Discussion Paper Series

No.213

A Simple Efficient Instrumental Variable Estimator in Panel AR(p) Models

Kazuhiko Hayakawa

May 2007

Hitotsubashi University Research Unit for Statistical Analysis in Social Sciences

A 21st-Century COE Program

Institute of Economic Research
Hitotsubashi University
Kunitachi, Tokyo, 186-8603 Japan
http://hi-stat.ier.hit-u.ac.jp/

A Simple Efficient Instrumental Variable Estimator in Panel AR(p) Models

Kazuhiko Hayakawa*†

Department of Economics, Hitotsubashi University

JSPS Research Fellow

May 13, 2007

Abstract

In this paper, we show that for panel AR(p) models with *iid* errors, an instrumental variable (IV) estimator with instruments in the backward orthogonal deviation has the same asymptotic distribution as the infeasible optimal IV estimator when both N and T, the dimensions of the cross section and the time series, are large. If we assume that the errors are normally distributed, the asymptotic variance of the proposed IV estimator is shown to attain the lower bound when both N and T are large. A simulation study is conducted to assess the estimator.

Keywords: panel AR(p) models, the optimal instruments, the backward orthogonal deviation.

JEL classification: C13, C23.

*E-mail : em031112@ yahoo.co.jp (Remove the space after @)

†The author is grateful to Ryo Okui who posed a question that inspired this paper.

1 Introduction

Since the work of Anderson and Hsiao (1981, 1982), instrumental variables have been widely used for the estimation of dynamic panel data models.[1] However, since the IV estimator is not generally efficient, Holtz-Eakin, Newey, and Rosen (1988) and Arellano and Bond (1991) proposed to use the generalized method of moments (GMM) estimator to improve efficiency. The GMM estimator has subsequently been refined in a number of studies, including Arellano and Bover (1995), Ahn and Schmidt (1995, 1997) and Blundell and Bond (1998). However, although the GMM estimator is generally more efficient than the IV estimator, it is well known that the GMM estimator is more biased than the IV estimator in finite sample.

In this paper, we focus on the IV estimator and address the efficiency problem of the IV estimator. Specifically, we show that, for panel AR(p) models with *iid* errors, a simple one-step IV estimator is obtained from the backward orthogonal deviation (BOD) transformation that has the same asymptotic distribution as the infeasible optimal IV estimator derived by Arellano (2003b) when both N and T are large. If normality is assumed on the errors, the proposed IV estimator is shown to be asymptotically efficient. Simulation results reveal that the proposed IV estimator is almost unbiased, and the difference in dispersions between the feasible optimal IV estimator and the proposed IV estimator is small when T is large.

The remainder of this paper is organized as follows. Section 2 provides the setup and the main result. Section 3 presents a Monte Carlo simulation and assess the theoretical result. Finally, Section 4 concludes.

A word on notation. For a vector $\boldsymbol{x}$ and a matrix $\boldsymbol{A}$, we define $\|\boldsymbol{x}\|^2 = \boldsymbol{x}'\boldsymbol{x}$ and $\|\boldsymbol{A}\|^2 = tr(\boldsymbol{A}'\boldsymbol{A})$ where $tr(\cdot)$ denotes the trace operator.

[1] Recent papers that discuss the IV estimator are Arellano (2003b) and Hahn, Hausman, and Kuersteiner (2007), proposing two-step efficient IV estimators and the long difference IV estimator are proposed, respectively.

2 Setup and Result

2.1 The model and assumptions

Let us consider the following panel AR(p) model:

$$y_{it} = \alpha_1 y_{i,t-1} + \alpha_2 y_{i,t-2} + \cdots + \alpha_p y_{i,t-p} + \eta_i + v_{it} \tag{1}$$

$$= \boldsymbol{\alpha}' \boldsymbol{x}_{it} + \eta_i + v_{it} \qquad (i = 1, ..., N, \quad t = 1, ..., T) \tag{2}$$

where $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)'$, $\boldsymbol{x}_{it} = (y_{i,t-1}, ..., y_{i,t-p})'$, v_{it} has zero mean given by $\eta_i, y_{i0}, ..., y_{i,t-1}$ and p is fixed and known.[2] For convenience, we assume that $y_{i,0}$,...,$y_{i,1-p}$ are observed.

(2) can be written in a companion form as

$$\boldsymbol{x}_{i,t+1} = \boldsymbol{\Pi} \boldsymbol{x}_{it} + \boldsymbol{d}_1 (\eta_i + v_{it}) \tag{3}$$

where $\boldsymbol{d}_1 = (1, 0, ..., 0)'$ of dimension p and $\boldsymbol{\Pi}$ is the $p \times p$ matrix given by

$$\boldsymbol{\Pi} = \begin{pmatrix} \alpha_1 & \cdots & \alpha_p \\ \boldsymbol{I}_{p-1} & | & \boldsymbol{O}_{(p-1)\times 1} \end{pmatrix} \tag{4}$$

where $\boldsymbol{I}_k$ is an identity matrix of order k and $\boldsymbol{O}_{k\times \ell}$ is a $k \times \ell$ matrix of zeros.

We make the following assumptions, which are part of the assumptions made by Lee (2005).

Assumption 1. *$\{v_{it}\}$ $(t = 1, ..., T, i = 1, ..., N)$ are iid over i and t and independent of η_i and $\boldsymbol{x}_{i1}$, with $E(v_{it}) = 0$, $var(v_{it}) = \sigma_v^2$ and finite fourth order moment. $\{\eta_i\}(i = 1, ..., N)$ are iid over i with $E(\eta_i) = 0$ and $var(\eta) = \sigma_\eta^2$.*

Assumption 2. *The initial observations satisfy*

$$\boldsymbol{x}_{i1} = (\boldsymbol{I}_p - \boldsymbol{\Pi})^{-1} \boldsymbol{d}_1 \eta_i + \boldsymbol{w}_{i0} \tag{5}$$

where $\boldsymbol{w}_{i0} = \left(\sum_{j=0}^{\infty} \boldsymbol{\Pi}^j v_{i,-j} \right) \boldsymbol{d}_1$.

Assumption 3. *$det\,[\boldsymbol{I}_p - \boldsymbol{\Pi} z] \neq 0$ for all $|z| \leq 1$.*

Assumption 4. *Let $m_j(i,t) = \boldsymbol{\Pi}^j \boldsymbol{d}_1 v_{i,t-1-j}$. For all i, t, and for any $r_1, ..., r_4 \in \{1, 2, \cdots p\}$,*

$$\sum_{j_1, \cdots, j_4 = 0}^{\infty} |cum_{r_1, \cdots, r_4} (m_{j_1}(i,t), m_{j_2}(i,t), m_{j_3}(i,t), m_{j_4}(i,t))| < \infty. \tag{6}$$

[2] The problem how to choose p is extensively discussed by Lee (2005).

Unlike Lee (2005), we do not need to impose the asymptotic relative ratio between N and T. Assumptions 1 and 2 are standard ones in the literature.[3] Although Assumption 2 can be relaxed to nonstationary initial conditions, we do not pursue this here for the purpose of simplicity. However, the main result of this paper is expected to hold since the initial conditions are negligible when T is large and since we do not use moment conditions that rely on stationary initial conditions as Blundell and Bond (1998) do. Assumption 3 is the stability condition, and Assumption 4 is necessary to use the central limit theorem for double indexed processes.[4]

Under Assumptions 2 and 3, $\boldsymbol{x}_{it}$ can be written as

$$\boldsymbol{x}_{i,t+1} = (\boldsymbol{I}_p - \boldsymbol{\Pi})^{-1}\boldsymbol{d}_1\eta_i + \boldsymbol{w}_{it} \tag{7}$$

where

$$\boldsymbol{w}_{it} = \left(\sum_{j=0}^{\infty} \boldsymbol{\Pi}^j v_{i,t-j}\right)\boldsymbol{d}_1. \tag{8}$$

The model to be estimated is given by

$$y_{it}^* = \boldsymbol{\alpha}'\boldsymbol{x}_{it}^* + v_{it}^* \qquad (i=1,...,N, \quad t=1,...,T-1) \tag{9}$$

where $y_{it}^* = c_t\left[y_{it} - (y_{i,t+1} + \cdots + y_{iT})/(T-t)\right]$, $\boldsymbol{x}_{it}^* = c_t\left[\boldsymbol{x}_{it} - (\boldsymbol{x}_{i,t+1} + \cdots + \boldsymbol{x}_{iT})/(T-t)\right]$, $v_{it}^* = c_t\left[v_{it} - (v_{i,t+1} + \cdots + v_{iT})/(T-t)\right]$, and $c_t^2 = (T-t)/(T-t+1)$.

2.2 The instrumental variable estimators

The infeasible optimal instruments

Following Arellano (2003a, b), the infeasible optimal IV estimator in a large N and small T context takes the following form:

$$\widetilde{\boldsymbol{\alpha}}_{opt} = \left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\boldsymbol{h}_{it}\boldsymbol{x}_{it}^{*\prime}\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\boldsymbol{h}_{it}y_{it}^*\right) \tag{10}$$

where $\boldsymbol{h}_{it} = E(\boldsymbol{x}_{it}^*|\boldsymbol{y}_i^{t-1})$ and $\boldsymbol{y}_i^{t-1} = (y_{i,t-1}, ..., y_{i,0})'$. One of the feasible optimal IV estimators is obtained using a sample linear projection of $\boldsymbol{h}_{it}$, which is given by

$$\widehat{\boldsymbol{h}}_{it}^{LEV} = \left(\sum_{i=1}^{N}\boldsymbol{x}_{it}^*\boldsymbol{y}_i^{t-1\prime}\right)\left(\sum_{i=1}^{N}\boldsymbol{y}_i^{t-1}\boldsymbol{y}_i^{t-1\prime}\right)^{-1}\boldsymbol{y}_i^{t-1}. \tag{11}$$

[3] See Alvarez and Arellano (2003) for the AR(1) case.

[4] See Phillips and Moon (1999) and Hahn and Kuersteiner (2002).

In this case, the feasible optimal IV estimator is equivalent to the GMM estimator using $\boldsymbol{y}_i^{t-1}$ as instruments:

$$\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV} = \left(\frac{1}{NT}\sum_{t=1}^{T-1}\boldsymbol{X}_t^{*\prime}\boldsymbol{M}_t^{LEV}\boldsymbol{X}_t^*\right)^{-1}\left(\frac{1}{NT}\sum_{t=1}^{T-1}\boldsymbol{X}_t^{*\prime}\boldsymbol{M}_t^{LEV}\boldsymbol{y}_t^*\right) \tag{12}$$

where $\boldsymbol{X}_t^* = (\boldsymbol{x}_{1t}^*, ..., \boldsymbol{x}_{Nt}^*)'$, $\boldsymbol{M}_t^{LEV} = \boldsymbol{Z}_t^{LEV}(\boldsymbol{Z}_t^{LEV\prime}\boldsymbol{Z}_t^{LEV})^{-1}\boldsymbol{Z}_t^{LEV\prime}$, $\boldsymbol{Z}_t^{LEV} = (\boldsymbol{y}_1^{t-1}, ..., \boldsymbol{y}_N^{t-1})'$, and $\boldsymbol{y}_t^* = (y_{1t}^*, ..., y_{Nt}^*)'$.

One problem of $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ is that if N and T increase at the same rate, the estimate of $\boldsymbol{h}_{it}^{LEV}$ is asymptotically biased (see Arellano 2003a, p.170). This causes a bias in $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$. In fact, for the case of $p = 1$, Alvarez and Arellano (2003) show that $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ has a bias of the order $O(1/N)$.[5]

Using the structure of AR(p) models, Arellano (2003b) shows that the infeasible optimal IV $\boldsymbol{h}_{it}$ can be rewritten in the following form:

$$E(\boldsymbol{x}_{it}^*|\boldsymbol{y}_i^{t-1}) = c_t\left[\boldsymbol{I}_p - \frac{1}{T-t}\boldsymbol{\Pi}(\boldsymbol{I}_p - \boldsymbol{\Pi}^{T-t})(\boldsymbol{I}_p - \boldsymbol{\Pi})^{-1}\right]\left[\boldsymbol{x}_{it} - \boldsymbol{\iota}_p E(\mu_i|\boldsymbol{y}_i^{t-1})\right]. \tag{13}$$

Under the assumption that $E(\mu_i|\boldsymbol{y}_i^{t-1})$ coincides with the linear projection, we have

$$E(\mu_i|\boldsymbol{y}_i^{t-1}) = \frac{\phi}{1+\phi(\boldsymbol{\iota}_t'\boldsymbol{V}_t^{-1}\boldsymbol{\iota}_t)}\boldsymbol{\iota}_t'\boldsymbol{V}_t^{-1}\boldsymbol{y}_i^{t-1} \tag{14}$$

where $\phi = \sigma_\mu^2/\sigma_v^2$, $\boldsymbol{V}_t = \sigma_v^{-2}E\left[(\boldsymbol{y}_i^{t-1} - \mu_i\boldsymbol{\iota_t})(\boldsymbol{y}_i^{t-1} - \mu_i\boldsymbol{\iota_t})\right]'$, $\mu_i = \eta_i/(1-\boldsymbol{\alpha}'\boldsymbol{\iota}_p)$, and $\sigma_\mu^2 = var(\mu_i)$. Hence, the infeasible optimal IV estimator is given by

$$\widehat{\boldsymbol{\alpha}}_{IV}^{OPT} = \left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\boldsymbol{h}_{it}^{OPT}\boldsymbol{x}_{it}^{*\prime}\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\boldsymbol{h}_{it}^{OPT}y_{it}^*\right) \tag{15}$$

$$= \boldsymbol{\alpha} + \left(\widehat{\boldsymbol{A}}_{IV}^{OPT}\right)^{-1}\widehat{\boldsymbol{b}}_{IV}^{OPT} \tag{16}$$

where

$$\boldsymbol{h}_{it}^{OPT} = c_t\left[\boldsymbol{I}_p - \frac{1}{T-t}\boldsymbol{\Pi}(\boldsymbol{I}_p - \boldsymbol{\Pi}^{T-t})(\boldsymbol{I}_p - \boldsymbol{\Pi})^{-1}\right]\left[\boldsymbol{x}_{it} - \boldsymbol{\iota}_p\frac{\phi}{1+\phi(\boldsymbol{\iota}_t'\boldsymbol{V}_t^{-1}\boldsymbol{\iota}_t)}\boldsymbol{\iota}_t'\boldsymbol{V}_t^{-1}\boldsymbol{y}_i^{t-1}\right]. \tag{17}$$

[5] Also see Bun and Kiviet (2006).

Instruments in the backward orthogonal deviation

We consider the IV estimator using instruments transformed by the BOD transformation.[6] Specifically, let us define the variables in the backward orthogonal deviation as follows:

$$\boldsymbol{x}_{it}^{**} = \frac{1}{c_t}\left[\boldsymbol{x}_{it} - \frac{\boldsymbol{x}_{i,t-1} + \cdots + \boldsymbol{x}_{i1}}{t-1}\right] \qquad t = 2, ..., T-1. \tag{18}$$

Since $\boldsymbol{x}_{it}^{**}$ contains all past values of $\boldsymbol{x}_{it}$, it is expected that linear projection of $\boldsymbol{x}_{it}^{*}$ on $\boldsymbol{x}_{it}^{**}$ has the same information as that of $\boldsymbol{x}_{it}^{*}$ on $\boldsymbol{y}_i^{t-1}$. Furthermore, we find that the second parenthesis in (13) can be regarded as demeaning, while the BOD transformation is a demeaning transformation.[7] Thus, we find that $\boldsymbol{x}_{it}^{**}$ has a similar structure as $\boldsymbol{h}_{it}^{OPT}$.

The IV estimator using $\boldsymbol{x}_{it}^{**}$ as instruments is given by

$$\begin{aligned}
\widehat{\boldsymbol{\alpha}}_{IV}^{BOD} &= \left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=2}^{T-1}\boldsymbol{x}_{it}^{**}\boldsymbol{x}_{it}^{*\prime}\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=2}^{T-1}\boldsymbol{x}_{it}^{**}y_{it}^{*}\right) & (20)\\
&= \boldsymbol{\alpha} + \left(\widehat{\boldsymbol{A}}_{IV}^{BOD}\right)^{-1}\widehat{\boldsymbol{b}}_{IV}^{BOD}. & (21)
\end{aligned}$$

The following proposition establishes the asymptotic equivalence of the infeasible optimal IV estimator, $\widehat{\boldsymbol{\alpha}}_{IV}^{OPT}$, and $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ in the sense that both estimators have the same asymptotic distribution.

Proposition 1. *Let Assumptions 1, 2, and 3 hold. Then, as both N and T tend to infinity, the infeasible optimal IV estimator $\widehat{\boldsymbol{\alpha}}_{IV}^{OPT}$ and the feasible IV estimator $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ are consistent. If we further assume that Assumption 4 holds, then, as both N and T tend to infinity, we have*

$$\sqrt{NT}\left(\widehat{\boldsymbol{\alpha}}_{IV} - \boldsymbol{\alpha}\right) \to^d N\left(0, \sigma_v^2\left[E(\boldsymbol{w}_{i,t-1}\boldsymbol{w}_{i,t-1}^{\prime})\right]^{-1}\right) \tag{22}$$

where $\widehat{\boldsymbol{\alpha}}_{IV}$ denotes $\widehat{\boldsymbol{\alpha}}_{IV}^{OPT}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$.

Note that the asymptotic variance $\sigma_v^2\left[E(\boldsymbol{w}_{i,t-1}\boldsymbol{w}_{i,t-1}^{\prime})\right]^{-1}$ is of the same form as the within groups (WG) estimator derived by Lee (2005).

[6] The BOD transformation was originally considered by So and Shin (1999) in a time series context.

[7] Note that $\boldsymbol{x}_{it}$ can be rewritten as

$$\boldsymbol{x}_{it} = \boldsymbol{\iota}_p\mu_i + \boldsymbol{w}_{i,t-1} \tag{19}$$

since $(\boldsymbol{I}_p - \boldsymbol{\Pi})^{-1}\boldsymbol{d}_1 = \frac{1}{1-\boldsymbol{\alpha}'\boldsymbol{\iota}_p}\boldsymbol{\iota}_p$.

Remark 1. For the case of $p = 1$, Alvarez and Arellano (2003) show that $\widehat{\alpha}_{GMM}^{LEV}$ and the WG estimator, $\widehat{\alpha}_{WG}$, has the following asymptotic distribution:

$$\sqrt{NT}\left[\widehat{\alpha}_{GMM}^{LEV} - \left(\alpha_1 - \frac{1}{N}(1+\alpha_1)\right)\right] \to^d N\left(0, 1-\alpha_1^2\right), \tag{23}$$

$$\sqrt{NT}\left[\widehat{\alpha}_{WG} - \left(\alpha_1 - \frac{1}{T}(1+\alpha_1)\right)\right] \to^d N\left(0, 1-\alpha_1^2\right). \tag{24}$$

Also, from Proposition 1, we have

$$\sqrt{NT}\left[\widehat{\alpha}_{IV}^{BOD} - \alpha_1\right] \to^d N\left(0, 1-\alpha_1^2\right). \tag{25}$$

Comparing (23), (24) and (25), we find that although all estimators have the same asymptotic variance, $\widehat{\alpha}_{GMM}^{LEV}$ and $\widehat{\alpha}_{WG}$ have asymptotic biases of the order $O(1/N)$ and $O(1/T)$, respectively, while $\widehat{\alpha}_{IV}^{BOD} - \alpha_1$ is centered at zero.

Remark 2. Hahn and Kuersteiner (2002) show that if we further assume normality on v_{it}, then $\sigma_v^2\left[E(\boldsymbol{w}_{i,t-1}\boldsymbol{w}_{i,t-1}')\right]^{-1}$ is equal to the lower bound under large N and large T asymptotics. Hence, $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ is an efficient IV estimator under large N and large T asymptotics without an asymptotic bias when v_{it} is normally distributed.

Remark 3. Another advantage of $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ is that since the individual effects are completely eliminated from both the model and instruments under stationary initial conditions, the performance of $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ is not affected by the variance ratio of the individual effects to the disturbances although the typical GMM estimators using instruments in levels are.[8]

3 Monte Carlo Simulation

In this section, we compare $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ with other estimators by Monte Carlo simulation. We consider AR(1) and AR(2) models. v_{it} and η_i are drawn from $N(0,1)$ independently. We consider the cases of $(T,N) = (10,100)$, $(10,500)$, $(15,100)$, $(15,300)$, $(20,100)$, $(20,200)$, $(50,100)$, and $(100,100)$. For the AR(1) model, we set $\alpha_1 = 0.3, 0.6, 0.9$, and for the AR(2) model, we set $(\alpha_1,\alpha_2) = (0.45, 0.45), (0.6, 0.3)$. We generate $T + p + 50$ observations for each i and discard the first 50 periods to diminish the effect of initial conditions. We compute the median (Median), the

[8] See Bun and Kiviet (2006), Hayakawa (2007a), and Bun and Windmeijer (2007).

interquartile range (IQR), and the median absolute error (MAE). The number of replications is 5000 for all cases.

The estimators to be compared are $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$, $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$, the GMM estimator using $\boldsymbol{x}_{it}^{**}$ as instruments, and the IV estimator using $\boldsymbol{x}_{it}$ as instruments. The GMM estimator where $\boldsymbol{x}_{it}^{**}$ are used as instruments is defined by

$$\widehat{\boldsymbol{\alpha}}_{GMM}^{BOD} = \left(\frac{1}{NT}\sum_{t=2}^{T-1}\boldsymbol{X}_t^{*\prime}\boldsymbol{M}_t^{BOD}\boldsymbol{X}_t^{*}\right)^{-1}\left(\frac{1}{NT}\sum_{t=2}^{T-1}\boldsymbol{X}_t^{*\prime}\boldsymbol{M}_t^{BOD}\boldsymbol{y}_t^{*}\right) \tag{26}$$

where $\boldsymbol{M}_t^{BOD} = \boldsymbol{Z}_t^{BOD}\left(\boldsymbol{Z}_t^{BOD\prime}\boldsymbol{Z}_t^{BOD}\right)^{-1}\boldsymbol{Z}_t^{BOD\prime}$, and $\boldsymbol{Z}_t^{BOD} = (\boldsymbol{x}_{1t}^{**}, ..., \boldsymbol{x}_{Nt}^{**})'$.

$\widehat{\boldsymbol{\alpha}}_{GMM}^{BOD}$ does not share the problem with $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ that the number of parameters increases as T gets larger. Although we suspect that discarding some available instruments results in an efficiency loss, for the case of $p = 1$, Hayakawa (2007b) shows that $\widehat{\boldsymbol{\alpha}}_{GMM}^{BOD}$ has the same asymptotic variance as $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$, while its asymptotic bias is of the order $O(1/NT)$.[9]

The IV estimator using $\boldsymbol{x}_{it}$ as instruments is given by

$$\widehat{\boldsymbol{\alpha}}_{IV}^{LEV} = \left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\boldsymbol{x}_{it}\boldsymbol{x}_{it}^{*\prime}\right)^{-1}\left(\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\boldsymbol{x}_{it}y_{it}^{*}\right). \tag{27}$$

Note that $\widehat{\boldsymbol{\alpha}}_{IV}^{LEV}$ is not exactly the same IV estimator as the one by Anderson and Hsiao (1981, 1982) since they used the first-difference to remove the individual effects from the model.

The simulation results for AR(1) and AR(2) model are provided in Tables 1 and 2, respectively. We first consider the AR(1) case. We find from Table 1 that, in terms of the bias, the IV estimators, $\widehat{\boldsymbol{\alpha}}_{IV}^{LEV}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$, have little bias for all cases, while the GMM estimators have non-negligible bias when $\alpha = 0.9$ and T is less than 15. Especially $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ has large bias for all cases. However, with regard to the IQR, $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ has the smallest dispersion and $\widehat{\boldsymbol{\alpha}}_{IV}^{LEV}$ has the largest dispersion. Also, we find that the differences in the IQR of $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$, $\widehat{\boldsymbol{\alpha}}_{GMM}^{BOD}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ become quite small when T is as large as 50. This result is consistent with Proposition 1 where $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$, which is a feasible optimal IV estimator, and $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ are shown to have the same asymptotic variance when N and T are large. For the median absolute error,

[9] Although we expect that similar results hold for AR(p) models, we do not provide a proof here since it would become quite long.

we find that $\widehat{\boldsymbol{\alpha}}_{GMM}^{BOD}$ has the smallest MAE in many cases. However, the difference in the MAE between $\widehat{\boldsymbol{\alpha}}_{GMM}^{BOD}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ is fairly small. Next, we discuss the results for the AR(2) case. The IV estimators are virtually median unbiased and $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ has the largest bias. In terms of the IQR, unlike in the AR(1) case, $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ is not least dispersed for all cases. For instance, in the case of $T \geq 20$, the IQR of $\widehat{\boldsymbol{\alpha}}_{GMM}^{BOD}$ is smaller than that of $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ in almost all cases. Also, we find that the difference in the IQR between $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$, $\widehat{\boldsymbol{\alpha}}_{GMM}^{BOD}$, and $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ becomes small when T is large. In terms of the MAE, although $\widehat{\boldsymbol{\alpha}}_{GMM}^{BOD}$ performs best in many cases, the difference between $\widehat{\boldsymbol{\alpha}}_{GMM}^{BOD}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ is quite small.

4 Conclusion

In this paper, we showed that the infeasible optimal IV estimator and the IV estimator using instruments in the backward orthogonal deviation are asymptotically equivalent in the sense that both estimators have the same asymptotic distribution when both N and T are large. We further showed that if we assume normality on the errors, the proposed IV estimator is asymptotically efficient when both N and T are large. Simulation results demonstrated that in terms of the bias and median absolute error, the new IV estimator outperforms the GMM and IV estimators using instruments in levels, which are commonly used in the literature.

Lastly, we note some possible extensions. Although we considered an AR(p) model with *iid* errors, it is of great interest to investigate whether the results obtained in this paper apply to more general models and errors, say, models that include additional regressors besides the lagged dependent variables (Arellano, 2003b) and/or heteroskedastic errors (Alvarez and Arellano, 2004). Also, it may be interesting to apply Okui's (2006) method, i.e., a procedure to select the number of moment conditions so as to minimize the MSE of the estimators, to improve the GMM/IV estimators using instruments in the backward orthogonal deviation. But these tasks are left for future research.

Appendix

Lemma A *Let Assumptions 1, 2, and 3 hold. Then,* $\boldsymbol{h}_{it}^{OPT}$ *and* $\boldsymbol{h}_{it}^{BOD}$ *can be written as*

$$\begin{aligned}\boldsymbol{h}_{it}^{OPT} &= c_t\left[\boldsymbol{I}_t - O\left(\frac{1}{T-t}\right)\right]\left[\boldsymbol{w}_{i,t-1} + \boldsymbol{g}_{it}^{OPT}\right], \\ \boldsymbol{h}_{it}^{BOD} &= \boldsymbol{w}_{it-1} - \boldsymbol{g}_{it}^{BOD} \qquad (28)\end{aligned}$$

where

$$\boldsymbol{g}_{it}^{OPT} = \boldsymbol{\iota}_p\left[\frac{\mu_i\left(1+\phi\boldsymbol{\kappa}_p'\boldsymbol{R}_i^{22}\boldsymbol{\kappa}_p\right) - \phi\left[(1-\boldsymbol{\alpha}'\boldsymbol{\iota}_p)(v_{i,t-1}+\cdots+v_{i,p}) + \boldsymbol{\kappa}_p'\boldsymbol{R}_i^{22}\boldsymbol{\zeta}_i\right]}{1+\phi\left\{(1-\boldsymbol{\alpha}'\boldsymbol{\iota}_p)^2(t-p) + \boldsymbol{\kappa}_p'\boldsymbol{R}_i^{22}\boldsymbol{\kappa}_p\right\}}\right], \qquad (29)$$

$$\boldsymbol{g}_{it}^{BOD} = \frac{\left(\boldsymbol{\Phi}_1 v_{i,t-2} + \cdots + \boldsymbol{\Phi}_{t-2}v_{i1}\right)\boldsymbol{d}_1 + \boldsymbol{\Phi}_{t-1}\boldsymbol{w}_{i0}}{t-1}, \qquad (30)$$

$$\boldsymbol{\Phi}_j = \boldsymbol{\Pi}^0 + \boldsymbol{\Pi}^1 + \cdots + \boldsymbol{\Pi}^{j-1} = (\boldsymbol{I}_p - \boldsymbol{\Pi})^{-1}(\boldsymbol{I}_p - \boldsymbol{\Pi}^j) \qquad (31)$$

and $\boldsymbol{\kappa}_p$, $\boldsymbol{R}_i^{22}$, *and* $\boldsymbol{\zeta}_i$ *are defined later.*

Proof of Lemma A Following Whittle (1951) and Wise (1955), let us define the $t \times t$ matrix $\boldsymbol{U}_t$ as follows.

$$\boldsymbol{U}_t = \begin{bmatrix} \boldsymbol{O}_{(t-1)\times 1} & \boldsymbol{I}_{t-1} \\ \boldsymbol{O}_{1\times 1} & \boldsymbol{O}_{1\times(t-1)} \end{bmatrix}. \qquad (32)$$

Then, we have

$$\boldsymbol{U}_t^2 = \begin{bmatrix} \boldsymbol{O}_{(t-1)\times 2} & \boldsymbol{I}_{t-2} \\ \boldsymbol{O}_{2\times 2} & \boldsymbol{O}_{2\times(t-2)} \end{bmatrix}, \quad \boldsymbol{U}_t^3 = \begin{bmatrix} \boldsymbol{O}_{(t-3)\times 3} & \boldsymbol{I}_{t-3} \\ \boldsymbol{O}_{3\times 3} & \boldsymbol{O}_{3\times(t-3)} \end{bmatrix}, \quad \cdots$$

$$\boldsymbol{U}_t^{p-1} = \begin{bmatrix} \boldsymbol{O}_{(t-p+1)\times(p-1)} & \boldsymbol{I}_{t-p+1} \\ \boldsymbol{O}_{(p-1)\times(p-1)} & \boldsymbol{O}_{(p-1)\times(t-p+1)} \end{bmatrix}, \quad \boldsymbol{U}_t^{p} = \begin{bmatrix} \boldsymbol{O}_{(t-p)\times p} & \boldsymbol{I}_{t-p} \\ \boldsymbol{O}_{p\times p} & \boldsymbol{O}_{p\times(t-p)} \end{bmatrix}.$$

Using these expressions, $\boldsymbol{y}_i^{t-1}$ can be written as

$$\boldsymbol{y}_i^{t-1} = \alpha_1\boldsymbol{U}_t\boldsymbol{y}_i^{t-1} + \alpha_2\boldsymbol{U}_t^2\boldsymbol{y}_i^{t-1} + \cdots + \alpha_p\boldsymbol{U}_t^p\boldsymbol{y}_i^{t-1} + \eta_i\begin{bmatrix}\boldsymbol{\iota}_{t-1} \\ 0\end{bmatrix} + \boldsymbol{v}_i^{t-1} + \boldsymbol{r}_i^{t-1} \qquad (33)$$

where $\boldsymbol{v}_i^{t-1} = (v_{i,t-1}, ..., v_{i,1}, 0)' = \left(\boldsymbol{v}_{(1),i}', \boldsymbol{v}_{(2),i}'\right)'$, $\boldsymbol{v}_{(1),i} = (v_{i,t-1}, ..., v_{i,p})'$, $\boldsymbol{v}_{(2),i} = (v_{i,p-1}, ..., v_{i1}, 0)'$, $\boldsymbol{r}_i^{t-1} = \left(\boldsymbol{r}_{(1),i}', \boldsymbol{r}_{(2),i}'\right)'$, $\boldsymbol{r}_{(1),i} = \boldsymbol{O}_{(t-p)\times 1}$, and $\boldsymbol{r}_{(2),i} = (\alpha_p y_{i,-1}, \alpha_{p-1}y_{i,-1}+$

$\alpha_1 y_{i,-2}, \cdots\cdots, \alpha_2 y_{i,-1} + \cdots + \alpha_p y_{i,-p+1}, y_{i,0})$. Since y_{it} is stationary and its conditional mean given by η_i is $\mu_i = \eta_i/(1-\boldsymbol{\alpha}'\boldsymbol{\iota_p})$,

$$
\begin{aligned}
(\boldsymbol{I}_t - \boldsymbol{\Delta}_t)\,\widetilde{\boldsymbol{y}}_i^{t-1} &= \eta_i \begin{bmatrix} \boldsymbol{\iota}_{t-1} \\ 0 \end{bmatrix} - \mu_i(\boldsymbol{I}_t - \boldsymbol{\Delta}_t)\boldsymbol{\iota}_t + \boldsymbol{v}_i^{t-1} + \boldsymbol{r}_i^{t-1} \\
&= \eta_i \begin{bmatrix} \boldsymbol{0}_{(t-p)\times 1} \\ \begin{pmatrix} \boldsymbol{\iota}_{p-1} \\ 0 \end{pmatrix} - \frac{\boldsymbol{\kappa}_p}{1-\boldsymbol{\alpha}'\boldsymbol{\iota}_p} \end{bmatrix} + \boldsymbol{v}_i^{t-1} + \boldsymbol{r}_i^{t-1} \\
&= \boldsymbol{v}_i^{t-1} + \widetilde{\boldsymbol{r}}_i^{t-1} \\
&= \boldsymbol{R}_i^{t-1}
\end{aligned}
$$

where $\widetilde{\boldsymbol{y}}_i^{t-1} = \boldsymbol{y}_i^{t-1} - \mu_i\boldsymbol{\iota}_t$, $\boldsymbol{\Delta}_t = \left(\alpha_1\boldsymbol{U}_t + \alpha_2\boldsymbol{U}_t^2 + \cdots + \alpha_p\boldsymbol{U}_t^p\right)$, $\widetilde{\boldsymbol{r}}_i^{t-1} = (\boldsymbol{0}_{1\times(t-p)}, \widetilde{\boldsymbol{r}}'_{(2),i})'$, and

$$
\widetilde{\boldsymbol{r}}_{(2),i} = \begin{bmatrix} \alpha_p(y_{i,-1}-\mu_i) \\ \alpha_{p-1}(y_{i,-1}-\mu_i) + \alpha_p(y_{i,-2}-\mu_i) \\ \vdots \\ \alpha_2(y_{i,-1}-\mu_i) + \cdots + \alpha_p(y_{i,-p+1}-\mu_i) \\ y_{i0}-\mu_i \end{bmatrix}.
$$

Then, it follows that

$$
\boldsymbol{V}_t^{-1} = \left(\boldsymbol{I}_t - \boldsymbol{\Delta_t}'\right) \begin{bmatrix} \boldsymbol{I}_{t-p} & O \\ O & \boldsymbol{R}_i^{22} \end{bmatrix} (\boldsymbol{I}_t - \boldsymbol{\Delta}_t) \tag{34}
$$

where $\boldsymbol{R}_i^{22} = \left[\sigma_v^{-2}E(\widetilde{\boldsymbol{r}}_{(2),i}\widetilde{\boldsymbol{r}}'_{(2),i})\right]^{-1}$.

Therefore, using

$$
(\boldsymbol{I}_t - \boldsymbol{\Delta}_t)\,\boldsymbol{\iota}_t = \begin{bmatrix} \boldsymbol{\iota}_{t-p}(1-\boldsymbol{\alpha}'\boldsymbol{\iota}_p) \\ 1-\alpha_1-\alpha_2-\cdots-\alpha_{p-2}-\alpha_{p-1} \\ 1-\alpha_1-\alpha_2-\cdots-\alpha_{p-2} \\ \vdots \\ 1-\alpha_1 \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{\iota}_{t-p}(1-\boldsymbol{\alpha}'\boldsymbol{\iota}_p) \\ \boldsymbol{\kappa}_p \end{bmatrix} \tag{35}
$$

we have

$$
\boldsymbol{\iota}_t'\boldsymbol{V}_t^{-1}\boldsymbol{\iota}_t = (1-\boldsymbol{\alpha}'\boldsymbol{\iota}_p)^2(t-p) + \boldsymbol{\kappa}_p'\boldsymbol{R}_i^{22}\boldsymbol{\kappa}_p, \tag{36}
$$

$$\boldsymbol{\iota}_t' \boldsymbol{V}_t^{-1} \boldsymbol{y}_i^{t-1} \quad = \quad (1 - \boldsymbol{\alpha}' \boldsymbol{\iota}_p) \left[\eta_i (t-p) + v_{i,t-1} + \cdots + v_{i,p} \right] + \boldsymbol{\kappa}_p' \boldsymbol{R}_i^{22} \boldsymbol{\zeta}_i \tag{37}$$

where

$$\boldsymbol{\zeta}_i = \eta_i \begin{pmatrix} \boldsymbol{\iota}_{p-1} \\ 0 \end{pmatrix} + \boldsymbol{v}_{(2)i} + \boldsymbol{r}_{(2)i}. \tag{38}$$

The result for $\boldsymbol{h}_{it}^{BOD}$ is readily obtained after a simple manipulation.

Lemma B *Let Assumptions 1, 2, and 3 hold. Then,* $\left\| E(\boldsymbol{g}_{it}^{OPT} \boldsymbol{w}_{i,t-1}') \right\|$ *and* $\left\| E(\boldsymbol{g}_{it}^{BOD} \boldsymbol{w}_{i,t-1}') \right\|$ *are* $O(1/t)$.

Proof of Lemma B First, note that $E(\mu_i \boldsymbol{w}_{i,t-1}) = \boldsymbol{O}_{p \times 1}$. Next, since p is fixed, we have

$$\begin{aligned}
&\left\| E\left[(v_{i,t-1} + \cdots + v_{i,p}) \boldsymbol{w}_{i,t-1}' \right] \right\| = \sigma_v^2 \left\| \boldsymbol{d}_1' \left[(\boldsymbol{I}_p - \boldsymbol{\Pi})^{-1} (\boldsymbol{I}_p - \boldsymbol{\Pi}^{t-p}) \right]' \right\| = O(1), \\
&\left\| E\left[\boldsymbol{\kappa}_p' \boldsymbol{R}_i^{22} (\boldsymbol{\zeta}_i) \boldsymbol{w}_{i,t-1}' \right] \right\| = O(1), \\
&\left\| E\left[(\boldsymbol{\Phi}_1 v_{i,t-2} + \cdots + \boldsymbol{\Phi}_{t-2} v_{i1}) \, \boldsymbol{d}_1 \boldsymbol{w}_{i,t-1}' \right] \right\| = \sigma_v^2 \left\| \sum_{j=1}^{t-2} \boldsymbol{\Phi}_j \boldsymbol{d}_1 \boldsymbol{d}_1' (\boldsymbol{\Pi}')^j \right\| = O(1).
\end{aligned}$$

The second result holds since all the elements are of dimension $p \times 1$ or $p \times p$. Then, the result follows from the fact that the denominators of $\boldsymbol{g}_{it}^{OPT}$ and $\boldsymbol{g}_{it}^{BOD}$ are $O(t)$.

Next, we derive the asymptotic properties of the IV estimators. Note that IV estimators $\widehat{\boldsymbol{\alpha}}_{IV}^{OPT}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$ can be written as

$$\sqrt{NT} \left(\widehat{\boldsymbol{\alpha}}_{IV} - \boldsymbol{\alpha} \right) = \widehat{\boldsymbol{A}}^{-1} \sqrt{NT} \widehat{\boldsymbol{b}} = \widehat{\boldsymbol{A}}^{-1} \widehat{\boldsymbol{c}} \tag{39}$$

where $\widehat{\boldsymbol{A}}$ denotes $\widehat{\boldsymbol{A}}_{IV}^{OPT}$, and $\widehat{\boldsymbol{A}}_{IV}^{BOD}$, and so on.

The asymptotic behavior of $\widehat{\boldsymbol{A}}$, $\widehat{\boldsymbol{b}}$ and $\widehat{\boldsymbol{c}}$ are given in the following lemma.

Lemma C *Let Assumptions 1, 2, and 3 hold. Then, as both N and T tend to infinity,*

$$(a) \quad \widehat{\boldsymbol{A}}_{IV}^{OPT}, \ \widehat{\boldsymbol{A}}_{IV}^{BOD} \to^p E\left(\boldsymbol{w}_{i,t-1} \boldsymbol{w}_{i,t-1}' \right), \tag{40}$$

$$(b) \quad \widehat{\boldsymbol{b}}_{IV}^{OPT}, \ \widehat{\boldsymbol{b}}_{IV}^{BOD} \to^p 0. \tag{41}$$

If we further assume that Assumption 4 holds, then as both N and T tend to infinity,

$$(c) \quad \widehat{\boldsymbol{c}}_{IV}^{OPT}, \ \widehat{\boldsymbol{c}}_{IV}^{BOD} \to^d N\left[0, \sigma_v^2 E\left(\boldsymbol{w}_{i,t-1} \boldsymbol{w}_{i,t-1}' \right) \right]. \tag{42}$$

Proof of Lemma C To derive the results, we use the following decomposition:

$$\begin{aligned} \boldsymbol{x}_{it}^{*} &= \boldsymbol{\Psi}_t \boldsymbol{w}_{i,t-1} - c_t \widetilde{\boldsymbol{v}}_{itT}, & (43)\\ \boldsymbol{\Psi}_t &= c_t \left(\boldsymbol{I}_p - \frac{1}{T-t} \boldsymbol{\Pi} \boldsymbol{\Phi}_{T-t} \right), & (44)\\ \widetilde{\boldsymbol{v}}_{itT} &= \frac{\left(\boldsymbol{\Phi}_{T-t} v_{it} + \boldsymbol{\Phi}_2 v_{i,T-2} + \cdots + \boldsymbol{\Phi}_1 v_{i,T-1} \right) \boldsymbol{d}_1}{T-t}. & (45) \end{aligned}$$

(a): First, we consider $\widehat{\boldsymbol{A}}_{IV}^{OPT}$. Using Lemma A, B, and the above decomposition, we have

$$\begin{aligned} E\left(\widehat{\boldsymbol{A}}_{IV}^{OPT}\right) &= \frac{1}{T}\sum_{t=1}^{T-1} E\left(\boldsymbol{h}_{it}^{OPT} \boldsymbol{x}_{it}^{*'}\right) \\ &= \frac{1}{T}\sum_{t=1}^{T-1} \left[\boldsymbol{I}_p - O\left(\frac{1}{T-t}\right)\right]\left[E\left(\boldsymbol{w}_{i,t-1}\boldsymbol{w}_{i,t-1}'\right) + O\left(\frac{1}{t}\right)\right] \\ &\to E\left(\boldsymbol{w}_{i,t-1}\boldsymbol{w}_{i,t-1}'\right). \end{aligned}$$

The last convergence comes from $T^{-1}\sum_{t=1}^{T-1} O(1/(T-t)) = O(\log T/T) \to 0$. $var\left(\widehat{\boldsymbol{A}}_{IV}^{OPT}\right)$ are easily shown to tend to zero. For $\widehat{\boldsymbol{\alpha}}_{IV}^{BOD}$, we have

$$\begin{aligned} E\left(\widehat{\boldsymbol{A}}_{IV}^{BOD}\right) &= \frac{1}{T}\sum_{t=1}^{T-1} \left[E(\boldsymbol{w}_{i,t-1}\boldsymbol{w}_{i,t-1}')\left\{\boldsymbol{I}_p + O\left(\frac{1}{T-t}\right)\right\} + O\left(\frac{1}{t}\right)\right] \\ &\to E(\boldsymbol{w}_{i,t-1}\boldsymbol{w}_{i,t-1}'). \end{aligned}$$

(b),(c): First, we consider $\widehat{\boldsymbol{c}}_{IV}^{OPT}$. Since $E\left(\widehat{\boldsymbol{c}}_{IV}^{OPT}\right) = 0$, and $E(\boldsymbol{h}_{it}^{OPT} v_{it}^{*} v_{is}^{*} \boldsymbol{h}_{it}^{OPT'}) = E\left(\boldsymbol{h}_{it}^{OPT} E_t(v_{it}^{*} v_{is}^{*}) \boldsymbol{h}_{it}^{OPT'}\right) = 0$ for $t > s$, where $E_t(\cdot)$ denotes the conditional expectation given η_i and $\{v_{i,t-j}\}_{j=1}^{\infty}$, we have

$$\begin{aligned} var\left(\widehat{\boldsymbol{c}}_{IV}^{OPT}\right) &= \frac{1}{T} var\left(\sum_{t=1}^{T-1} \boldsymbol{h}_{it}^{OPT} v_{it}^{*}\right) = \frac{\sigma_v^2}{T}\sum_{t=1}^{T-1} E\left(\boldsymbol{h}_{it}^{OPT}\boldsymbol{h}_{it}^{OPT'}\right) \\ &= \frac{\sigma_v^2}{T}\sum_{t=1}^{T-1}\left[E\left(\boldsymbol{w}_{i,t-1}\boldsymbol{w}_{i,t-1}'\right) + O\left(\frac{1}{t}\right)\right] \\ &\to \sigma_v^2 E\left(\boldsymbol{w}_{i,t-1}\boldsymbol{w}_{i,t-1}'\right). \end{aligned}$$

Then, using the similar argument as Hahn and Kuersteiner (2002) and Lee (2005), we have

$$\widehat{\boldsymbol{c}}_{IV}^{OPT} \to^{d} N\left[0, \sigma_v^2 E\left(\boldsymbol{w}_{i,t-1}\boldsymbol{w}_{i,t-1}'\right)\right]. \tag{46}$$

The result for $\widehat{\boldsymbol{c}}_{IV}^{BOD}$ is obtained in a similar way.

From (c), it is straightforward to show that $\widehat{\boldsymbol{b}}_{IV}^{BOD}, \widehat{\boldsymbol{b}}_{IV}^{OPT} \to^p 0$.

Proof of Proposition 1 Using Lemma C, the results are easily obtained.

References

[1] Ahn, S.C. and P. Schmidt (1995): "Efficient Estimation of Models for Dynamic Panel Data," *Journal of Econometrics*, 68, 5-27.

[2] Ahn, S.C. and P. Schmidt (1997): "Efficient Estimation of Dynamic Panel Data Models: Alternative Assumptions and Simplified Estimation," *Journal of Econometrics*, 76, 309-321.

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

[4] Alvarez, J. and M. Arellano (2004): "Robust Likelihood Estimation of Dynamic Panel Data Models," mimeo.

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

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

[7] Arellano, M. (2003a): *Panel Data Econometrics*, Oxford University Press, New York.

[8] Arellano, M. (2003b): "Modelling Optimal Instrumental Variables for Dynamic Panel Data Models," Working Paper No. 0310, CEMFI, Madrid.

[9] Arellano, M. and S.R. Bond (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 277-297.

[10] Arellano, M. and O. Bover (1995): "Another Look at the Instrumental Variable Estimation of Error-Component Models," *Journal of Econometrics*, 68, 29-45.

[11] Bun, M.J.G. and J.F. Kiviet (2006): "The Effects of Dynamic Feedbacks on LS and MM Estimator Accuracy in Panel Data Models," *Journal of Econometrics*, 132, 409-444.

[12] Bun, M.J.G. and F. Windmeijer (2007): "The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models," University of Bristol, Discussion Paper No. 07/595.

[13] Blundell, R. and S. Bond (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115-43.

[14] Hahn, J. and G. Kuersteiner (2002): "Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects When Both n and T are Large," *Econometrica*, 70, 1639-1657.

[15] Hahn, J., J. Hausman, and G. Kuersteiner (2007): "Long Difference Instrumental Variables Estimation for Dynamic Panel Models with Fixed Effects," forthcoming in *Journal of Econometrics*.

[16] Hayakawa, K. (2007a): "Small Sample Bias Properties of the System GMM Estimator in Dynamic Panel Data Models," *Economics Letters*, 95, 32-38.

[17] Hayakawa, K. (2007b): "On the Effect of Nonstationary Initial Conditions in Dynamic Panel Data Models," mimeo.

[18] Holtz-Eakin, D., W. Newey, and H. Rosen: (1988) "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 56, 1371-1395.

[19] Lee, Y. (2005): "A General Approach to Bias Correction in Dynamic Panels under Time Series Misspecification," mimeo.

[20] Okui, R. (2006) "The Optimal Choice of Moments in Dynamic Panel Data Models", mimeo.

[21] Phillips, P.C.B. and H. R. Moon: (1999) "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica*, 67, 1057-1111.

[22] So, B.S. and D.W. Shin (1999) "Recursive Mean Adjustment in Time Series Inferences", *Statistics and Probability Letters*, 43, 65–73.

[23] Whittle, P. (1951): *Hypothesis Testing in Time-Series Analysis*, Almqvist and Wiksell, Upsala.

[24] Wise, J. (1955): "The Autocorrelation Function and the Spectral Density Function," *Biometrika*, 42, 151-159.

Table 1: Simulation Results for the AR(1) model

			Median				IQR				MAE			
α_1	T	N	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{BOD}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{BOD}$	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{BOD}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{BOD}$	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{BOD}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{BOD}$
0.3	10	100	0.279	0.296	0.299	0.299	0.067	0.075	0.090	0.075	0.036	0.037	0.045	0.038
0.3	10	500	0.295	0.299	0.300	0.300	0.030	0.033	0.042	0.034	0.015	0.016	0.021	0.017
0.3	15	100	0.281	0.298	0.299	0.300	0.046	0.050	0.070	0.051	0.026	0.025	0.035	0.025
0.3	15	300	0.294	0.299	0.301	0.300	0.027	0.029	0.040	0.029	0.014	0.014	0.020	0.014
0.3	20	100	0.283	0.299	0.301	0.300	0.037	0.041	0.058	0.042	0.022	0.021	0.029	0.021
0.3	20	200	0.291	0.299	0.300	0.300	0.027	0.029	0.041	0.029	0.015	0.014	0.021	0.015
0.3	50	100	0.286	0.300	0.300	0.301	0.020	0.021	0.034	0.021	0.015	0.010	0.017	0.010
0.3	100	100	0.287	0.300	0.300	0.300	0.013	0.014	0.023	0.014	0.013	0.007	0.011	0.007
0.6	10	100	0.556	0.590	0.599	0.599	0.080	0.100	0.129	0.101	0.052	0.050	0.064	0.051
0.6	10	500	0.590	0.598	0.600	0.600	0.038	0.043	0.059	0.045	0.020	0.022	0.030	0.023
0.6	15	100	0.566	0.595	0.598	0.600	0.052	0.059	0.092	0.062	0.037	0.030	0.046	0.031
0.6	15	300	0.588	0.598	0.600	0.600	0.031	0.034	0.054	0.035	0.018	0.017	0.027	0.018
0.6	20	100	0.571	0.598	0.601	0.600	0.039	0.043	0.074	0.046	0.030	0.022	0.037	0.023
0.6	20	200	0.585	0.598	0.600	0.599	0.029	0.031	0.053	0.032	0.018	0.016	0.027	0.016
0.6	50	100	0.580	0.600	0.600	0.600	0.018	0.020	0.039	0.021	0.020	0.010	0.020	0.010
0.6	100	100	0.583	0.600	0.600	0.600	0.012	0.013	0.026	0.013	0.017	0.006	0.013	0.007
0.9	10	100	0.690	0.806	0.897	0.896	0.163	0.237	0.398	0.290	0.210	0.135	0.199	0.145
0.9	10	500	0.825	0.876	0.898	0.898	0.086	0.115	0.172	0.131	0.077	0.060	0.087	0.065
0.9	15	100	0.767	0.862	0.904	0.899	0.087	0.120	0.241	0.143	0.133	0.065	0.121	0.071
0.9	15	300	0.832	0.886	0.900	0.899	0.059	0.073	0.138	0.082	0.068	0.037	0.069	0.041
0.9	20	100	0.803	0.882	0.904	0.900	0.059	0.077	0.170	0.089	0.097	0.039	0.084	0.045
0.9	20	200	0.837	0.890	0.899	0.901	0.045	0.055	0.117	0.063	0.064	0.029	0.059	0.031
0.9	50	100	0.860	0.898	0.900	0.900	0.018	0.022	0.066	0.023	0.040	0.011	0.032	0.012
0.9	100	100	0.875	0.899	0.900	0.900	0.009	0.011	0.037	0.011	0.025	0.005	0.018	0.005

Table 2: Simulation Results for the AR(2) model

				Median				IQR				MAE			
T	N	α_1	α_2	$\widehat{\alpha}^{LEV}_{GMM}$	$\widehat{\alpha}^{BOD}_{GMM}$	$\widehat{\alpha}^{LEV}_{IV}$	$\widehat{\alpha}^{BOD}_{IV}$	$\widehat{\alpha}^{LEV}_{GMM}$	$\widehat{\alpha}^{BOD}_{GMM}$	$\widehat{\alpha}^{LEV}_{IV}$	$\widehat{\alpha}^{BOD}_{IV}$	$\widehat{\alpha}^{LEV}_{GMM}$	$\widehat{\alpha}^{BOD}_{GMM}$	$\widehat{\alpha}^{LEV}_{IV}$	$\widehat{\alpha}^{BOD}_{IV}$
10	100	0.45		0.141	0.259	0.448	0.434	0.193	0.228	0.508	0.414	0.309	0.195	0.254	0.208
10	100		0.45	0.324	0.345	0.444	0.433	0.116	0.125	0.282	0.228	0.126	0.108	0.142	0.116
10	500	0.45		0.274	0.375	0.446	0.449	0.140	0.140	0.244	0.186	0.176	0.090	0.122	0.093
10	500		0.45	0.387	0.409	0.448	0.448	0.067	0.078	0.138	0.099	0.063	0.050	0.069	0.049
15	100	0.45		0.245	0.362	0.457	0.449	0.115	0.137	0.310	0.201	0.205	0.095	0.155	0.101
15	100		0.45	0.372	0.399	0.451	0.447	0.078	0.079	0.183	0.115	0.078	0.056	0.092	0.057
15	300	0.45		0.308	0.408	0.450	0.448	0.093	0.089	0.184	0.110	0.142	0.054	0.092	0.055
15	300		0.45	0.404	0.425	0.448	0.448	0.050	0.051	0.107	0.064	0.047	0.032	0.053	0.032
20	100	0.45		0.297	0.401	0.459	0.450	0.082	0.089	0.215	0.121	0.153	0.056	0.108	0.060
20	100		0.45	0.396	0.419	0.453	0.447	0.057	0.056	0.131	0.074	0.055	0.036	0.066	0.037
20	200	0.45		0.328	0.419	0.452	0.450	0.071	0.068	0.157	0.083	0.122	0.041	0.079	0.041
20	200		0.45	0.413	0.431	0.450	0.449	0.043	0.043	0.096	0.051	0.038	0.026	0.048	0.026
50	100	0.45		0.391	0.444	0.451	0.450	0.032	0.027	0.075	0.029	0.059	0.014	0.037	0.015
50	100		0.45	0.431	0.445	0.449	0.449	0.026	0.023	0.053	0.024	0.020	0.012	0.026	0.012
100	100	0.45		0.420	0.449	0.450	0.450	0.017	0.014	0.038	0.015	0.030	0.007	0.018	0.007
100	100		0.45	0.437	0.449	0.450	0.450	0.015	0.013	0.030	0.014	0.013	0.007	0.015	0.007
10	100	0.6		0.294	0.425	0.601	0.589	0.195	0.218	0.471	0.374	0.306	0.178	0.234	0.187
10	100		0.3	0.236	0.226	0.298	0.290	0.119	0.095	0.202	0.158	0.076	0.077	0.101	0.080
10	500	0.6		0.424	0.537	0.596	0.600	0.144	0.129	0.219	0.166	0.176	0.079	0.109	0.083
10	500		0.3	0.282	0.274	0.299	0.298	0.061	0.056	0.094	0.068	0.033	0.034	0.047	0.034
15	100	0.6		0.397	0.523	0.607	0.600	0.121	0.128	0.280	0.183	0.203	0.087	0.140	0.092
15	100		0.3	0.265	0.264	0.300	0.298	0.086	0.063	0.134	0.082	0.049	0.042	0.067	0.041
15	300	0.6		0.456	0.564	0.599	0.599	0.098	0.083	0.163	0.101	0.144	0.048	0.082	0.050
15	300		0.3	0.292	0.283	0.299	0.299	0.054	0.039	0.078	0.047	0.027	0.023	0.039	0.023
20	100	0.6		0.449	0.558	0.608	0.600	0.087	0.084	0.196	0.110	0.151	0.050	0.098	0.055
20	100		0.3	0.278	0.278	0.302	0.297	0.067	0.047	0.100	0.058	0.037	0.028	0.050	0.029
20	200	0.6		0.476	0.575	0.602	0.600	0.078	0.063	0.141	0.075	0.124	0.037	0.071	0.038
20	200		0.3	0.294	0.287	0.299	0.298	0.050	0.036	0.071	0.040	0.025	0.020	0.035	0.020
50	100	0.6		0.543	0.595	0.601	0.600	0.035	0.026	0.068	0.028	0.057	0.014	0.033	0.014
50	100		0.3	0.291	0.296	0.299	0.299	0.031	0.022	0.043	0.023	0.016	0.011	0.022	0.011
100	100	0.6		0.572	0.599	0.600	0.600	0.018	0.015	0.035	0.015	0.028	0.007	0.017	0.008
100	100		0.3	0.291	0.299	0.299	0.299	0.017	0.014	0.026	0.014	0.011	0.007	0.013	0.007