OLS and feasible IV-OLS estimators are more efficient than the GLS estimators except the case $\rho = .7$ and $(N,T) = (150, 50)$. The efficiency gain becomes more conspicuous as T gets larger. However, part (ii) of Table 1 shows that the relative efficiency gain diminishes as errors have higher serial correlations and heterogeneity (i.e., when $\rho = 0.7$).

Third, for both coefficients, the Within-IV, Within-IV-OLS and feasible Within-IV-OLS show similar performance, so that the theoretical efficiency advantage of Within-IV-OLS over Within-IV does not manifest itself at least in this experiment.

Fourth, the feasible IV-GLS and Within-IV-GLS estimators show similar performance, but the feasible IV-GLS tends to be slightly more efficient than the feasible Within-IV-GLS.

Table 2 reports the efficiency comparison results for the case of one instrument for each regressor. The data for the results in Table 2 were generated by equations (24) and (29). Note that the Within-IV-OLS estimator is equivalent to the Within-IV estimator in this case.

Table 2: Relative Mean Squared Errors: The Case of One Instrument

(i) $\rho = .5$

(N,T)	WIV	WIVGLS	FIVGLS	FWIVGLS
(30,50)	(1.15,1.28)	(1.00,1.08)	(1.04,1.11)	(1.04,1.19)
(30,100)	(1.19,1.21)	(1.00,1.05)	(1.02,1.04)	(1.02,1.10)
(30,200)	(1.17,1.19)	(1.00,1.02)	(1.00,1.04)	(1.00,1.07)
(80,50)	(1.17,1.20)	(1.00,1.09)	(1.04,1.08)	(1.05,1.18)
(80,100)	(1.18,1.19)	(1.00,1.05)	(1.02,1.06)	(1.02,1.11)
(80,200)	(1.20,1.16)	(1.00,1.03)	(1.01,1.04)	(1.01,1.07)
(150,50)	(1.21,1.21)	(1.00,1.07)	(1.03,1.11)	(1.03,1.19)
(150,100)	(1.24,1.17)	(1.00,1.04)	(1.02,1.07)	(1.02,1.12)
(150,200)	(1.24,1.21)	(1.00,1.03)	(1.00,1.04)	(1.00,1.07)

(ii) $\rho = .7$

(N,T)	WIV	WIVGLS	FIVGLS	FWIVGLS
(30,50)	(1.37,1.62)	(1.00,1.10)	(1.03,1.13)	(1.04,1.23)
(30,100)	(1.44,1.51)	(1.00,1.06)	(1.02,1.05)	(1.02,1.12)
(30,200)	(1.42,1.51)	(1.00,1.03)	(1.00,1.05)	(1.00,1.08)
(80,50)	(1.40,1.47)	(1.00,1.11)	(1.04,1.09)	(1.05,1.21)
(80,100)	(1.44,1.49)	(1.00,1.06)	(1.02,1.07)	(1.02,1.13)
(80,200)	(1.46,1.45)	(1.00,1.03)	(1.01,1.04)	(1.01,1.08)
(150,50)	(1.48,1.50)	(1.00,1.09)	(1.03,1.12)	(1.03,1.22)
(150,100)	(1.52,1.47)	(1.00,1.05)	(1.02,1.08)	(1.02,1.14)
(150,200)	(1.54,1.56)	(1.00,1.03)	(1.00,1.05)	(1.00,1.08)

The results in Table 2 can be summarized as follows. First, the GLS estimators are shown to be more efficient that the Within-IV estimator for both coefficients. The efficiency gain of GLS over Within-IV for the coefficient corresponding to the nonstationary regressor contrasts the results in Table 1. The efficiency gain of GLS over Within-IV increases as there are more serial correlations in the errors. Second, the IV-GLS estimator is most efficient with respect to both regressors, and feasible IV-GLS estimator tend to be more efficient than the feasible Within-IV-GLS estimator.

6.2 Wald Tests

Tables 3 and 4 report the empirical size and power of the Wald tests using the Within-IV, feasible Within-IV-OLS, feasible IV-GLS and feasible Within-IV-GLS for the cases of two instruments for each regressor and one instrument for each regressor, respectively. The data were generated as for Tables 1 and 2, except that we used $\alpha = [1.01, 1.01]'$ under the alternative. The nominal size was set at 0.05, and the null hypothesis chosen is $H_0 : \alpha_1 = 1, \alpha_2 = 1$.

Table 3 Empirical Size and Power of Wald Tests: The Case of Two Instruments

(i) $\rho = .5$

	Size				Power			
(N,T)	W_{WIV}	W_{FWIVO}	W_{FIVG}	W_{FWIVG}	W_{WIV}	W_{FWIVO}	W_{FIVG}	W_{FWIVG}
(30,50)	.08	.08	.09	.09	.18	.18	.15	.15
(30,100)	.08	.08	.07	.07	.41	.41	.24	.24
(30,200)	.08	.08	.07	.06	.89	.89	.51	.50
(80,50)	.06	.06	.07	.07	.29	.29	.25	.24
(80,100)	.07	.07	.07	.07	.75	.75	.54	.54
(80,200)	.07	.07	.06	.06	1.0	1.0	.88	.88
(150,50)	.05	.05	.08	.08	.44	.45	.42	.42
(150,100)	.05	.05	.06	.06	.95	.95	.84	.83
(150,200)	.05	.05	.06	.07	1.0	1.0	.98	.98

(ii) $\rho = .7$

	Size				Power			
(N,T)	W_{WIV}	W_{FWIVO}	W_{FIVG}	W_{FWIVG}	W_{WIV}	W_{FWIVO}	W_{FIVG}	W_{FWIVG}
(30,50)	.08	.07	.09	.09	.15	.15	.15	.14
(30,100)	.08	.08	.07	.07	.32	.31	.24	.24
(30,200)	.08	.08	.07	.07	.75	.75	.53	.53
(80,50)	.06	.06	.08	.08	.23	.23	.25	.24
(80,100)	.07	.07	.07	.07	.58	.58	.52	.51
(80,200)	.07	.07	.06	.06	.99	.99	.90	.89
(150,50)	.05	.05	.08	.09	.33	.33	.41	.39
(150,100)	.05	.05	.06	.06	.83	.83	.83	.82
(150,200)	.05	.05	.06	.06	1.0	1.0	.99	.99

Notes: 1. The nominal size is set at 0.05.
2. Powers were calculated at $\alpha = [1.01, 1.01]'$.
3. The number of iterations was set at 3,000 when $N \leq 80$, and 1,000 when $N = 150$.
4. The autoregressive coefficients for regression errors were randomly chosen from the uniform distribution U[0,ρ].
5. These notes also apply to Table 4.

Table 4 Empirical Size and Power of Wald Tests: The Case of One Instrument

(i) $\rho = .5$

	Size			Power		
(N,T)	W_{WIV}	W_{FIVG}	W_{FWIVG}	W_{WIV}	W_{FIVG}	W_{FWIVG}
(30,50)	.08	.11	.12	.17	.22	.22
(30,100)	.08	.10	.11	.37	.43	.43
(30,200)	.08	.10	.10	.82	.87	.86
(80,50)	.06	.09	.09	.24	.33	.33
(80,100)	.07	.09	.09	.66	.75	.74
(80,200)	.07	.07	.07	1.0	1.0	1.0
(150,50)	.05	.08	.08	.41	.52	.50
(150,100)	.06	.06	.07	.89	.93	.93
(150,200)	.06	.05	.05	1.0	1.0	1.0

(ii) $\rho = .7$

	Size			Power		
(N,T)	W_{WIV}	W_{FIVG}	W_{FWIVG}	W_{WIV}	W_{FIVG}	W_{FWIVG}
(30,50)	.08	.12	.12	.14	.22	.21
(30,100)	.08	.10	.11	.29	.40	.39
(30,200)	.08	.10	.10	.67	.82	.82
(80,50)	.06	.09	.09	.19	.30	.31
(80,100)	.07	.09	.09	.50	.69	.67
(80,200)	.07	.07	.07	.95	.99	.99
(150,50)	.06	.08	.08	.30	.47	.44
(150,100)	.06	.07	.07	.74	.89	.89
(150,200)	.06	.06	.06	1.0	1.0	1.0

We may summarize the results in Tables 3 and 4 as follows. First, all of the Wald tests tend to reject slightly more often than is required especially when N is as small as 30. But as N gets larger, the empirical size improves. Second, when there are two instruments for each regressor, W_{WIV} and W_{FWIVO} tend to reject more often than W_{FIVG} and W_{FWIVG}, though such higher rates of rejection dissipate as errors have higher serial correlations and heterogeneity . But when there is only one instrument for each regressor, W_{FIVG} and W_{FWIVG} tend to reject more often than W_{WIV}, which is in accordance with the results in Tables 1 and 2[10]. Third, as errors have higher serial correlations and heterogeneity, the power of the Wald tests based on IV estimator tends to decrease appreciably while those based on GLS estimators do not.

6.3 DWH Tests

Tables 5 and 6 report the empirical size and power of the DWH_G, $GDWH_{WIV}$ and $GDWH_{WIVO}$ for the cases of two instruments for each regressor and one instrument

[10] When an estimator has lower asymptotic variance than another estimator, the Wald test based on the more efficient estimator has higher asymptotic local power than the other. The results in Tables 3 and 4 seem to reflect this.

for each regressor, respectively. The data were generated as for Tables 1 and 2, except that we used $\tau = .0$ under the null and $\tau = .05$ under the alternative.

Table 5 Empirical Size and Power of the DWH Tests: The Case of Two Instruments

(i) $\rho = .5$

	Size ($\tau = .0$)			Power ($\tau = .05$)		
(N, T)	DWH_G	$GDWH_{WIV}$	$GDWH_{WIVO}$	DWH_G	$GDWH_{WIV}$	$GDWH_{WIVO}$
(30,50)	.12	.06	.06	.50	.43	.43
(30,100)	.11	.07	.07	.76	.71	.71
(30,200)	.11	.07	.07	.94	.95	.95
(80,50)	.09	.05	.05	.88	.83	.83
(80,100)	.08	.06	.06	.99	.98	.98
(80,200)	.09	.06	.06	.99	1.0	1.0
(150,50)	.08	.05	.05	.99	.99	.99
(150,100)	.07	.06	.06	1.0	1.0	1.0
(150,200)	.06	.05	.05	1.0	1.0	1.0

(ii) $\rho = .7$

	Size ($\tau = .0$)			Power ($\tau = .05$)		
(N, T)	DWH_G	$GDWH_{WIV}$	$GDWH_{WIVO}$	DWH_G	$GDWH_{WIV}$	$GDWH_{WIVO}$
(30,50)	.13	.06	.06	.50	.39	.39
(30,100)	.11	.07	.07	.75	.67	.67
(30,200)	.12	.07	.07	.93	.92	.92
(80,50)	.09	.05	.05	.87	.78	.78
(80,100)	.09	.05	.05	.99	.97	.97
(80,200)	.09	.06	.06	1.00	1.00	1.00
(150,50)	.08	.04	.04	.99	.96	.96
(150,100)	.07	.06	.06	1.00	1.00	1.00
(150,200)	.06	.05	.05	1.00	1.00	1.00

Notes: 1. The nominal size is set at 0.05.
2. Powers were calculated at $\tau = .05$.
3. The number of iterations was set at 3,000 when $N \leq 80$, and 1,000 when $N = 150$.
4. The autoregressive coefficients for regression errors were randomly chosen from the uniform distribution U$[0,\rho]$.
5. These notes also apply to Table 6

Table 6 Empirical Size and Power of the DWH Tests: The Case of One Instrument

(i) $\rho = .5$

	Size ($\tau = .0$)		Power ($\tau = .05$)	
(N,T)	DWH_G	$GDWH_{WIV}$	DWH_G	$GDWH_{WIV}$
(30,50)	.15	.06	.45	.32
(30,100)	.14	.06	.38	.55
(30,200)	.14	.07	.88	.87
(80,50)	.11	.06	.77	.67
(80,100)	.10	.06	.96	.94
(80,200)	.10	.06	1.0	1.0
(150,50)	.08	.05	.97	.92
(150,100)	.07	.05	1.0	1.0
(150,200)	.07	.05	1.0	1.0

(ii) $\rho = .7$

	Size ($\tau = .0$)		Power ($\tau = .05$)	
(N,T)	DWH_G	$GDWH_{WIV}$	DWH_G	$GDWH_{WIV}$
(30,50)	.16	.06	.44	.30
(30,100)	.14	.06	.65	.55
(30,200)	.13	.07	.88	.83
(80,50)	.12	.05	.77	.62
(80,100)	.10	.06	.96	.90
(80,200)	.10	.06	1.0	1.0
(150,50)	.09	.05	.94	.89
(150,100)	.07	.05	1.0	1.0
(150,200)	.07	.05	1.0	1.0

We may summarize the results in Tables 5 and 6 as follows. First, the $GDWH_{WIV}$ and $GDWH_{WIVO}$ tests tend to show better size properties than the DWH_G in all the cases. The empirical size of the $GDWH_{WIV}$ and $GDWH_{WIVO}$ tests is reasonably close to .05 in most cases. Second, the DWH_G tends to reject more often than the $GDWH_{WIV}$ and $GDWH_{WIVO}$ tests, but this reflects its size distortions.[11] Third, as errors have higher serial correlation and heterogeneity, the $GDWH_{WIV}$ and $GDWH_{WIVO}$ tests tend to reject less often under the alternative, while the DWH_G test does not. This is accordance with the power properties of the Wald tests.

7 Summary and Further Remarks

We have studied asymptotic properties of the Within-IV, Within-IV-OLS, Within-IV-GLS and IV-GLS estimators for an error component model with stationary and nearly nonstationary regressors under the assumption of $N \to \infty$ and $T \to \infty$. Autoregressive disturbances are assumed for the error component model, the structure of which may vary with individuals. We have shown that all of the estimators have

[11] We also calculated size-adjusted power of the DWH tests, but found that the $GDWH_{WIV}$ and $GDWH_{WIVO}$ tests have higher power than the DWH_G test.

normal distributions in the limit under proper conditions. Additionally, Wald tests for coefficient vectors and DWH tests for exogeneity are studied. Simulation results regarding the estimator efficiency show that the efficiency ranking of the estimators crucially depends on the type of regressor and the number of instruments. Simulation results for the Wald and DWH tests show that these tests keep nominal size reasonably well when N is large. The power properties of the Wald tests depend on the number of instruments, as in the case of efficiency comparison, and the degree of serial correlation and heterogeneity in the errors. The power properties of the DWH tests are also shown to depend on the degree of serial correlation and heterogeneity in the errors.

8 Appendix: Proofs

Proof of Lemma 1: The vector diffusion process $K_{a_{2i}}(r)$ $(a = x, z)$ can be written as

$$K_{a_{2i}}(r) = B_{a_{2i}}(r) + C_{a_{2i}} \int_0^r e^{(r-s)C_{a_{2i}}} B_{a_{2i}}(s)ds \tag{A.1}$$

(cf. Phillips, 1988). Thus,

$$\begin{aligned} E\left(K_{x_{2i}}(r)K'_{z_{2i}}(t)\right) &= E\left(B_{x_{2i}}(r)B'_{z_{2i}}(t)\right) + E\left(B_{x_{2i}}(r)\int_0^t B'_{z_{2i}}(s)e^{(t-s)C'_{z_{2i}}}dsC'_{z_{2i}}\right) \\ &\quad + E\left(C_{x_{2i}}\int_0^r e^{(r-s)C_{x_{2i}}}B_{x_{2i}}(s)dsB'_{z_{2i}}(t)\right) \qquad \text{(A.2)} \\ &\quad + E\left(C_{x_{2i}}\int_0^r e^{(r-s)C_{x_{2i}}}B_{x_{2i}}(s)ds\int_0^t B'_{z_{2i}}(u)e^{(t-u)C'_{z_{2i}}}duC'_{z_{2i}}\right). \end{aligned}$$

Fubini's theorem allows us to interchange the expectation operator and the integral in (A.2), and hence the result follows.

Proof of Lemma 2: (a) Writing

$$\begin{aligned} E\left(\bar{K}_{x_{2i}}(r)\bar{K}'_{z_{2i}}(r)dr\right) &= E\left(K_{x_{2i}}(r)K'_{z_{2i}}(r)\right) - \int_0^r E\left(K_{x_{2i}}(s)K'_{z_{2i}}(r)\right)ds \\ &\quad - \int_0^r E\left(K_{x_{2i}}(r)K'_{z_{2i}}(s)\right)ds + \int_0^r\int_0^r E\left(K_{x_{2i}}(s)K'_{z_{2i}}(t)\right)dsdt \end{aligned}$$

and using Lemma 1 gives the stated result.
(b) This is proven in the same manner as in part (a).

Lemma 4 *Under Assumption 5,*
(a)

$$\| \frac{1}{T}\Sigma_{t=1}^T(x_{1it} - \bar{x}_{1i.})(z_{1it} - \bar{z}_{1i.})' \|, \ \| \frac{1}{T^{3/2}}\Sigma_{t=1}^T(x_{1it} - \bar{x}_{1i.})(z_{2it} - \bar{z}_{2i.})' \|,$$

$$\| \frac{1}{T^{3/2}}\Sigma_{t=1}^T(x_{2it} - \bar{x}_{2i.})(z_{1it} - \bar{z}_{1i.})' \|, \ \| \frac{1}{T^2}\Sigma_{t=1}^T(x_{2it} - \bar{x}_{2i.})(z_{2it} - \bar{z}_{2i.})' \|,$$

$$\| \frac{1}{T}\Sigma_{t=1}^{T}(z_{1it}-\bar{z}_{1i.})(z_{1it}-\bar{z}_{1i.})' \|, \ \| \frac{1}{T^{3/2}}\Sigma_{t=1}^{T}(z_{1it}-\bar{z}_{1i.})(z_{2it}-\bar{z}_{2i.})' \|$$

and

$$\| \frac{1}{T^2}\Sigma_{t=1}^{T}(z_{2it}-\bar{z}_{2i.})(z_{2it}-\bar{z}_{2i.})' \|$$

are uniformly integrable in T for all i;

(b)

$$\| \frac{1}{\sqrt{T}}\Sigma_{t=1}^{T}(z_{1it}-\bar{z}_{1i.})v_{it} \|^2 \quad \text{and} \| \frac{1}{T}\Sigma_{t=1}^{T}(z_{2it}-\bar{z}_{2i.})v_{it} \|^2$$

are uniformly integrable in T for all i.

Proof) (a) $\| \frac{1}{T}\Sigma_{t=1}^{T}(x_{1it}-\bar{x}_{1i.})(z_{1it}-\bar{z}_{1i.})' \|$ is uniformly integrable in T for all i if

$$\sup_T E \| \frac{1}{T}\Sigma_{t=1}^{T}(x_{1it}-\bar{x}_{1i.})(z_{1it}-\bar{z}_{1i.})' \|^{1+\varepsilon} < \infty \ (\varepsilon > 0) \tag{A.3}$$

for all i (see, for example, Billingsley,1968, p. 32). But relation (A.3) holds if for all i

$$\begin{aligned} \sup_T E \quad & | \quad \left[\frac{1}{T}\Sigma_{t=1}^{T}(x_{1it}-\bar{x}_{1i.})(z_{1it}-\bar{z}_{1i.})'\right]_{(h,j)} |^{1+\varepsilon} \\ &= \quad \sup_T E \ | \left[\frac{1}{T}\Sigma_{t=1}^{T}x_{1it}z_{1it}' - \bar{x}_{1i.}\bar{z}_{1i.}'\right]_{(h,j)} |^{1+\varepsilon} < \infty \text{ for all } h \text{ and } j \end{aligned} \tag{A.4}$$

due to part (b) of Lemma 9 in Phillips and Moon (1998)[12]. Furthermore, relation (A.4) holds due to the Minkowski inequality if for all i

$$\sup_T E \ | \left[\frac{1}{T}\Sigma_{t=1}^{T}x_{1it}z_{1it}'\right]_{(h,j)} |^{1+\varepsilon} < \infty \text{ for all } h \text{ and } j \tag{A.5}$$

and

$$\sup_T E \ | \left[\bar{x}_{1i.}\bar{z}_{1i.}'\right]_{(h,j)} |^{1+\varepsilon} < \infty \text{ for all } h \text{ and } j. \tag{A.6}$$

Again due to the Minkowski inequality, relation (A.5) hold if

$$\sup_T \sup_{1\le t\le T} E \ | \left[x_{1it}z_{1it}'\right]_{(h,j)} |^{1+\varepsilon} < \infty \text{ for all } h \text{ and } j \tag{A.7}$$

and relation (A.6) if

$$\sup_T \sup_{1\le t,s\le T} E \ | \left[x_{1it}z_{1is}'\right]_{(h,j)} |^{1+\varepsilon} < \infty \text{ for all } h \text{ and } j, \tag{A.8}$$

[12] Part (b) of Lemma 9 in Phillips and Moon (1998) is: *For any $p \ge 1$ and any $m \times n$ matrix A, there exists a constant $M > 0$ such that $\| A \|^p \le M\Sigma_{i=1}^{m}\Sigma_{j=1}^{n} | a_{i,j} |^p$, where $a_{i,j}$ is the (i,j)-th element of A.*

which implies relation (A.7). Thus, inequality (A.8) is a necessary condition for $\| \frac{1}{T}\Sigma_{t=1}^{T}(x_{1it}-\bar{x}_{1i.})(z_{1it}-\bar{z}_{1i.})' \|$ to be uniformly integrable. Using the same methods, we find that $\| \frac{1}{T^{3/2}}\Sigma_{t=1}^{T}(x_{1it}-\bar{x}_{1i.})(z_{2it}-\bar{z}_{2i.})' \|$ is uniformly integrable if

$$\sup_T \sup_{1\leq t,s\leq T} E \mid \left[\frac{1}{\sqrt{T}} x_{1it} z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty.$$

But

$$\begin{aligned}\sup_T \sup_{1\leq t,s\leq T} E \quad \mid \quad \left[\frac{1}{\sqrt{T}} x_{1it} z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon} &= \sup_T \sup_{1\leq t,s\leq T} \frac{s^{1+\varepsilon}}{T^{1+\varepsilon}} E \mid \left[\frac{1}{\sqrt{s}} x_{1it} z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon} \\ &\leq \sup_T \sup_{1\leq t,s\leq T} E \mid \left[\frac{1}{\sqrt{s}} x_{1it} z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty\end{aligned}$$

due to the given assumption, which proves that $\| \frac{1}{T^{3/2}}\Sigma_{t=1}^{T}(x_{1it}-\bar{x}_{1i.})(z_{2it}-\bar{z}_{2i.})' \|$ is uniformly integrable. The other parts can be shown by the same methods.
(b) We may apply Theorem 5.4 in Billingsley (1968) due to conditions (h)-(2) and (g)-(2) in Assumption 5. Thus, the stated results follow.

Proof of Theorem 1: (a) Apply the law of large numbers, the central limit theorem and Lemma 2 to the weak limits in Assumption 4 which are independent due to Assumption 3. Then, the result follows. It is trivial to show by using equation (A.1) and Lemmas A10 and A12 in Choi (1998) that the moment conditions for the law of large numbers and the central limit theorem for independent but not identically distributed random vectors hold.
(b) Lemma 4 shows that the uniform integrability conditions for Corollary 1 and Theorem 3 in Phillips and Moon (1998) are satisfied under Assumption 5. Moreover, it is straightforward to show that the other conditions for Corollary 1 and Theorem 3 in Phillips and Moon are also satisfied under the given conditions. Thus, the sequential and joint limits are identical, which proves part (b).

Lemma 5 *Under Assumption 5 (b), (e) and (f),* $\| \frac{1}{T}\bar{Z}'_{1i}\bar{V}_i\bar{Z}_{1i} \|$, $\| \frac{1}{T^{3/2}}\bar{Z}'_{1i}\bar{V}_i\bar{Z}_{2i} \|$ *and* $\| \frac{1}{T^2}\bar{Z}'_{2i}\bar{V}_i\bar{Z}_{2i} \|$ *are uniformly integrable in* T *for all* i.

Proof) As in the proof of part (a) of Lemma 4, $\| \frac{1}{T}\bar{Z}'_{1i}\bar{V}_i\bar{Z}_{1i} \|$ is uniformly integrable in T for all i if

$$\sup_T E \mid \left[\frac{1}{T}\bar{Z}'_{1i}\bar{V}_i\bar{Z}_{1i}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty \tag{A.9}$$

for all h and j. The (h,j)-th element of $\bar{Z}'_{1i}\bar{V}_i\bar{Z}_{1i}$ can be written as

$$\begin{aligned}&\Sigma_{t=1}^{T}(z_{1ith}-\bar{z}_{1i.h})(z_{1itj}-\bar{z}_{1i.j}) + \Sigma_{t=1}^{T-1}(z_{1ith}-\bar{z}_{1i.h})(z_{1i(t+1)j}-\bar{z}_{1i.j})\gamma_{i1} + ... \\ &+\Sigma_{t=1}^{1}(z_{1ith}-\bar{z}_{1i.h})(z_{1i(t+1)j}-\bar{z}_{1i.j})\gamma_{i(T-1)} + \Sigma_{t=1}^{T-1}(z_{1itj}-\bar{z}_{1i.j})(z_{1i(t+1)h}-\bar{z}_{1i.h})\gamma_{i1} + ... \\ &+\Sigma_{t=1}^{1}(z_{1itj}-\bar{z}_{1i.j})(z_{1i(t+1)h}-\bar{z}_{1i.h})\gamma_{i(T-1)}.\end{aligned}$$

Because γ_{it} are all finite, relation (A.9) holds under part (b) of Assumption 5 as in the proof of part (a) of Lemma 4. The other parts can be proven in an analogous manner.

Proof of Theorem 2: (a) This is proven in the same manner as in part (a) of Theorem 1.
(b) Due to Lemmas 4 and 5, the uniform integrability conditions for Corollary 1 and Theorem 3 in Phillips and Moon (1998) are satisfied; and the other conditions for Corollary 1 and Theorem 3 in Phillips and Moon can straightforwardly be shown to be satisfied. Therefore, the sequential and joint limit are identical as desired.

Proof of Lemma 3: See the proof of Lemma 3 in Choi (1998).

Lemma 6 *Under Assumption 9, we have the followings as* $T \to \infty$.

(a) $\quad i_T' \Phi_i^{-1} i_T \to \frac{1}{\sigma_\mu^2}$.

(b) $\quad \left[\begin{array}{cc} \frac{1}{\sqrt{T}} i_T' \Phi_i^{-1} S_{1i} & \frac{1}{T} i_T' \Phi_i^{-1} S_{2i} \end{array} \right] \Rightarrow \left[\begin{array}{cc} 0, & 0 \end{array} \right]$ where $S = Z$ or X.

(c)

$$P_T^{-1} Z_i' \Phi_i^{-1} X_i Q_T^{-1} \Rightarrow \left[\begin{array}{cc} \frac{1}{\sigma_w^2} \delta_i(z,x) & 0 \\ 0 & \frac{\mathfrak{s}_{p_i}^2}{\sigma_w^2} \int_0^1 \bar{K}_{z2i}(r) \bar{K}_{x2i}'(r) dr \end{array} \right],$$

where $P_T \triangleq \left[\begin{array}{cc} \sqrt{T} I_{l_1} & 0 \\ 0 & T I_{l_2} \end{array} \right]$, $Q_T \triangleq \left[\begin{array}{cc} \sqrt{T} I_{k_1} & 0 \\ 0 & T I_{k_2} \end{array} \right]$ and $\delta_i(z,x)$ *is defined in Theorem 3.*

(d)

$$P_T^{-1} Z_i' \Phi_i^{-1} Z_i P_T^{-1} \Rightarrow \left[\begin{array}{cc} \frac{1}{\sigma_w^2} \delta_i(z,z) & 0 \\ 0 & \frac{\mathfrak{s}_{p_i}^2}{\sigma_w^2} \int_0^1 \bar{K}_{z2i}(r) \bar{K}_{z2i}'(r) dr \end{array} \right],$$

where $\delta_i(z,z)$ *is defined in Assumption 9 (f).*

(e) $\quad i_T' \Phi_i^{-1} u_i \xrightarrow{p} \frac{\mu_i}{\sigma_\mu^2}$.

(f) $\quad \frac{1}{\sqrt{T}} Z_{1i}' \Phi_i^{-1} u_i \Rightarrow \frac{1}{\sigma_w^2} N\left(0, \sigma_w^2 \delta_i(z_1, z_1)\ \right)$.

(g) $\quad \frac{1}{T} Z_{2i}' \Phi_i^{-1} u_i \Rightarrow \frac{\mathfrak{s}_{p_i}}{\sigma_w^2} \left(\int_0^1 \bar{K}_{z2i}(r) dB_{w_i}(r) \right)$.

Proof) (a) This is proven in Lemma A8 of Choi (1998).
(b) As in the proof of Lemma A8 in Choi (1998), we have

$$i_T' \Phi_i^{-1} X_i = (\mathfrak{l}_i' \tilde{X}_i - \mathfrak{m}_i' \ddot{X}_i) / \{\sigma_w^2 + \sigma_\mu^2 (\mathfrak{l}_i' \mathfrak{l}_i - \mathfrak{m}_i' \mathfrak{m}_i)\}. \tag{A.10}$$

where, letting $x_{it}' \triangleq \left[\begin{array}{cc} x_{1it}' & x_{2it}' \end{array} \right]$,

$$\begin{aligned} \mathfrak{l}_i' \tilde{X}_i &= \mathfrak{s}_{p_i} \left[\sum_{k=0}^{p_i - 1} \delta_k \left(\sum_{j=0}^{k} \rho_{ij} x_{i(k+1-j)}' \right) + \sum_{k=p_i+1}^{T} \left(x_{ik}' + \rho_{i1} x_{i(k-1)}' + \ldots + \rho_{ip_i} x_{i(k-p_i)}' \right) \right] \\ &= \mathfrak{s}_{p_i} \left[\sum_{k=p_i+1}^{T} \left(x_{ik}' + \rho_{i1} x_{i(k-1)}' + \ldots + \rho_{ip_i} x_{i(k-p_i)}' \right) \right] + O_p(1) \end{aligned} \tag{A.11}$$

and

$$\mathfrak{m}_i'\ddot{X}_i = \rho_{i1}\mathfrak{t}_1 x_{i1}' + \rho_{i2}(\mathfrak{t}_2 x_{i1}' + \mathfrak{t}_1 x_{i2}')... + \rho_{ip_i}(\mathfrak{t}_{p_i} x_{i1}' + ... + \mathfrak{t}_1 x_{ip_i}') = O_p(1). \quad \text{(A.12)}$$

But due to Assumption 9,

$$\mathfrak{l}_i'\tilde{X}_{1i} = O_p(\sqrt{T}) \text{ and } \mathfrak{l}_i'\tilde{X}_{2i} = O_p(T^{3/2}), \quad \text{(A.13)}$$

where $\mathfrak{l}_i'\tilde{X}_i = \begin{bmatrix} \mathfrak{l}_i'\tilde{X}_{1i} & \mathfrak{l}_i'\tilde{X}_{2i} \end{bmatrix}$. Because $\mathfrak{l}_i'\mathfrak{l}_i - \mathfrak{m}_i'\mathfrak{m}_i$ behaves in the limit as if it were $\mathfrak{s}_{p_i}^2 T$, the result follows from (A.11)-(A.13). The proof for $S = Z$ is exactly the same.
(c) Using the definitions of $\tilde{S}_i$ and $\ddot{S}_i$ $(S = Z, X)$, we obtain

$$\begin{aligned}\tilde{Z}_i'\tilde{X}_i &= \sum_{k=1}^{p_i}(\sum_{j=1}^{k}\rho_{i(k-j)}z_{ij})(\sum_{j=1}^{k}\rho_{i(k-j)}x_{ij}') \quad \text{(A.14)}\\ &+ \sum_{t=p_i+1}^{T}(z_{it} + \rho_{i1}z_{i(t-1)} + ... + \rho_{ip_i}z_{i(t-p_i)})(x_{it}' + \rho_{i1}x_{i(t-1)}' + ... + \rho_{ip_i}x_{i(t-p_i)}')\\ &= \sum_{t=p_i+1}^{T}(z_{it} + \rho_{i1}z_{i(t-1)} + ... + \rho_{ip_i}z_{i(t-p_i)})(x_{it}' + \rho_{i1}x_{i(t-1)}' + ... + \rho_{ip_i}x_{i(t-p_i)}') + O_p(1)\end{aligned}$$

and

$$\ddot{Z}_i'\ddot{X}_i = \sum_{k=1}^{p_i}\{\rho_{i(p_i-k+j)}z_{ij}\}\{\rho_{i(p_i-k+j)}x_{ij}'\} = O_p(1). \quad \text{(A.15)}$$

Assumption 9 gives

$$P_T^{-1}\tilde{Z}_i'\tilde{X}_iQ_T^{-1} \Rightarrow \begin{bmatrix} \gamma_{zx_i}(0)\sum_{j=0}^{p_i}\rho_{ij}^2 + 2\sum_{k=1}^{p_i}\gamma_{zx_i}(k)\sum_{j=0}^{p_i-k}\rho_{ij}\rho_{i(j+k)} & 0 \\ 0 & \mathfrak{s}_{p_i}^2\int_0^1 K_{z2i}(r)K_{x2i}'(r)dr \end{bmatrix}. \quad \text{(A.16)}$$

Also, because τ_i behaves in the limit as if it were $\frac{1}{\sigma_w^2\mathfrak{s}_{p_i}^2 T}$, Assumption 9 gives

$$P_T^{-1}\tau_i[\mathfrak{l}_i'\tilde{Z}_i - \mathfrak{m}_i'\ddot{Z}_i]'[\mathfrak{l}_i'\tilde{X}_i - \mathfrak{m}_i'\ddot{X}_i]Q_T^{-1} \Rightarrow \begin{bmatrix} 0 & 0 \\ 0 & \frac{\mathfrak{s}_{p_i}^2}{\sigma_w^2}\int_0^1 K_{z2i}(r)dr\int_0^1 K_{x2i}'(r)dr \end{bmatrix}. \quad \text{(A.17)}$$

Now, plugging (A.15), (A.16) and (A.17) into part (c) of Lemma 3 gives the required result.
(d) This is proven in the same way as for part (c).
(e) This is proven in the proof of Lemma 8 in Choi (1998).
(f), (g) As in the proof of Lemma 8 in Choi (1998), we have

$$\begin{aligned}&P_T^{-1}Z_i'\Phi_i^{-1}u_i \quad \text{(A.18)}\\ &= P_T^{-1}\sum_{t=p_i+1}^{T}\frac{1}{\sigma_w^2}(z_{it} + \rho_{i1}z_{i(t-1)} + ... + \rho_{ip_i}z_{i(t-p_i)})(w_{it} + \mathfrak{s}_{p_i}\mu_i)\end{aligned}$$

$$-\mathfrak{s}_{p_i}^2 \tau_i P_T^{-1} \sum_{s=p_i+1}^{T} (z_{is} + \rho_{i1} z_{i(s-1)} + ... + \rho_{ip_i} z_{i(s-p_i)})$$
$$\times(\sum_{s=p_i+1}^{T} w_{it} + \mathfrak{s}_{p_i}(T-p_i)\mu_i) + o_p(1).$$

But because τ_i behaves in the limit as if it were $\frac{1}{\sigma_w^2 \mathfrak{s}_{p_i}^2 T}$, the terms involving μ_i in (A.18) become negligible in the limit. Therefore, Assumption 9 gives the required result.

Lemma 7 *: (a) Under Assumption 10,* $\parallel \frac{1}{\sqrt{T}} i_T' \Phi_i^{-1} S_{1i} \parallel$, $\parallel \frac{1}{T} i_T' \Phi_i^{-1} S_{2i} \parallel$ $(S = Z$ *or* $X)$, $\parallel P_T^{-1} Z_i' \Phi_i^{-1} X_i Q_T^{-1} \parallel$, $\parallel P_T^{-1} Z_i' \Phi_i^{-1} Z_i P_T^{-1} \parallel$ *are uniformly integrable.*

(b) Under Assumptions 1, 2 and 10, $\parallel i_T' \Phi_i^{-1} u_i \parallel^2$, $\parallel \frac{1}{\sqrt{T}} Z_{1i}' \Phi_i^{-1} u_i \parallel^2$ *and* $\parallel \frac{1}{T} Z_{2i}' \Phi_i^{-1} u_i \parallel^2$ *are uniformly integrable.*

Proof) (a) As in part (a) of proof for Lemma 4, $\parallel \frac{1}{\sqrt{T}} i_T' \Phi_i^{-1} S_{1i} \parallel$ is uniformly integrable in T for all i if

$$\sup_T E \mid \left[\frac{1}{\sqrt{T}} i_T' \Phi_i^{-1} S_{1i} \right]_{(h)} \mid^{1+\varepsilon} < \infty \ (\varepsilon > 0)$$

for all h and i. But relation (A.10) yields

$$\begin{aligned} \sup_T E \quad & \mid \quad \left[\frac{1}{\sqrt{T}} i_T' \Phi_i^{-1} S_{1i} \right]_{(h)} \mid^{1+\varepsilon} \quad \text{(A.19)} \\ & \leq \sup_T \mid \frac{1}{\sqrt{T}(\sigma_\mu^2 (\mathfrak{l}_i' \mathfrak{l}_i - \mathfrak{m}_i' \mathfrak{m}_i))} \mid^{1+\varepsilon} E \mid \left[(\mathfrak{l}_i' \tilde{S}_{1i} - \mathfrak{m}_i' \ddot{S}_{1i}) \right]_{(h)} \mid^{1+\varepsilon} \\ & = \sup_T \mid \frac{1}{\sqrt{T}(\sigma_\mu^2 (\mathfrak{l}_i' \mathfrak{l}_i - \mathfrak{m}_i' \mathfrak{m}_i))} \mid^{1+\varepsilon} E \mid \left[\sum_{t=1}^{T} c_{it} s_{1it}' \right]_{(h)} \mid^{1+\varepsilon}, \end{aligned}$$

where c_{ik} are finite constants. Moreover, due to the Minkowski inequality and Assumption 10,

$$\begin{aligned} E \quad & \mid \quad \left[\sum_{t=1}^{T} c_{it} s_{1it}' \right]_{(h)} \mid^{1+\varepsilon} \leq \left[\sum_{t=1}^{T} \left(E \mid c_{it} s_{1it}' \mid_{(h)}^{1+\epsilon} \right)^{\frac{1}{1+\varepsilon}} \right]^{1+\varepsilon} \\ & \leq \left[T \left(\sup_{1 \leq i \leq N, 1 \leq t \leq T} \mid c_{it} \mid^{1+\epsilon} \sup_{1 \leq t \leq T} E \mid s_{1it}' \mid_{(h)}^{1+\epsilon} \right)^{\frac{1}{1+\varepsilon}} \right]^{1+\varepsilon} \\ & = T^{1+\varepsilon} \sup_{1 \leq i \leq N, 1 \leq t \leq T} \mid c_{it} \mid^{1+\epsilon} \sup_{1 \leq t \leq T} E \mid s_{1it}' \mid_{(h)}^{1+\epsilon} . \quad \text{(A.20)} \end{aligned}$$

Letting $\check{c} = \sup_{1 \le i \le N, 1 \le t \le T} \mid c_{it} \mid^{1+\epsilon}$ and noting that $\mid \frac{T}{\sqrt{T}(\sigma_\mu^2(\mathfrak{l}_i'\mathfrak{l}_i - \mathfrak{m}_i'\mathfrak{m}_i))} \mid^{1+\varepsilon} < \infty$ for all T, relations (A.19) and (A.20) give the inequality

$$\begin{aligned} \sup_T E \quad & \mid \quad \left[\frac{1}{\sqrt{T}} i_T' \Phi_i^{-1} S_{1i} \right]_{(h)} \mid^{1+\varepsilon} \\ & \le \quad \check{c} \sup_T \mid \frac{T}{\sqrt{T}(\sigma_\mu^2(\mathfrak{l}_i'\mathfrak{l}_i - \mathfrak{m}_i'\mathfrak{m}_i))} \mid^{1+\varepsilon} \sup_T \sup_{1 \le t \le T} E \mid s_{1it}' \mid_{(h)}^{1+\epsilon} < \infty \end{aligned}$$

as required. In the same manner, the uniform integrability of $\| \frac{1}{T} i_T' \Phi_i^{-1} S_{2i} \|$ follows from the inequality

$$\begin{aligned} \sup_T E \quad & \mid \quad \left[\frac{1}{T} i_T' \Phi_i^{-1} S_{2i} \right]_{(h)} \mid^{1+\varepsilon} \\ & \le \quad \check{d} \sup_T \sup_{1 \le t \le T} \mid \frac{\sqrt{t}}{\sigma_\mu^2(\mathfrak{l}_i'\mathfrak{l}_i - \mathfrak{m}_i'\mathfrak{m}_i)} \mid^{1+\varepsilon} \sup_{1 \le t \le T} E \mid \frac{1}{\sqrt{t}} s_{2it}' \mid_{(h)}^{1+\epsilon} \\ & \le \quad \check{d} \sup_T \mid \frac{\sqrt{T}}{\sigma_\mu^2(\mathfrak{l}_i'\mathfrak{l}_i - \mathfrak{m}_i'\mathfrak{m}_i)} \mid^{1+\varepsilon} \sup_T \sup_{1 \le t \le T} E \mid \frac{1}{\sqrt{t}} s_{2it}' \mid_{(h)}^{1+\epsilon} < \infty, \end{aligned}$$

where $\check{d}$ is a finite constant. For the rest, we prove the uniform integrability result only for $\| \frac{1}{T^2} Z_{2i}' \Phi_i^{-1} Z_{2i} \|$ because proofs for the other parts are completely analogous. Part (c) of Lemma 3 provides

$$Z_{2i}' \Phi_i^{-1} Z_{2i} = [\tilde{Z}_{2i}' \tilde{Z}_{2i} - \ddot{Z}_{2i}' \ddot{Z}_{2i}] / \sigma_w^2 - \tau_i (\mathfrak{l}_i' \tilde{Z}_{2i} - \mathfrak{m}_i' \ddot{Z}_{2i})' (\mathfrak{l}_i' \tilde{Z}_{2i} - \mathfrak{m}_i' \ddot{Z}_{2i})$$

where $\tau_i = \sigma_\mu^2 / \{\sigma_w^4 + \sigma_\mu^2 \sigma_w^2 (\mathfrak{l}_i'\mathfrak{l}_i - \mathfrak{m}_i'\mathfrak{m}_i)\}$. Thus, as in the proof of Lemma 4, $\| \frac{1}{T^2} Z_{2i}' \Phi_i^{-1} Z_{2i} \|$ is uniformly integrable if

$$\sup_T \; E \mid \frac{1}{T^2} \left[\tilde{Z}_{2i}' \tilde{Z}_{2i} - \ddot{Z}_{2i}' \ddot{Z}_{2i} \right]_{(h,j)} \mid^{1+\varepsilon} < \infty \tag{A.21}$$

and

$$\sup_T \; E \mid \frac{1}{T^2} \left[\tau_i (\mathfrak{l}_i' \tilde{Z}_{2i} - \mathfrak{m}_i' \ddot{Z}_{2i})' (\mathfrak{l}_i' \tilde{Z}_{2i} - \mathfrak{m}_i' \ddot{Z}_{2i}) \right]_{(h,j)} \mid^{1+\varepsilon} < \infty \tag{A.22}$$

for all h, j and i. Relation (A.21) holds if

$$\sup_T \; E \mid \frac{1}{T^2} \left[\tilde{Z}_{2i}' \tilde{Z}_{2i} \right]_{(h,j)} \mid^{1+\varepsilon} < \infty \tag{A.23}$$

and

$$\sup_T \; E \mid \frac{1}{T^2} \left[\ddot{Z}_{2i}' \ddot{Z}_{2i} \right]_{(h,j)} \mid^{1+\varepsilon} < \infty. \tag{A.24}$$

But using (A.14) and (A.15), we find that

$$\sup_T E \mid \frac{1}{T^2}\left[\tilde{Z}'_{2i}\tilde{Z}_{2i}\right]_{(h,j)} \mid^{1+\varepsilon}$$
$$\leq \sup_{0\leq j,k\leq p_i} \mid \rho_{ij}\rho_{ik} \mid \sup_T \mid \sup_{1\leq t,s\leq T} \frac{\sqrt{ts}d_i}{T^2}(T-f_i) \mid^{1+\varepsilon} E \mid \left[\frac{1}{\sqrt{ts}} z_{2it}z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon}$$
$$< \sup_{0\leq j,k\leq p_i} \mid \rho_{ij}\rho_{ik} \mid d_i{}^{1+\varepsilon} \sup_T \sup_{1\leq t,s\leq T} E \mid \left[\frac{1}{\sqrt{ts}} z_{2it}z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty,$$

where d_i and f_i are positive constants, which shows that relation (A.23) holds. In the same manner, relation (A.24) can be shown to hold. These establish relation (A.21). Moreover, relations (A.11) and (A.12) yield $\mathfrak{l}'_i\tilde{Z}_{2i} - \mathfrak{m}'_i\ddot{Z}_{2i} = \sum_{t=1}^T g_{it}z'_{2ti}$ where g_{it} are finite constants. Therefore,

$$\sup_T E \mid \frac{1}{T^2}\left[\tau_i(\mathfrak{l}'_i\tilde{Z}_{2i} - \mathfrak{m}'_i\ddot{Z}_{2i})'(\mathfrak{l}'_i\tilde{Z}_{2i} - \mathfrak{m}'_i\ddot{Z}_{2i})\right]_{(h,j)} \mid^{1+\varepsilon}$$
$$= \sup_T \frac{\tau_i}{T^2} E \mid \left[(\sum_{t=1}^T g_{it}z_{2it})(\sum_{t=1}^T g_{it}z'_{2it})\right]_{(h,j)} \mid^{1+\varepsilon}$$
$$\leq \sup_T \sup_{1\leq t,s\leq T} \mid g_{it}g_{is} \mid \tau_i\sqrt{ts} E \mid \left[\frac{1}{\sqrt{ts}} z_{2it}z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon}$$
$$\leq \sup_T \sup_{1\leq t,s\leq T} \mid g_{it}g_{is} \mid \tau_i\sqrt{ts} \sup_T \sup_{1\leq t,s\leq T} E \mid \left[\frac{1}{\sqrt{ts}} z_{2it}z'_{2is}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty,$$

which proves relation (A.22). Thus, $\parallel \frac{1}{T^2}Z'_{2i}\Phi_i^{-1}X_{2i} \parallel$ is uniformly integrable.

(b) The results are obtained under Assumption 10 by applying Theorem 5.4 in Billingsley (1968). Note that conditions for $i'_T\Phi_i^{-1}u_i$ similar to (e)-(1), (e)-(2), (f)-(1) and (f)-(2) in Assumption 10 are not required, because these can straightforwardly be shown to hold under Assumptions 1 and 2.

Proof of Theorem 3: (a) When $T \to \infty$, we have due to Lemma 6

$$\sqrt{N}(\hat{\beta}_{IVG} - \beta_{IVG}) \Rightarrow \frac{1}{\sqrt{N}}\sum_{i=1}^N \mu_i, \tag{A.25}$$

$$\sqrt{NT}(\hat{\alpha}_{1IVG} - \alpha_1) \Rightarrow \left[\frac{1}{N}\sum_{i=1}^N \delta_i(x_1,z_1)\left(\frac{1}{N}\sum_{i=1}^N \delta_i(z_1,z_1)\right)^{-1}\frac{1}{N}\sum_{i=1}^N \delta_i(z_1,x_1)\right]^{-1}$$
$$\times\frac{1}{N}\sum_{i=1}^N \delta_i(x_1,z_1)\left(\frac{1}{N}\sum_{i=1}^N \delta_i(z_1,z_1)\right)^{-1}\frac{1}{\sqrt{N}}\sum_{i=1}^N N\left(0,\ \sigma_w^2\delta_i(z_1,z_1)\right). \tag{A.26}$$

Now, apply the law of large numbers and the central limit theorem to (A.25) and (A.26) by sending N to infinity. Then, parts (i) and (ii) follow. Moreover, due to

Lemma 6, as $T \to \infty$

$$
\begin{aligned}
&\sqrt{N}T(\hat{\alpha}_{2IVG}-\alpha_2) \Rightarrow \\
&[\frac{1}{N}\sum_{i=1}^{N}\mathfrak{s}_{p_i}^2\int_0^1 \bar{K}_{z_{2i}}(r)\bar{K}'_{x_{2i}}(r)dr\left(\frac{1}{N}\sum_{i=1}^{N}\mathfrak{s}_{p_i}^2\int_0^1 \bar{K}_{z_{2i}}(r)\bar{K}'_{z_{2i}}(r)dr\right)^{-1} \\
&\times\frac{1}{N}\sum_{i=1}^{N}\mathfrak{s}_{p_i}^2\int_0^1 \bar{K}_{x_{2i}}(r)\bar{K}'_{z_{2i}}(r)dr]^{-1} \\
&\times\frac{1}{N}\sum_{i=1}^{N}\mathfrak{s}_{p_i}^2\int_0^1 \bar{K}_{z_{2i}}(r)\bar{K}'_{x_{2i}}(r)dr\left(\frac{1}{N}\sum_{i=1}^{N}\mathfrak{s}_{p_i}^2\int_0^1 \bar{K}_{z_{2i}}(r)\bar{K}'_{z_{2i}}(r)dr\right)^{-1} \\
&\times\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\mathfrak{s}_{p_i}\left(\int_0^1 \bar{K}_{z_{2i}}(r)dB_{w_i}(r)\right).
\end{aligned}
\tag{A.27}
$$

Applying the central limit theorem and the law of large numbers to (A.27) by sending N to infinity, we obtain the wanted result. As in Theorem 1, conditions for the law of large numbers and the central limit theorem are trivially satisfied.
(b) Due to Lemmas 7, the uniform integrability conditions for Corollary 1 and Theorem 3 in Phillips and Moon (1998) are satisfied. Moreover, it is straightforward to show that the other conditions for Corollary 1 and Theorem 3 in Phillips and Moon are satisfied. Therefore, the sequential and joint limits are identical as stated.

Lemma 8 *Under Assumption 10,*
(a) $\| Q_T^{-1}\bar{X}_i'\bar{V}_i^{-1}\bar{Z}_iP_T^{-1} \|$ *and* $\| P_T^{-1}\bar{Z}_i'\bar{V}_i^{-1}\bar{Z}_iP_T^{-1} \|$ *are uniformly integrable;*
(b) $\| \frac{1}{\sqrt{T}}\bar{Z}_{1i}'\bar{V}_i^{-1}\bar{v}_i \|^2$ *and* $\| \frac{1}{T}\sum_{i=1}^{N}\bar{Z}_{2i}'\bar{V}_i^{-1}\bar{v}_i \|^2$ *are uniformly integrable.*

Proof) (a) We prove this only for $\| \frac{1}{T^2}\bar{Z}_{2i}'\bar{V}_i^{-1}\bar{Z}_{2i} \|$, because the rest are analogous. Since

$$
\begin{aligned}
&\| \frac{1}{T^2}\bar{Z}_{2i}'\bar{V}_i^{-1}\bar{Z}_{2i} \| \\
= &\| \frac{1}{T^2}(Z_{2i}-i_T\otimes\bar{z}_{2i.}')'\bar{V}_i^{-1}(Z_{2i}-i_T\otimes\bar{z}_{2i.}') \| \\
= &\| \frac{1}{T^2}Z_{2i}'\bar{V}_i^{-1}Z_{2i}-\frac{1}{T^2}\bar{z}_{2i.}i_T'\bar{V}_i^{-1}Z_{2i}-\frac{1}{T^2}Z_{2i}'\bar{V}_i^{-1}i_T\bar{z}_{2i.}' \\
&+\frac{1}{T^2}i_T'\bar{V}_i^{-1}i_T\bar{z}_{2i.}\bar{z}_{2i.}' \|,
\end{aligned}
$$

we obtain the required result if

$$\sup_T E \mid \frac{1}{T^2}\left[Z_{2i}'\bar{V}_i^{-1}Z_{2i}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty, \tag{A.28}$$

$$\sup_T E \mid \frac{1}{T^2}\left[\bar{z}_{2i.}i_T'\bar{V}_i^{-1}Z_{2i}\right]_{(h,j)} \mid^{1+\varepsilon} < \infty \tag{A.29}$$

and

$$\sup_T E \mid \frac{1}{T^2}\left[i_T'\bar{V}_i^{-1}i_T\bar{z}_{2i.}\bar{z}_{2i.}'\right]_{(h,j)} \mid^{1+\varepsilon} < \infty. \tag{A.30}$$

But relation (A.28) is equivalent to (A.21) which was already proven. Letting c_{ik} finite constants, we have

$$\begin{aligned}
\sup_T E \quad \mid \quad & \frac{1}{T^2}\left[\bar{z}_{2i.}i_T'V_i^{-1}Z_{2i}\right]_{(h,j)} \mid^{1+\varepsilon} \\
= \quad & \sup_T \frac{1}{T^{2(1+\varepsilon)}} E \mid \left[\bar{z}_{2i.}\sum_{k=1}^{T} c_{ik}z_{2ik}\right]_{(h,j)} \mid^{1+\varepsilon} \\
\leq \quad & \sup_T \sup_{1\leq i\leq N, 1\leq t\leq T} \mid c_{it} \mid \sup_{1\leq t,s\leq T} \frac{(ts)^{\frac{1+\varepsilon}{2}}}{T^{1+\varepsilon}} E \mid \left[\frac{1}{\sqrt{ts}}z_{2it}z_{2is}'\right]_{(h,j)} \mid^{1+\varepsilon} \\
\leq \quad & \sup_T \sup_{1\leq i\leq N, 1\leq t\leq T} \mid c_{it} \mid \sup_T \sup_{1\leq t,s\leq T} E \mid \left[\frac{1}{\sqrt{ts}}z_{2it}z_{2is}'\right]_{(h,j)} \mid^{1+\varepsilon} < \infty.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\sup_T E \quad \mid \quad & \frac{1}{T^2}\left[i_T'\bar{V}_i^{-1}i_T\bar{z}_{2i.}\bar{z}_{2i.}'\right]_{(h,j)} \mid^{1+\varepsilon} \\
= \quad & \sup_T \mid \frac{1}{T}i_T'\bar{V}_i^{-1}i_T \mid^{1+\varepsilon} E \mid \frac{1}{T}\left[\bar{z}_{2i.}\bar{z}_{2i.}'\right]_{(h,j)} \mid^{1+\varepsilon} \\
\leq \quad & \sup_T \mid \frac{1}{T}i_T'\bar{V}_i^{-1}i_T \mid^{1+\varepsilon} \sup_{1\leq t,s\leq T} \frac{(ts)^{\frac{1+\varepsilon}{2}}}{T^{1+\varepsilon}} E \mid \left[\frac{1}{\sqrt{ts}}z_{2it}z_{2is}'\right]_{(h,j)} \mid^{1+\varepsilon} \\
\leq \quad & \sup_T \mid \frac{1}{T}i_T'\bar{V}_i^{-1}i_T \mid^{1+\varepsilon} \sup_T \sup_{1\leq t,s\leq T} E \mid \left[\frac{1}{\sqrt{ts}}z_{2it}z_{2is}'\right]_{(h,j)} \mid^{1+\varepsilon} < \infty.
\end{aligned}$$

Thus, (A.28), (A.29) and (A.30) hold, from which the result follows.

(b) This follows from Assumption 12 and Theorem 5.4 of Billingsley (1968).

Proof of Theorem 4: (a) The sequential limit results can be proven by adapting the proofs for Lemma A9 and Theorem 4 in Choi (1998), the details for which are omitted.

(b) Using Lemma 8 as in the proof for part (b) of Theorem 3 gives the required result.

Proof of Theorem 5: We prove the result only for W_{WIVO} because the rest can be proven in the same manner. Let $RP_{NT}^{-1} = \tilde{R}$ where $P_{NT} = \begin{bmatrix} \sqrt{NT}I_{k_1} & 0 \\ 0 & \sqrt{N}TI_{k_2} \end{bmatrix}$. Then, we obtain

$$\begin{aligned}
A_{NT} \quad = \quad & \left(\tilde{R}P_{NT}Q_{\hat{V}}\sum_{i=1}^{N}\left[\bar{Z}_i'\hat{U}_{iWIVG}\right]\left[\bar{Z}_i'\hat{U}_{iWIVG}\right]'Q_{\hat{V}}'P_{NT}'\tilde{R}\right)^{-1/2}\tilde{R}P_{NT}(\hat{\alpha}_{WIVO}-r) \mid_{\tilde{R}} \\
\Rightarrow \quad & N(0, I_J)
\end{aligned} \tag{A.31}$$

by using Theorem 2 and the fact that $P_{NT}Q_{\hat{V}}\sum_{i=1}^{N}\left[\bar{Z}_i'\hat{U}_{iWIVG}\right]\left[\bar{Z}_i'\hat{U}_{iWIVG}\right]'Q_{\hat{V}}'P_{NT}'$ is a consistent estimator of the asymptotic variance-covariance matrix for $\hat{\alpha}_{WIVO}$. The conditional result (A.31) is obtained by treating $\tilde{R}$ as if it were a fixed constant matrix. But because the limiting distribution in (A.31) does not depend on $\tilde{R}$, it is also an unconditional distribution. Now, the required result follows from the relation $W_{WIVO} = A_{NT}'A_{NT}$.

Proof of Theorem 6: Result for the DWH_G test can be proven by using the same arguments as given in Hausman (1978) and Hausman and Taylor (1981), the details for which are omitted. Additionally, the other results can be proven exactly as in Choi and Yu (1999).

9 References

Baltagi, B. (1995): *Econometric Analysis of Panel Data*, New York: John Wiley.

Billingsley, P. (1968): *Convergence of Probability Measures*, New York: John Wiley.

Bowden, R. J. and D. A. Turkington (1984): *Instrumental Variables*, New York: Cambridge University Press.

Cavanagh, C. L., G. Elliott and J. H. Stock (1995): "Inference in Models with Nearly Integrated Regressors," *Econometric Theory* 11, 1131-1147.

Choi, I. (1997): "Unit Root Tests for Panel Data," *Journal of International Money and Finance* (forthcoming).

Choi, I. (1998): "Asymptotic Analysis of a Nonstationary Error Component Model," manuscript, Kookmin University, Seoul.

Choi, I. and B. Yu (1999): "Generalized Durbin-Wu-Hausman Tests in the Time and Frequency Domains," manuscript in preparation.

Durbin, J. (1954): "Errors in Variables," *Review of the International Statistical Institute* 22, 23-32.

Fuller, W. (1976): *Introduction to Statistical Time Series,* New York: John Wiley.

Godolphin, E. J. and J. M. Unwin (1983): "Evaluation of the Covariance Matrix for the Maximum Likelihood Estimator of a Gaussian Autoregressive Moving Average Process," *Biometrika* 70, 279-284.

Hannan, E. J. (1970) *Multiple Time Series*, New York: John Wiley.

Hausman, J. (1978): "Specification Tests in Econometrics," *Econometrica* 46, 1251-1271.

Hausman, J. A. and W. E. Taylor (1981): "Panel Data and Unobservable Individual Effects," *Econometrica* 49, 1377-1398.

Im, K. S., M. H. Pesaran and S. Shin (1995): "Testing for Unit Roots in Heterogeneous Panels," mimeo, University of Cambridge.

Kao, C. and M.-H. Chiang (1997): "On the Estimation and Inference of a Cointegrated Regression in Panel Data,"mimeo, Syracuse University.

Levin, A. and C.-F. Lin (1992): "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," UC, San Diego, Discussion Paper 92-23.

Maddala, G. S. and S. Wu (1996), "A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test," mimeo, The Ohio State University.

Newey, W. K. and D. McFadden (1994), "Large Sample Estimation and Hypothesis Testing." In R. Engle and D. McFadden (eds.) Handbook of Econometrics Vol. IV, 2113-2245.

Pedroni, P. (1996): "Fully Modified OLS for Heterogeneous Cointegrated Panels and the Case of Purchasing Power Parity," mimeo, Indiana University.

Pedroni, P. (1997): "On the Estimation and Inference of a Cointegrated Regression in Panel Data," mimeo, Indiana University.

Phillips, P. C. B. (1988): "Multiple Regression with Integrated Processes," in N. U. Prabhu (ed.), *Contemporary Mathematics* Vol. 80. Providence: American Mathematical Society.

Phillips, P. C. B. and B. Hansen (1990): "Statistical Inference in Instrumental Variables Regression with I(1) Processes," *Review of Economic Studies* 57, 99–125.

Phillips, P. C. B. and H. Moon (1999): "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica* 67, 1057-1112.

Saikkonen, P. (1991): "Asymptotically Efficient Estimation of Cointegration Regressions," *Econometric Theory* 9, 1–21.

Wansbeek, T. J. and A. Kapteyn (1982): "A Simple Way to Obtain the Spectral Decomposition of Variance Component Models for Balanced Data," *Communications in Statistics* A11, 2105-2112.

Wansbeek, T. J. and A. Kapteyn (1983): "A Note on Spectral Decomposition and Maximum Likelihood Estimation of ANOVA Models with Balanced Data," *Statistics and Probability Letters* 1, 213-215.

Wu, D.-M. (1973): "Alternative Tests of Independence Between Stochastic Regressors and Disturbances," *Econometrica* 41, 733-750.