A Simple Efficient Instrumental Variable Estimator for Panel AR(p) Models When Both N and T are Large

Kazuhiko Hayakawa*†
Department of Economics, Hitotsubashi University
JSPS Research Fellow

First Draft: May 2007
This version: February 9, 2008

Abstract

In this paper, we show that for panel AR(p) models, an instrumental variable (IV) estimator with instruments deviated from past means 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, instruments deviated from past means.

JEL classification: C13, C23.

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

†The author is deeply grateful to two anonymous referees, Kaddour Hadri, Cheng Hsiao, Naoto Kunitomo, Eiji Kurozumi, Kosuke Oya, Donggyu Sul, Taku Yamamoto, and the participants of the 14th International Conference of Panel Data at Xiamen University, the Fall meeting of Japanese Economic Association at Nihon University and Hitotsubashi Conference on Econometrics 2007 for helpful comments. I also acknowledge Ryo Okui who posed a question that inspired this paper. This research benefited from the JSPS fellowship. All the remaining errors are mine.

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, a simple one-step IV estimator using instruments deviated from past means 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. Compared to the existing estimators, there are two advantages in the proposed IV estimator. The first is that although the WG and GMM estimators are consistent only when T and N is large, respectively, the proposed IV estimator is consistent under large N and fixed T, fixed N and large T, or large N and large T asymptotics. This implies that the proposed IV estimator can be used for large N and small T, small N and large T, or large N and large T panel data. The second advantage is that the proposed IV estimator is more efficient than Anderson and Hsiao's (1981) IV estimator, and as efficient as the WG and GMM estimators when both N and T are large.

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 $\mathbf{x}$ and a matrix $\mathbf{A}$, we define $\|\mathbf{x}\|^2 = \mathbf{x}'\mathbf{x}$ and $\|\mathbf{A}\|^2 = tr(\mathbf{A}'\mathbf{A})$ 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, respectively.

2 Setup and Result

2.1 The model and assumptions

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

$$\begin{aligned} y_{it} &= \alpha_1 y_{i,t-1} + \alpha_2 y_{i,t-2} + \cdots + \alpha_p y_{i,t-p} + \eta_i + v_{it} \\ &= \boldsymbol{\alpha}' \mathbf{x}_{it} + \eta_i + v_{it} \qquad (i = 1, ..., N, \quad t = 1, ..., T) \end{aligned} \tag{1}$$

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

(1) can be written in a companion form as

$$\mathbf{x}_{i,t+1} = \boldsymbol{\Pi}\mathbf{x}_{it} + \mathbf{d}_1(\eta_i + v_{it}) \tag{2}$$

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

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

where $\mathbf{I}_k$ is an identity matrix of order k and $\mathbf{O}_{k\times\ell}$ 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 $\mathbf{x}_{i1}$, 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 $var(\eta_i) = \sigma_\eta^2$.*

Assumption 2. *The initial observations satisfy*

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

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

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

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

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

[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, $\mathbf{x}_{it}$ can be written as

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

where

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

To remove the individual effects, η_i, we use the forward orthogonal deviation (FOD) since the errors transformed by the FOD are serially uncorrelated and homoskedastic if the original errors are.[5] Specifically, the model to be estimated is given by

$$y_{it}^* = \boldsymbol{\alpha}'\mathbf{x}_{it}^* + v_{it}^* \qquad (i = 1, ..., N, \quad t = 1, ..., T-1) \quad (3)$$

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

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

where $\mathbf{h}_{it} = E(\mathbf{x}_{it}^*|\mathbf{y}_i^{t-1})$ and $\mathbf{y}_i^{t-1} = (y_{i,t-1}, ..., y_{i,1-p})'$. $\widehat{\boldsymbol{\alpha}}_{IV}^{OPT}$ is an ideal estimator since it is consistent and asymptotically efficient when N is large and T is fixed.

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

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

[5]Note that taking a first difference induces a serial correlation in the errors, and this correlation changes the form of the optimal instruments defined in the next section.

However, the drawback of this estimator is that it is infeasible since the optimal instruments $\mathbf{h}_{it}$ is unknown. A standard approach to obtain a feasible optimal IV estimator is to use a sample linear projection of $\mathbf{h}_{it}$, which is given by

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

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

$$\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV} = \left(\frac{1}{NT} \sum_{t=1}^{T-1} \mathbf{X}_t^{*\prime} \mathbf{M}_t^{LEV} \mathbf{X}_t^* \right)^{-1} \left(\frac{1}{NT} \sum_{t=1}^{T-1} \mathbf{X}_t^{*\prime} \mathbf{M}_t^{LEV} \mathbf{y}_t^* \right) \tag{4}$$

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

However, one problem of $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ is that if N and T increase at the same rate, the estimate of $\mathbf{h}_{it}$ 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 $O(1/N)$.[6]

Thus, in this paper, we propose an alternative approach. Instead of estimating the optimal instruments, we propose to use an *observable* variable that has the same structure as the optimal instruments, $\mathbf{h}_{it}$. Hence, we need to investigate the structure of $\mathbf{h}_{it}$. Arellano (2003b) shows that, under the assumption that $E(\mu_i | \mathbf{y}_i^{t-1})$ coincides with the linear projection, the infeasible optimal instruments can be rewritten in the following form:

$$\begin{aligned} E(\mathbf{x}_{it}^* | \mathbf{y}_i^{t-1}) &= c_t \left[\mathbf{I}_p - \frac{\boldsymbol{\Pi}(\mathbf{I}_p - \boldsymbol{\Pi}^{T-t})(\mathbf{I}_p - \boldsymbol{\Pi})^{-1}}{T-t} \right] \left[\mathbf{x}_{it} - \boldsymbol{\iota}_p E(\mu_i | \mathbf{y}_i^{t-1}) \right] \\ &= c_t \left[\mathbf{I}_p - \frac{\boldsymbol{\Pi}(\mathbf{I}_p - \boldsymbol{\Pi}^{T-t})(\mathbf{I}_p - \boldsymbol{\Pi})^{-1}}{T-t} \right] \left[\mathbf{w}_{i,t-1} + \boldsymbol{\iota}_p \left\{ \mu_i - E(\mu_i | \mathbf{y}_i^{t-1}) \right\} \right] \qquad (5) \\ &= c_t \left[\mathbf{I}_p - O\left(\frac{1}{T-t} \right) \right] \left[\mathbf{w}_{i,t-1} + O_p\left(\frac{1}{\sqrt{t}} \right) \right]. \qquad (6) \end{aligned}$$

where the second equality comes from the fact that $\mathbf{x}_{it} = \boldsymbol{\iota}_p \mu_i + \mathbf{w}_{i,t-1}$, $\mu_i = \eta_i / (1 - \boldsymbol{\alpha}' \boldsymbol{\iota}_p)$, and the third equality is proved in Lemma A (see Appendix).

From (5) and (6), we find that (i) the individual effect μ_i is demeaned in (5), (ii) when t is large, $\mathbf{w}_{i,t-1}$ is the dominating term in (6).

Our next task is to find an observable variable that has the same structure as (6).

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

Instruments deviated from past means

We consider instruments $\mathbf{z}_{it(\ell)} = (z_{it(\ell)}^{t-1}, z_{it(\ell)}^{t-2}, ..., z_{it(\ell)}^{t-p})'$ as follows:

$$\begin{aligned} z_{it(\ell)}^{t-1} &= c_t \left[y_{i,t-1} - \frac{y_{i,t-1-\ell} + \cdots + y_{i,-p+1}}{t+p-\ell-1} \right] \\ z_{it(\ell)}^{t-2} &= c_t \left[y_{i,t-2} - \frac{y_{i,t-2-\ell} + \cdots + y_{i,-p+1}}{t+p-\ell-2} \right] \\ &\vdots \\ z_{it(\ell)}^{t-p} &= c_t \left[y_{i,t-p} - \frac{y_{i,t-p-\ell} + \cdots + y_{i,-p+1}}{t-\ell} \right] \end{aligned}$$

where $\ell \geq 0$ is fixed.[7] Since $\mathbf{z}_{it(\ell)}$ is deviated from past means, it can be seen as a modification of the recursive mean adjustment (RMA) method by So and Shin (1999).[8]

Now, we show that $\mathbf{z}_{it(\ell)}$ meets the above two requirements. For the first requirement, it is straightforward to show that the individual effects are demeaned since $\mathbf{z}_{it(\ell)}$ is deviated from past means. For the second requirement, we show in Appendix that $\mathbf{z}_{it(\ell)}$ can be written as

$$\mathbf{z}_{it(\ell)} = c_t \left[\mathbf{w}_{i,t-1} + O_p \left(\frac{1}{\sqrt{t}} \right) \right] \tag{7}$$

Thus, comparing (6) and (7), we find that unobservable $\mathbf{h}_{it}$ and observable $\mathbf{z}_{it(\ell)}$ have the same structure, i.e., (i) demeaning individual effects, (ii) $\mathbf{w}_{i,t-1}$ is dominating.

The IV estimator using $\mathbf{z}_{it(\ell)}$ as instruments is given by

$$\begin{aligned} \widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell} &= \left(\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=t_0}^{T-1} \mathbf{z}_{it(\ell)} \mathbf{x}_{it}^{*\prime} \right)^{-1} \left(\frac{1}{NT} \sum_{i=1}^{N} \sum_{t=t_0}^{T-1} \mathbf{z}_{it(\ell)} y_{it}^{*} \right) \\ &= \boldsymbol{\alpha} + \left(\widehat{\mathbf{A}}_{IV}^{RMA\ell} \right)^{-1} \widehat{\mathbf{b}}_{IV}^{RMA\ell} \end{aligned} \tag{8}$$

where $t_0 = 2$ for $\ell = 0, 1$ and $t_0 = \ell + 1$ for $\ell \geq 2$.

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

Proposition 1. *Let Assumptions 1, 2, and 3 hold. Then, for fixed $\ell \geq 0$, as both N and T tend to infinity, the infeasible optimal IV estimator $\widehat{\boldsymbol{\alpha}}_{IV}^{OPT}$ and the feasible IV estimator $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ are consistent. If we further assume that Assumption 4 holds, then, as both N and T tend to infinity, we have*

$$\sqrt{NT} \left(\widehat{\boldsymbol{\alpha}}_{IV} - \boldsymbol{\alpha} \right) \xrightarrow{d} \mathcal{N} \left(0, \sigma_v^2 \left[E(\mathbf{w}_{i,t-1} \mathbf{w}_{i,t-1}') \right]^{-1} \right) \tag{9}$$

[7] For the choice of ℓ, we consider $\ell = 0, ..., 4$ in simulation studies (see Section 3).

[8] The case $\ell = 1$ corresponds to the original RMA method.

where $\widehat{\boldsymbol{\alpha}}_{IV}$ *denotes* $\widehat{\boldsymbol{\alpha}}_{IV}^{OPT}$ *and* $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$.

Note that the asymptotic variance $\sigma_v^2 \left[E(\mathbf{w}_{i,t-1}\mathbf{w}'_{i,t-1})\right]^{-1}$ is of the same form as the within groups (WG) estimator derived by Lee (2005).

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:

$$\sqrt{NT}\left[\widehat{\alpha}_{GMM}^{LEV} - \left(\alpha_1 - \frac{1}{N}(1+\alpha_1)\right)\right] \xrightarrow{d} \mathcal{N}\left(0, 1-\alpha_1^2\right), \tag{10}$$

$$\sqrt{NT}\left[\widehat{\alpha}_{WG} - \left(\alpha_1 - \frac{1}{T}(1+\alpha_1)\right)\right] \xrightarrow{d} \mathcal{N}\left(0, 1-\alpha_1^2\right). \tag{11}$$

Also, from Proposition 1, we have

$$\sqrt{NT}\left[\widehat{\alpha}_{IV}^{RMA\ell} - \alpha_1\right] \xrightarrow{d} \mathcal{N}\left(0, 1-\alpha_1^2\right). \tag{12}$$

Comparing (10), (11) and (12), 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 the distribution of $\widehat{\alpha}_{IV}^{RMA\ell} - \alpha_1$ is centered at zero. This is because $\widehat{\alpha}_{GMM}^{LEV}$ and $\widehat{\alpha}_{WG}$ suffer from the "many instruments problem" and "incidental parameter problem", respectively, while $\widehat{\alpha}_{IV}^{RMA\ell}$ does not.[9]

Remark 2. Hahn and Kuersteiner (2002) show that if we further assume normality on v_{it}, then $\sigma_v^2 \left[E(\mathbf{w}_{i,t-1}\mathbf{w}'_{i,t-1})\right]^{-1}$ is equal to the lower bound under large N and T asymptotics.[10] Hence, $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ is an efficient IV estimator under large N and T asymptotics without an asymptotic bias when v_{it} is normally distributed.

Remark 3. Another feature of $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ is that since the individual effects are completely eliminated from both the model and instruments under stationary initial conditions, the performance of $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ is not affected by the variance ratio of the individual effects to the disturbances although the typical GMM estimators using instruments in levels are.[11]

Remark 4. Although we use large N and T asymptotics in deriving the properties, consistency and asymptotic normality are also obtained under large N and fixed T, or fixed N and large T asymptotics. Especially, under fixed N and large T asymptotics, the same expression as (9) is obtained. This is in marked contrast to the GMM estimator where large N is required. Furthermore, although the GMM estimator can be computed only when $T - 1 \leq N$, the proposed IV estimator can be computed for any N and T.

[9] This interpretation was suggested by a referee.

[10] Note that the AR(p) model (1) can be written as the VAR(1) model (2).

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

3 Monte Carlo Simulation

In this section, we compare $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ 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 $\alpha_1 = 0.5, 0.9$, and for the AR(2) model, we set $(\alpha_1, \alpha_2) = (0.6, -0.1), (0.6, 0.3)$. 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 interquartile range (IQR), and the median absolute error (MAE). The number of replications is 5000 for all cases.

We consider the GMM and IV estimators using instruments in levels or deviated from past means. The GMM estimator using $\mathbf{y}_i^{t-1}$ as instruments is defined as (4). The GMM estimator using $\mathbf{z}_{it(1)}$ as instruments is defined by[12]

$$\widehat{\boldsymbol{\alpha}}_{GMM}^{RMA1} = \left(\frac{1}{NT} \sum_{t=2}^{T-1} \mathbf{X}_t^{*\prime} \mathbf{M}_t^{RMA1} \mathbf{X}_t^* \right)^{-1} \left(\frac{1}{NT} \sum_{t=2}^{T-1} \mathbf{X}_t^{*\prime} \mathbf{M}_t^{RMA1} \mathbf{y}_t^* \right)$$

where $\mathbf{M}_t^{RMA1} = \mathbf{Z}_t^{RMA1} \left(\mathbf{Z}_t^{RMA1\prime} \mathbf{Z}_t^{RMA1} \right)^{-1} \mathbf{Z}_t^{RMA1\prime}$, and $\mathbf{Z}_t^{RMA1} = (\mathbf{z}_{1t(1)}, ..., \mathbf{z}_{Nt(1)})'$. $\widehat{\boldsymbol{\alpha}}_{GMM}^{RMA1}$ does not share the problem with $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ that the number of instruments is too large. In fact, the number of instruments used in $\widehat{\boldsymbol{\alpha}}_{GMM}^{RMA1}$ is $O(T)$, while that in $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ is $O(T^2)$.

For the proposed IV estimators, we consider $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA0}, ..., \widehat{\boldsymbol{\alpha}}_{IV}^{RMA4}$ as defined by (8). Also, for the purpose of comparison, we consider an IV estimator using $\mathbf{x}_{it}$ as instruments as follows:

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

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.

For the choice of ℓ, we find that, in terms of MAE, $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA1}$ performs best in many cases. To describe the intuition behind this result, we consider the AR(1)

[12] The reason why we consider the GMM estimator using $\mathbf{z}_{it(1)}$ as instruments is that, in terms of MAE, the GMM estimator may perform better than the IV estimator since the GMM estimator is more efficient than the IV estimator under large N and fixed T asymptotics. Also, the reason why we choose $\ell = 1$ is that the IV estimator with $\ell = 1$ performs best as will be shown.

model and $\ell = 0, 2$. In this case, the instruments are

$$\begin{aligned} z_{it(0)} &= c_t \left[y_{i,t-1} - \frac{y_{i,t-1} + \cdots + y_{i,0}}{t} \right], \\ z_{it(2)} &= c_t \left[y_{i,t-1} - \frac{y_{i,t-3} + \cdots + y_{i,0}}{t-2} \right]. \end{aligned}$$

For the case of $\ell = 2$, we find that $y_{i,t-2}$ is not used and this causes an efficiency loss. The same result applies to the case $\ell \geq 2$. For the case of $\ell = 0$, although $z_{it(0)}$ uses all information, $y_{i,t-1}$ induces an additional correlation and make the second term larger although its order is $O(1/\sqrt{t})$.

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 $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$, have little bias for all cases, while the GMM estimators have non-negligible bias when $\alpha = 0.9$. 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 $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$, $\widehat{\boldsymbol{\alpha}}_{GMM}^{RMA1}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ become quite small when T is as large as 50. This result is consistent with Proposition 1 where $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ are shown to have the same asymptotic variance when N and T are large. Also, asymptotic variances in (10) and (11) suggest that for given N and T, the dispersion of $\widehat{\boldsymbol{\alpha}}_{GMM}^{RMA1}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ becomes smaller as α_1 grows. However, the simulation results do not show such a tendency when $T \leq 50$. This may be due to the weak identification problem as α_1 approaches unity.[13] It is of interest how much efficiency of the proposed IV estimator is lost compared to the infeasible optimal IV estimator. Looking at the table, we find that the infeasible optimal IV estimator is slightly less efficient than $\widehat{\alpha}_{GMM}^{LEV}$, which is a feasible optimal IV estimator. Although the proposed IV estimators are less efficient than the infeasible optimal IV estimators, the difference becomes negligible as T gets larger. For the median absolute error, we find that $\widehat{\boldsymbol{\alpha}}_{GMM}^{RMA1}$ has the smallest MAE in many cases. However, the difference in the MAE between $\widehat{\boldsymbol{\alpha}}_{GMM}^{RMA1}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ 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, $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$ has the smallest dispersion in all cases. Also, we find that the difference in the IQR between $\widehat{\boldsymbol{\alpha}}_{GMM}^{LEV}$, $\widehat{\boldsymbol{\alpha}}_{GMM}^{RMA1}$, and $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ becomes small when T is large. For $\widehat{\boldsymbol{\alpha}}_{IV}^{OPT}$, we find that it performs well for the case $(\alpha_1, \alpha_2) = (0.6, -0.1)$. However, for the case $(\alpha_1, \alpha_2) = (0.6, 0.3)$, $\widehat{\boldsymbol{\alpha}}_{IV}^{OPT}$ does not work well and the reason is unclear. Therefore, it is difficult to investigate how much efficiency is lost in the proposed IV estimator. In terms of the MAE, although $\widehat{\boldsymbol{\alpha}}_{GMM}^{RMA1}$ performs best in many cases, the difference

[13] Note that similar results are also reported in Alvarez and Arellano (2003).

between $\widehat{\boldsymbol{\alpha}}_{GMM}^{RMA1}$ and $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$ is quite small.

The simulation results suggest that, in terms of the bias and MAE, the IV and GMM estimators using the proposed instruments perform better than the commonly used GMM estimator, $\widehat{\boldsymbol{\alpha}}$, even when T is as large as 10.

4 Conclusion

In this paper, we showed that the infeasible optimal IV estimator and the IV estimator using instruments deviated from past means 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 deviated from past means. But these tasks are left for future research.

Appendix

Lemma A *Let Assumptions 1, 2, and 3 hold. Then,* $\mathbf{h}_{it}$ *and* $\mathbf{z}_{it(\ell)}$ *can be written as*

$$\begin{aligned}(a)\quad \mathbf{h}_{it} = E(\mathbf{x}_{it}^{*}|\mathbf{y}_i^{t-1}) &= c_t\left[\mathbf{I}_p - \frac{\boldsymbol{\Pi}(\mathbf{I}_p-\boldsymbol{\Pi}^{T-t})(\mathbf{I}_p-\boldsymbol{\Pi})^{-1}}{T-t}\right][\mathbf{w}_{i,t-1}+\boldsymbol{g}_{it}] \\ &= c_t\left[\mathbf{I}_t - O\left(\frac{1}{T-t}\right)\right]\left[\mathbf{w}_{i,t-1}+O_p\left(\frac{1}{\sqrt{t}}\right)\right]\end{aligned}$$

$$(b)\quad \mathbf{z}_{it(\ell)} = c_t\left[\mathbf{w}_{i,t-1}+\widetilde{\boldsymbol{g}}_{it}\right] \tag{13}$$

$$= c_t\left[\mathbf{w}_{i,t-1}+O_p\left(\frac{1}{\sqrt{t}}\right)\right] \tag{14}$$

where

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

$$\widetilde{\boldsymbol{g}}_{it} = \mu_i \boldsymbol{G}_p \boldsymbol{\iota}_p - (\mathbf{I}_p - \boldsymbol{G}_p) \left\{ \frac{(\boldsymbol{\Phi}_1 v_{i,t-1-\ell} + \cdots + \boldsymbol{\Phi}_{t-1-\ell} v_{i1})\, \mathbf{d}_1 + \boldsymbol{\Phi}_{t-\ell} \mathbf{w}_{i0} + \mathbf{q}_i}{t - \ell} \right\}, \quad (15)$$

$$\begin{aligned}
\boldsymbol{\Phi}_j &= \boldsymbol{\Pi}^0 + \boldsymbol{\Pi}^1 + \cdots + \boldsymbol{\Pi}^{j-1} = (\mathbf{I}_p - \boldsymbol{\Pi})^{-1} (\mathbf{I}_p - \boldsymbol{\Pi}^j) \\
\mathbf{q}_i &= \begin{bmatrix} y_{i,-1} + y_{i,-2} + \cdots + y_{i,-p+1} \\ y_{i,-2} + \cdots + y_{i,-p+1} \\ \vdots \\ y_{i,-p+1} \end{bmatrix} \\
\mathbf{G}_p &= diag \left[\frac{p-1}{t-\ell+p-1}, \frac{p-2}{t-\ell+p-2}, \cdots, 1 \right]
\end{aligned}$$

and $\boldsymbol{\kappa}_p$, $\mathbf{R}$, *and* $\mathbf{r}_i$ *are defined later.*

Proof of Lemma A. (a) First, note that under the assumption that $E(\mu_i | \mathbf{y}_i^{t-1})$ coincides with the linear projection, we have

$$E(\mathbf{x}_{it}^* | \mathbf{y}_i^{t-1}) = c_t \left[\mathbf{I}_p - \frac{\boldsymbol{\Pi}(\mathbf{I}_p - \boldsymbol{\Pi}^{T-t})(\mathbf{I}_p - \boldsymbol{\Pi})^{-1}}{T-t} \right] \left[\mathbf{x}_{it} - \boldsymbol{\iota}_p \frac{\phi \left(\boldsymbol{\iota}_{t+p-1}' \mathbf{V}_{t-1}^{-1} \mathbf{y}_i^{t-1}\right)}{1 + \phi(\boldsymbol{\iota}_{t+p-1}' \mathbf{V}_{t-1}^{-1} \boldsymbol{\iota}_{t+p-1})} \right]$$

where $\phi = \sigma_\mu^2 / \sigma_v^2$, $\mathbf{V}_{t-1} = \sigma_v^{-2} E\left[(\mathbf{y}_i^{t-1} - \mu_i \boldsymbol{\iota}_{t+p-1})(\mathbf{y}_i^{t-1} - \mu_i \boldsymbol{\iota}_{t+p-1})'\right]$, $\mu_i = \eta_i / (1 - \boldsymbol{\alpha}'\boldsymbol{\iota}_p)$, and $\sigma_\mu^2 = var(\mu_i)$. In the following, we derive formulas of $\boldsymbol{\iota}_{t+p-1}' \mathbf{V}_{t-1}^{-1} \mathbf{y}_i^{t-1}$ and $\boldsymbol{\iota}_{t+p-1}' \mathbf{V}_{t-1}^{-1} \boldsymbol{\iota}_{t+p-1}$.

Following Whittle (1951) and Wise (1955), let us define the $(t+p-1) \times (t+p-1)$ matrix $\mathbf{U}$ as follows.

$$\mathbf{U} = \begin{bmatrix} \mathbf{O}_{(t+p-2)\times 1} & \mathbf{I}_{t+p-2} \\ \mathbf{O}_{1\times 1} & \mathbf{O}_{1\times(t+p-2)} \end{bmatrix}.$$

Then, we have

$$\mathbf{U}^2 = \begin{bmatrix} \mathbf{O}_{(t+p-3)\times 2} & \mathbf{I}_{t+p-3} \\ \mathbf{O}_{2\times 2} & \mathbf{O}_{2\times(t+p-3)} \end{bmatrix}, \quad \mathbf{U}^3 = \begin{bmatrix} \mathbf{O}_{(t+p-4)\times 3} & \mathbf{I}_{t+p-4} \\ \mathbf{O}_{3\times 3} & \mathbf{O}_{3\times(t+p-4)} \end{bmatrix}, \quad \cdots$$

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

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

$$\mathbf{y}_i^{t-1} = \alpha_1 \mathbf{U} \mathbf{y}_i^{t-1} + \alpha_2 \mathbf{U}^2 \mathbf{y}_i^{t-1} + \cdots + \alpha_p \mathbf{U}^p \mathbf{y}_i^{t-1}$$

$$+\eta_i \begin{bmatrix} \boldsymbol{\iota}_{t-1} \\ \mathbf{O}_{p\times 1} \end{bmatrix} + \begin{bmatrix} \mathbf{v}_i^{t-1} \\ \mathbf{O}_{p\times 1} \end{bmatrix} + \begin{bmatrix} \mathbf{O}_{(t-1)\times 1} \\ \mathbf{r}_i \end{bmatrix}$$

where $\mathbf{v}_i^{t-1} = (v_{i,t-1}, ..., v_{i,1})'$, and

$$\mathbf{r}_i = \begin{bmatrix} y_{i0} - \alpha_1 y_{i,-1} - \alpha_2 y_{i,-2} - \cdots - \alpha_{p-1} y_{i,-p+1} \\ \vdots \\ y_{i,-p+3} - \alpha_1 y_{i,-p+2} - \alpha_2 y_{i,-p+1} \\ y_{i,-p+2} - \alpha_1 y_{i,-p+1} \\ y_{i,-p+1} \end{bmatrix}$$

. Since y_{it} is stationary and its conditional mean given η_i is $\mu_i = \eta_i/(1-\boldsymbol{\alpha}'\boldsymbol{\iota_p})$,

$$\begin{aligned} (\mathbf{I}_{t+p-1} - \Delta)\, \widetilde{\mathbf{y}}_i^{t-1} &= \eta_i \begin{bmatrix} \boldsymbol{\iota}_{t-1} \\ \mathbf{O}_{p\times 1} \end{bmatrix} - \mu_i (\mathbf{I}_{t+p-1} - \Delta)\boldsymbol{\iota}_{t+p-1} + \begin{bmatrix} \mathbf{v}_i^{t-1} \\ \mathbf{O}_{p\times 1} \end{bmatrix} + \begin{bmatrix} \mathbf{O}_{p\times 1} \\ \mathbf{r}_i \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{v}_i^{t-1} \\ \mathbf{O}_{p\times 1} \end{bmatrix} + \begin{bmatrix} \mathbf{O}_{(t-1)\times 1} \\ \mathbf{r}_i - \boldsymbol{\kappa}_p \mu_i \end{bmatrix} \\ &= \widetilde{\mathbf{v}}_i^{t-1} + \widetilde{\mathbf{r}}_i \end{aligned}$$

where $\widetilde{\mathbf{y}}_i^{t-1} = \mathbf{y}_i^{t-1} - \mu_i \boldsymbol{\iota}_{t+p-1}$, $\Delta = \left(\alpha_1 \mathbf{U} + \alpha_2 \mathbf{U}^2 + \cdots + \alpha_p \mathbf{U}^p\right)$, and

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

Then, it follows that

$$\mathbf{V}_{t-1}^{-1} = \left(\mathbf{I}_{t+p-1} - \Delta'\right) \begin{bmatrix} \mathbf{I}_{t-1} & O \\ O & \mathbf{R}^{-1} \end{bmatrix} (\mathbf{I}_{t+p-1} - \Delta)$$

where $\mathbf{R} = \sigma_v^{-2} E\left[(\mathbf{r}_i - \boldsymbol{\kappa}_p \mu_i)(\mathbf{r}_i - \boldsymbol{\kappa}_p \mu_i)'\right]$.

Therefore, we have

$$\begin{aligned} \boldsymbol{\iota}_{t+p-1}' \mathbf{V}_{t-1}^{-1} \boldsymbol{\iota}_{t+p-1} &= (1-\boldsymbol{\alpha}'\boldsymbol{\iota}_p)^2 (t-1) + \boldsymbol{\kappa}_p' \mathbf{R}^{-1} \boldsymbol{\kappa}_p, \\ \boldsymbol{\iota}_{t+p-1}' \mathbf{V}_{t-1}^{-1} \mathbf{y}_i^{t-1} &= (1-\boldsymbol{\alpha}'\boldsymbol{\iota}_p)\left[\eta_i (t-1) + v_{i,t-1} + \cdots + v_{i,1}\right] + \boldsymbol{\kappa}_p' \mathbf{R}^{-1} \mathbf{r}_i \end{aligned}$$

(b) Since $\mathbf{x}_{i,t-\ell} = (y_{i,t-1-\ell}, ..., y_{i,t-p-\ell})'$, we have

$$\mathbf{z}_{it(\ell)} = \begin{bmatrix} z_{it(\ell)}^{t-1} \\ z_{it(\ell)}^{t-2} \\ \vdots \\ z_{it(\ell)}^{t-p} \end{bmatrix} = c_t \left\{ \begin{bmatrix} y_{i,t-1} \\ y_{i,t-2} \\ \vdots \\ y_{i,t-p} \end{bmatrix} - \begin{bmatrix} \frac{y_{i,t-\ell-1} + \cdots + y_{i,-p+1}}{t-\ell+p-1} \\ \frac{y_{i,t-\ell-2} + \cdots + y_{i,-p+1}}{t-\ell+p-2} \\ \vdots \\ \frac{y_{i,t-\ell-p} + \cdots + y_{i,-p+1}}{t-\ell} \end{bmatrix} \right\}$$

$$
= c_t \left\{ \begin{bmatrix} y_{i,t-1} \\ y_{i,t-2} \\ \vdots \\ y_{i,t-p} \end{bmatrix} - (\mathbf{I} - \mathbf{G}_p) \begin{bmatrix} \frac{(y_{i,t-\ell-1}+\cdots+y_{i0})+y_{i,-1}+\cdots+y_{i,-p+1}}{t-\ell} \\ \frac{(y_{i,t-\ell-2}+\cdots+y_{i,-1})+y_{i,-2}+\cdots+y_{i,-p+1}}{t-\ell} \\ \vdots \\ \frac{(y_{i,t-\ell-p}+\cdots+y_{i,-p+1})}{t-\ell} \end{bmatrix} \right\}
$$

$$
= c_t \left\{ \mathbf{x}_{it} - (\mathbf{I} - \mathbf{G}_p) \frac{\mathbf{x}_{i,t-\ell} + \cdots + \mathbf{x}_{i1} + \mathbf{q}_i}{t-\ell} \right\}
$$

Since $\mathbf{x}_{it} = \mu_i \boldsymbol{\iota}_p + \mathbf{w}_{i,t-1}$, after some algebra, we get

$$
\mathbf{x}_{i,t-\ell} + \cdots + \mathbf{x}_{i1} = (t-\ell)\mu_i \boldsymbol{\iota}_p + (\boldsymbol{\Phi}_1 v_{i,t-1-\ell} + \cdots + \boldsymbol{\Phi}_{t-1-\ell} v_{i1})\, \mathbf{d}_1 + \boldsymbol{\Phi}_{t-\ell} \mathbf{w}_{i0}
$$

Thus, we get (13). To prove (14), we have to show that $\widetilde{\mathbf{g}}_{it}$ is $O_p(1/\sqrt{t})$. However, since $\|\mathbf{I}_p - \mathbf{G}_p\| = O(1)$, $v_{i,t-\ell-1}, ..., v_{i1}$ are independent random variables, and p is fixed, the second term in (15) is $O_p(1/\sqrt{t})$. For the first term, since p is fixed, $\|E(\mu_i \mathbf{G}_p \boldsymbol{\iota_p})\|^2 = \sigma_\mu^2 \sum_{j=1}^p [(p-j)/(t-\ell+p-j)]^2 = O(1/t)$, the result follows.

□

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

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

$$
\begin{aligned}
&\left\| E\left[(v_{i,t-1} + \cdots + v_{i,p}) \mathbf{w}'_{i,t-1} \right] \right\| = \sigma_v^2 \left\| \mathbf{d}'_1 \left[(\mathbf{I}_p - \boldsymbol{\Pi})^{-1} (\mathbf{I}_p - \boldsymbol{\Pi}^{t-p}) \right]' \right\| = O(1), \\
&\left\| E\left[\boldsymbol{\kappa}'_p \mathbf{R}^{-1} \mathbf{r}_i \mathbf{w}'_{i,t-1} \right] \right\| = O(1), \\
&\left\| E\left[(\boldsymbol{\Phi}_1 v_{i,t-1-\ell} + \cdots + \boldsymbol{\Phi}_{t-1-\ell} v_{i1})\, \mathbf{d}_1 \mathbf{w}'_{i,t-1} \right] \right\| = \sigma_v^2 \left\| \sum_{j=1}^{t-2} \boldsymbol{\Phi}_j \mathbf{d}_1 \mathbf{d}'_1 (\boldsymbol{\Pi}')^j \right\| = O(1) \\
&\left\| E\left(\boldsymbol{\Phi}_{t-\ell} \mathbf{w}_{i,0} + \mathbf{q}_i \right) \mathbf{w}'_{i,t-1} \right\| = O(1).
\end{aligned}
$$

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 $\boldsymbol{g}_{it}$ and $\widetilde{\boldsymbol{g}}_{it}$ are $O(t)$.

□

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

$$
\sqrt{NT}\,(\widehat{\boldsymbol{\alpha}}_{IV} - \boldsymbol{\alpha}) = \widehat{\mathbf{A}}^{-1} \sqrt{NT} \widehat{\mathbf{b}}
$$

where $\widehat{\mathbf{A}}$ denotes $\widehat{\mathbf{A}}_{IV}^{OPT}$, $\widehat{\mathbf{A}}_{IV}^{RMA\ell}$, and $\widehat{\mathbf{b}}$ denotes $\widehat{\mathbf{b}}_{IV}^{OPT}$, and $\widehat{\mathbf{b}}_{IV}^{RMA\ell}$

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

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

$$\begin{array}{ll} (a) & \widehat{\mathbf{A}}_{IV}^{OPT},\ \widehat{\mathbf{A}}_{IV}^{RMA\ell} \xrightarrow{p} E\left(\mathbf{w}_{i,t-1}\mathbf{w}_{i,t-1}'\right), \\ (b) & \widehat{\mathbf{b}}_{IV}^{OPT},\ \widehat{\mathbf{b}}_{IV}^{RMA\ell} \xrightarrow{p} 0. \end{array}$$

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

$$(c) \quad \sqrt{NT}\widehat{\mathbf{b}}_{IV}^{OPT},\ \sqrt{NT}\widehat{\mathbf{b}}_{IV}^{RMA\ell} \xrightarrow{d} \mathcal{N}\left[0, \sigma_v^2 E\left(\mathbf{w}_{i,t-1}\mathbf{w}_{i,t-1}'\right)\right].$$

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

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

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

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

The last convergence comes from $T^{-1}\sum_{t=1}^{T-1} O(1/(T-t)) = O(\log T/T) \to 0$. $var\left(\widehat{\mathbf{A}}_{IV}^{OPT}\right)$ is shown to tend to zero as follows:

$$\left\| var\left(vec\left\{\widehat{\mathbf{A}}_{IV}^{OPT}\right\}\right)\right\| = \frac{1}{NT^2}\left\| var\left(\sum_{t=1}^{T-1} vec\left\{\mathbf{h}_{it}\mathbf{x}_{it}^{*\prime}\right\}\right)\right\| = O\left(\frac{1}{N}\right) \to 0.$$

For $\widehat{\boldsymbol{\alpha}}_{IV}^{RMA\ell}$, we have

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

$var\left(\widehat{\mathbf{A}}_{IV}^{RMA\ell}\right)$ is shown to tend to zero in a similar way to $\widehat{\mathbf{A}}_{IV}^{OPT}$.

(b),(c): First, we consider $\sqrt{NT}\widehat{\mathbf{b}}_{IV}^{OPT}$.

$$\sqrt{NT}\widehat{\mathbf{b}}_{IV}^{OPT} = \frac{1}{\sqrt{NT}}\sum_{i=1}^{N}\sum_{t=1}^{T-1}\mathbf{h}_{it}v_{it}^*$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T-1} c_t \left[\mathbf{I}_p - O\left(\frac{1}{T-t} \right) \right] \left(\mathbf{w}_{i,t-1} + \boldsymbol{g}_{it} \right) v_{it}^{*} \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T-1} c_t \mathbf{w}_{i,t-1} v_{it}^{*} + o_p(1) \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T-1} \left(1 - \frac{1}{T-t+1} \right) \\
&\qquad \times \left(\mathbf{w}_{i,t-1} v_{it} - \frac{\mathbf{w}_{i,t-1}(v_{i,t+1} + \cdots + v_{iT})}{T-t} \right) + o_p(1) \\
&= \frac{1}{\sqrt{NT}} \sum_{i=1}^{N} \sum_{t=1}^{T-1} \mathbf{w}_{i,t-1} v_{it} + o_p(1)
\end{aligned}
$$

Then, using the central limit theorem of Phillips and Moon (1999), we have[14]

$$\sqrt{NT}\widehat{\mathbf{b}}_{IV}^{OPT} \xrightarrow{d} \mathcal{N}\left[0, \sigma_v^2 E\left(\mathbf{w}_{i,t-1} \mathbf{w}_{i,t-1}' \right) \right].$$

The result for $\sqrt{NT}\widehat{\mathbf{b}}_{IV}^{RMA\ell}$ is obtained in a similar way.

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

□

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.

[14] See also Hahn and Kuersteiner (2002) and Lee (2005)

[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," *Journal of Econometrics*, 127, 574-617.

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

[17] Hayakawa, K. (2008): "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										
α_1	T	N	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{RMA1}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{RMA0}$	$\widehat{\alpha}_{IV}^{RMA1}$	$\widehat{\alpha}_{IV}^{RMA2}$	$\widehat{\alpha}_{IV}^{RMA3}$	$\widehat{\alpha}_{IV}^{RMA4}$
0.5	10	100	0.467	0.493	0.500	0.499	0.499	0.498	0.497	0.497
0.5	10	500	0.493	0.499	0.501	0.500	0.500	0.500	0.499	0.500
0.5	15	100	0.472	0.495	0.499	0.498	0.498	0.498	0.498	0.499
0.5	15	300	0.490	0.499	0.501	0.500	0.500	0.499	0.499	0.500
0.5	20	100	0.476	0.498	0.500	0.500	0.500	0.499	0.499	0.499
0.5	20	200	0.488	0.499	0.500	0.500	0.500	0.500	0.500	0.500
0.5	50	100	0.482	0.499	0.500	0.500	0.500	0.500	0.500	0.500
0.5	100	100	0.484	0.500	0.500	0.500	0.500	0.500	0.500	0.500
0.9	10	100	0.691	0.805	0.906	0.891	0.894	0.889	0.884	0.882
0.9	10	500	0.827	0.878	0.903	0.897	0.897	0.899	0.898	0.898
0.9	15	100	0.767	0.862	0.902	0.896	0.896	0.897	0.896	0.897
0.9	15	300	0.833	0.885	0.902	0.899	0.899	0.899	0.900	0.900
0.9	20	100	0.802	0.881	0.900	0.899	0.899	0.898	0.898	0.898
0.9	20	200	0.837	0.892	0.902	0.900	0.900	0.900	0.900	0.899
0.9	50	100	0.860	0.897	0.899	0.900	0.900	0.899	0.899	0.899
0.9	100	100	0.875	0.899	0.900	0.900	0.900	0.900	0.900	0.900

IQR										
α_1	T	N	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{RMA1}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{RMA0}$	$\widehat{\alpha}_{IV}^{RMA1}$	$\widehat{\alpha}_{IV}^{RMA2}$	$\widehat{\alpha}_{IV}^{RMA3}$	$\widehat{\alpha}_{IV}^{RMA4}$
0.5	10	100	0.075	0.084	0.117	0.088	0.086	0.092	0.103	0.118
0.5	10	500	0.034	0.038	0.050	0.039	0.039	0.041	0.046	0.052
0.5	15	100	0.050	0.055	0.082	0.057	0.056	0.057	0.062	0.067
0.5	15	300	0.030	0.032	0.048	0.033	0.032	0.034	0.035	0.038
0.5	20	100	0.040	0.043	0.069	0.044	0.044	0.044	0.046	0.047
0.5	20	200	0.028	0.030	0.048	0.030	0.031	0.031	0.032	0.033
0.5	50	100	0.020	0.021	0.038	0.021	0.021	0.021	0.021	0.021
0.5	100	100	0.013	0.014	0.025	0.014	0.014	0.014	0.014	0.014
0.9	10	100	0.156	0.222	0.407	0.291	0.278	0.289	0.314	0.367
0.9	10	500	0.086	0.114	0.174	0.127	0.125	0.130	0.139	0.159
0.9	15	100	0.085	0.117	0.237	0.143	0.139	0.137	0.142	0.150
0.9	15	300	0.058	0.073	0.136	0.082	0.079	0.078	0.081	0.086
0.9	20	100	0.059	0.078	0.176	0.088	0.087	0.086	0.088	0.090
0.9	20	200	0.045	0.057	0.120	0.064	0.062	0.062	0.062	0.065
0.9	50	100	0.018	0.021	0.066	0.023	0.023	0.023	0.023	0.023
0.9	100	100	0.008	0.010	0.037	0.011	0.011	0.011	0.011	0.011

MAE										
α_1	T	N	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{RMA1}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{RMA0}$	$\widehat{\alpha}_{IV}^{RMA1}$	$\widehat{\alpha}_{IV}^{RMA2}$	$\widehat{\alpha}_{IV}^{RMA3}$	$\widehat{\alpha}_{IV}^{RMA4}$
0.5	10	100	0.043	0.042	0.058	0.044	0.043	0.046	0.051	0.059
0.5	10	500	0.018	0.019	0.025	0.019	0.020	0.020	0.023	0.026
0.5	15	100	0.032	0.028	0.040	0.028	0.028	0.028	0.031	0.033
0.5	15	300	0.016	0.016	0.024	0.016	0.016	0.017	0.017	0.019
0.5	20	100	0.027	0.022	0.034	0.022	0.022	0.022	0.023	0.024
0.5	20	200	0.016	0.015	0.024	0.015	0.015	0.015	0.016	0.017
0.5	50	100	0.018	0.010	0.019	0.010	0.010	0.011	0.011	0.011
0.5	100	100	0.016	0.007	0.012	0.007	0.007	0.007	0.007	0.007
0.9	10	100	0.209	0.131	0.203	0.146	0.139	0.146	0.158	0.185
0.9	10	500	0.075	0.059	0.087	0.064	0.063	0.065	0.070	0.080
0.9	15	100	0.133	0.064	0.118	0.072	0.070	0.069	0.071	0.075
0.9	15	300	0.067	0.037	0.067	0.041	0.039	0.039	0.040	0.043
0.9	20	100	0.098	0.040	0.087	0.044	0.043	0.043	0.044	0.045
0.9	20	200	0.063	0.029	0.060	0.032	0.031	0.031	0.031	0.033
0.9	50	100	0.040	0.011	0.032	0.012	0.011	0.012	0.012	0.011
0.9	100	100	0.025	0.005	0.018	0.005	0.005	0.005	0.005	0.005

Table 2: Simulation results for an AR(2) model

Median											
T	N	α_1	α_2	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{RMA1}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{RMA0}$	$\widehat{\alpha}_{IV}^{RMA1}$	$\widehat{\alpha}_{IV}^{RMA2}$	$\widehat{\alpha}_{IV}^{RMA3}$	$\widehat{\alpha}_{IV}^{RMA4}$
10	100	0.6		0.565	0.586	0.599	0.598	0.597	0.598	0.597	0.598
10	100		-0.1	-0.117	-0.108	-0.101	-0.102	-0.102	-0.103	-0.104	-0.102
10	500	0.6		0.593	0.598	0.600	0.600	0.600	0.600	0.600	0.599
10	500		-0.1	-0.104	-0.102	-0.101	-0.101	-0.101	-0.101	-0.101	-0.101
15	100	0.6		0.577	0.595	0.601	0.600	0.600	0.600	0.600	0.599
15	100		-0.1	-0.114	-0.103	-0.101	-0.101	-0.101	-0.101	-0.101	-0.101
15	300	0.6		0.591	0.598	0.600	0.600	0.600	0.600	0.600	0.599
15	300		-0.1	-0.105	-0.101	-0.100	-0.100	-0.100	-0.100	-0.100	-0.100
20	100	0.6		0.590	0.599	0.601	0.600	0.600	0.600	0.600	0.600
20	100		-0.1	-0.106	-0.101	-0.100	-0.100	-0.100	-0.100	-0.100	-0.100
20	200	0.6		0.590	0.599	0.601	0.600	0.600	0.600	0.600	0.600
20	200		-0.1	-0.106	-0.101	-0.100	-0.100	-0.100	-0.100	-0.100	-0.100
50	100	0.6		0.588	0.600	0.601	0.601	0.601	0.600	0.601	0.601
50	100		-0.1	-0.110	-0.101	-0.100	-0.100	-0.100	-0.100	-0.100	-0.100
100	100	0.6		0.590	0.600	0.600	0.600	0.600	0.600	0.600	0.600
100	100		-0.1	-0.109	-0.100	-0.101	-0.100	-0.100	-0.100	-0.100	-0.100

IQR											
T	N	α_1	α_2	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{RMA1}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{RMA0}$	$\widehat{\alpha}_{IV}^{RMA1}$	$\widehat{\alpha}_{IV}^{RMA2}$	$\widehat{\alpha}_{IV}^{RMA3}$	$\widehat{\alpha}_{IV}^{RMA4}$
10	100	0.6		0.067	0.077	0.101	0.079	0.079	0.084	0.094	0.111
10	100		-0.1	0.051	0.056	0.064	0.058	0.059	0.063	0.068	0.076
10	500	0.6		0.032	0.035	0.046	0.036	0.036	0.038	0.042	0.050
10	500		-0.1	0.024	0.026	0.029	0.027	0.026	0.029	0.031	0.035
15	100	0.6		0.046	0.050	0.070	0.050	0.050	0.051	0.054	0.059
15	100		-0.1	0.038	0.041	0.050	0.043	0.043	0.045	0.047	0.051
15	300	0.6		0.028	0.029	0.041	0.030	0.030	0.031	0.032	0.034
15	300		-0.1	0.023	0.024	0.029	0.024	0.025	0.026	0.027	0.029
20	100	0.6		0.027	0.029	0.041	0.029	0.029	0.029	0.030	0.031
20	100		-0.1	0.024	0.025	0.031	0.026	0.026	0.026	0.027	0.028
20	200	0.6		0.027	0.029	0.041	0.029	0.029	0.029	0.030	0.031
20	200		-0.1	0.024	0.025	0.031	0.026	0.026	0.026	0.027	0.028
50	100	0.6		0.020	0.021	0.031	0.021	0.021	0.021	0.022	0.022
50	100		-0.1	0.019	0.020	0.026	0.020	0.020	0.020	0.021	0.021
100	100	0.6		0.014	0.014	0.020	0.014	0.014	0.014	0.014	0.014
100	100		-0.1	0.014	0.014	0.019	0.014	0.014	0.014	0.014	0.014

MAE											
T	N	α_1	α_2	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{RMA1}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{RMA0}$	$\widehat{\alpha}_{IV}^{RMA1}$	$\widehat{\alpha}_{IV}^{RMA2}$	$\widehat{\alpha}_{IV}^{RMA3}$	$\widehat{\alpha}_{IV}^{RMA4}$
10	100	0.6		0.043	0.040	0.050	0.040	0.040	0.042	0.047	0.055
10	100		-0.1	0.028	0.029	0.032	0.029	0.029	0.032	0.034	0.038
10	500	0.6		0.016	0.017	0.023	0.018	0.018	0.019	0.021	0.025
10	500		-0.1	0.012	0.013	0.014	0.013	0.013	0.014	0.016	0.018
15	100	0.6		0.029	0.025	0.035	0.025	0.025	0.026	0.027	0.030
15	100		-0.1	0.021	0.021	0.025	0.021	0.021	0.022	0.023	0.025
15	300	0.6		0.015	0.015	0.021	0.015	0.015	0.015	0.016	0.017
15	300		-0.1	0.012	0.012	0.014	0.012	0.012	0.013	0.014	0.015
20	100	0.6		0.015	0.014	0.021	0.015	0.015	0.015	0.015	0.015
20	100		-0.1	0.013	0.013	0.015	0.013	0.013	0.013	0.014	0.014
20	200	0.6		0.015	0.014	0.021	0.015	0.015	0.015	0.015	0.015
20	200		-0.1	0.013	0.013	0.015	0.013	0.013	0.013	0.014	0.014
50	100	0.6		0.014	0.011	0.015	0.011	0.011	0.011	0.011	0.011
50	100		-0.1	0.012	0.010	0.013	0.010	0.010	0.010	0.010	0.010
100	100	0.6		0.011	0.007	0.010	0.007	0.007	0.007	0.007	0.007
100	100		-0.1	0.010	0.007	0.010	0.007	0.007	0.007	0.007	0.007

Table 2 (Cont.)

Median											
T	N	α_1	α_2	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{RMA1}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{RMA0}$	$\widehat{\alpha}_{IV}^{RMA1}$	$\widehat{\alpha}_{IV}^{RMA2}$	$\widehat{\alpha}_{IV}^{RMA3}$	$\widehat{\alpha}_{IV}^{RMA4}$
10	100	0.6		0.357	0.430	0.600	0.583	0.583	0.586	0.580	0.577
10	100		0.3	0.196	0.225	0.298	0.285	0.287	0.286	0.284	0.285
10	500	0.6		0.482	0.541	0.597	0.595	0.595	0.593	0.596	0.593
10	500		0.3	0.250	0.274	0.299	0.297	0.297	0.296	0.297	0.296
15	100	0.6		0.445	0.524	0.608	0.596	0.597	0.597	0.597	0.597
15	100		0.3	0.227	0.263	0.303	0.295	0.295	0.296	0.296	0.295
15	300	0.6		0.508	0.566	0.602	0.600	0.600	0.600	0.600	0.600
15	300		0.3	0.258	0.284	0.300	0.299	0.299	0.299	0.299	0.299
20	100	0.6		0.521	0.575	0.605	0.600	0.600	0.600	0.600	0.600
20	100		0.3	0.261	0.288	0.301	0.299	0.299	0.298	0.299	0.299
20	200	0.6		0.521	0.575	0.605	0.600	0.600	0.600	0.600	0.600
20	200		0.3	0.261	0.288	0.301	0.299	0.299	0.298	0.299	0.299
50	100	0.6		0.561	0.596	0.601	0.600	0.600	0.601	0.600	0.601
50	100		0.3	0.275	0.296	0.300	0.299	0.299	0.299	0.299	0.299
100	100	0.6		0.579	0.599	0.599	0.600	0.600	0.600	0.600	0.600
100	100		0.3	0.284	0.299	0.299	0.300	0.300	0.300	0.300	0.300

IQR											
T	N	α_1	α_2	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{RMA1}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{RMA0}$	$\widehat{\alpha}_{IV}^{RMA1}$	$\widehat{\alpha}_{IV}^{RMA2}$	$\widehat{\alpha}_{IV}^{RMA3}$	$\widehat{\alpha}_{IV}^{RMA4}$
10	100	0.6		0.161	0.225	0.474	0.365	0.353	0.373	0.390	0.465
10	100		0.3	0.073	0.098	0.207	0.160	0.157	0.165	0.170	0.192
10	500	0.6		0.101	0.129	0.222	0.161	0.157	0.161	0.178	0.200
10	500		0.3	0.044	0.056	0.092	0.069	0.067	0.070	0.074	0.083
15	100	0.6		0.093	0.126	0.292	0.166	0.161	0.162	0.169	0.183
15	100		0.3	0.050	0.062	0.140	0.083	0.082	0.082	0.085	0.090
15	300	0.6		0.066	0.083	0.166	0.098	0.095	0.097	0.099	0.107
15	300		0.3	0.032	0.041	0.076	0.047	0.047	0.048	0.049	0.051
20	100	0.6		0.052	0.063	0.141	0.074	0.072	0.072	0.075	0.075
20	100		0.3	0.030	0.036	0.071	0.040	0.039	0.040	0.040	0.041
20	200	0.6		0.052	0.063	0.141	0.074	0.072	0.072	0.075	0.075
20	200		0.3	0.030	0.036	0.071	0.040	0.039	0.040	0.040	0.041
50	100	0.6		0.024	0.026	0.066	0.028	0.028	0.028	0.028	0.028
50	100		0.3	0.020	0.021	0.044	0.023	0.023	0.023	0.023	0.023
100	100	0.6		0.014	0.015	0.034	0.015	0.015	0.015	0.015	0.015
100	100		0.3	0.013	0.014	0.027	0.014	0.014	0.014	0.014	0.015

MAE											
T	N	α_1	α_2	$\widehat{\alpha}_{GMM}^{LEV}$	$\widehat{\alpha}_{GMM}^{RMA1}$	$\widehat{\alpha}_{IV}^{LEV}$	$\widehat{\alpha}_{IV}^{RMA0}$	$\widehat{\alpha}_{IV}^{RMA1}$	$\widehat{\alpha}_{IV}^{RMA2}$	$\widehat{\alpha}_{IV}^{RMA3}$	$\widehat{\alpha}_{IV}^{RMA4}$
10	100	0.6		0.243	0.176	0.236	0.184	0.179	0.187	0.196	0.233
10	100		0.3	0.104	0.079	0.104	0.081	0.079	0.083	0.087	0.099
10	500	0.6		0.118	0.078	0.110	0.081	0.078	0.081	0.089	0.101
10	500		0.3	0.050	0.034	0.046	0.035	0.034	0.035	0.037	0.042
15	100	0.6		0.155	0.084	0.146	0.083	0.080	0.081	0.084	0.092
15	100		0.3	0.073	0.042	0.070	0.041	0.041	0.041	0.043	0.045
15	300	0.6		0.092	0.047	0.083	0.049	0.047	0.048	0.049	0.053
15	300		0.3	0.042	0.023	0.038	0.024	0.023	0.024	0.024	0.025
20	100	0.6		0.079	0.036	0.070	0.037	0.036	0.036	0.037	0.037
20	100		0.3	0.039	0.019	0.035	0.020	0.020	0.020	0.020	0.021
20	200	0.6		0.079	0.036	0.070	0.037	0.036	0.036	0.037	0.037
20	200		0.3	0.039	0.019	0.035	0.020	0.020	0.020	0.020	0.021
50	100	0.6		0.039	0.014	0.033	0.014	0.014	0.014	0.014	0.014
50	100		0.3	0.025	0.011	0.022	0.011	0.011	0.011	0.011	0.012
100	100	0.6		0.021	0.007	0.016	0.007	0.007	0.007	0.007	0.007
100	100		0.3	0.016	0.007	0.013	0.007	0.007	0.007	0.007	0.007