Indirect inference and bias-corrected pooled least squares estimators for dynamic panels with short T *

In Choi† and Sanghyun Jung‡

July 15, 2019

Abstract

This paper proposes new estimators for panel autoregressive (PAR) models with short time dimensions (T) and large cross sections. These estimators are based on the cross-sectional regression model using the first time series observations as a regressor and the last as a dependent variable. The regressors and errors of this regression model are correlated. The first estimator is the maximum likelihood estimator (MLE) under the assumption of normal distributions. This estimator is called the cross-sectional MLE (CSMLE). The second estimator is the indirect inference estimator (IIE) that improves the finite-sample bias of the CSMLE. The third estimator is the bias-corrected pooled least squares estimator (BCPLSE) that eliminates the asymptotic bias of the pooled least squares estimator by using the CSMLE. The CSMLE, IIE and BCPLSE are extended to the PAR model with endogenous time-invariant regressors. The CSMLE, IIE and BCPLSE provide consistent estimates of the PAR coefficients for stationary, unit root and explosive PAR models, estimate the coefficients of time-invariant

*This paper is an expanded and revised version of the first author's paper entitled as "Cross-sectional maximum likelihood and bias-corrected pooled least squares estimators for dynamic panels with short T." Some results in this paper were presented at the 22nd Panel Data conference held in Perth, Australia in 2016. The author thanks the participants, especially Stéphanne Bonhomme, for helpful comments. This paper was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2015S1A5A2A01009333).

†Sogang University, Seoul, Korea. Personal Webpage: http://inchoi.sogang.ac.kr. E-mail: inchoi@gmail.com, inchoi@sogang.ac.kr.

‡Sogang University, Seoul, Korea. E-mail: sanghnjung@gmail.com

regressors consistently and can be computed as long as $T \geq 2$. Their finite sample properties are compared with those of some other estimators for the PAR model of order 1. The estimators of this paper are shown to perform quite well in finite samples.

Keywords: dynamic panels, maximum likelihood estimator, indirect inference estimator, pooled least squares estimator, stationarity, unit root, explosive root

1 Introduction

Panel autoregressive (PAR) models have been the focus of much research in recent years. When the number of time series observations (T) is small, the PAR model of order 1 (PAR(1) model) is often used in practice and, accordingly, much emphasis has been given to the PAR(1) model. There are several approaches to the estimation of the PAR(1) model. The most popular approach uses generalized methods of moments (GMM) estimators. These estimators improve on Anderson and Hsiao's (1981) instrumental variables (IV) estimator in terms of efficiency. The improved efficiency stems from additional moment conditions these estimators employ. Most notable papers studying the GMM estimators for the PAR(1) model are Arellano and Bond (1991), Ahn and Schmidt (1995) and Blundell and Bond (1998). But when the PAR(1) coefficient is close to unity, Arellano and Bond's and Ahn and Schmidt's GMM estimators are subject to the problem of weak instruments as analyzed in Blundell and Bond. Blundell and Bond's estimator does not share the same problem, is known to be more efficient than the other two, and has been used widely in applications.[1] Hahn (1999) shows that the efficiency gain of Blundell and Bond's estimator comes from the initial condition they use. In addition, Hahn, Hausman and Kuersteiner (2007) introduce a long difference regression model for the PAR(1) process and devise a GMM estimator. They report via simulation that it is sometimes more efficient than Blundell and Bond's GMM estimator. Ashley and Sun (2016) improve the GMM estimators for the case of stationarity using Hansen, Heaton and Yaron's

[1] But they can also be subject to the weak instrument problem as reported in Hayakawa (2007) and Bun and Windmeijer (2010).

(1996) continuous-updating method. Baltagi (2008) and Bun and Sarafidis (2015) provide nice reviews on the literature related to the GMM estimators for the PAR model.

The second approach to the estimation of the PAR model is the maximum likelihood estimation in differences. As alternatives to GMM estimation, Hsiao, Pesaran and Tahmiscioglu (2002) and Kruiniger (2008) have proposed maximum likelihood estimators (MLEs) using the differenced model, which are called the first difference MLEs (FDMLEs). Differencing eliminates individual effects and a unit root, if any, so that standard asymptotic theories seem to be applicable. But the FDMLEs have a few problems that deserve our attention. First, functional forms of the likelihood functions invite a host of issues. Kruiniger (2008) assumes a stable coefficient[2], while Hsiao, Pesaran and Tahmiscioglu (2002) do not necessarily do so. When the coefficient is greater than 1, however, the likelihood functions of Hsiao, Pesaran and Tahmiscioglu (2002) and Kruiniger (2008) are no longer true likelihood functions as indicated by Han and Phillips (2013, p.37, footnote 2). Asymptotic results may also depend on the likelihood function. Han and Phillips (2013, pp. 41-42) prove that Kruiniger's (2008) MLE for the unit root case has a normal distribution in the limit only when the stationary likelihood is extended outside its natural domain of definition. Non-normal distributions follow if the likelihood function is formulated in different ways. In other words, Kruiniger's limiting normal distribution for the case of a unit root is the consequence of using an incorrect likelihood function. Second, because the FDMLE employs the differenced PAR(1) model, it becomes impossible to estimate the coefficients of time-invariant regressors. Third, as indicated by Han and Phillips (2013, p.36), the likelihood function under the stationarity assumption behaves so wildly when the PAR coefficient is near and above unity that the global maximum is often unidentified by numerical optimization procedures. This is a practical problem researchers need to pay attention to in applications.

The third approach employs least squares estimators (LSEs). Han and Phillips (2010) and Han, Phillips and Sul (2014) introduce transformations of the PAR(1)

[2] Here, a stable coefficient implies that its value belongs to the open interval (-1,1).

model that make the regressors and errors uncorrelated. Then, they apply LSEs to estimate the PAR(1) coefficient. Being able to use LSEs is a convenient aspect of this approach, but validity of their transformations depends on their assumptions on the model and may not hold under different assumptions (see Subsubsection 7.1.2 for more discussion). Hahn and Kuersteiner (2002) propose a bias-corrected Within-OLS estimator of the PAR(1) coefficient. Gourieroux, Phillips and Yu (2010) also introduce a bias-corrected Within-OLS estimator using the indirect inference method of Gourieroux, Monfort and Renault (1993). In Hahn and Kuersteiner's and Gourieroux, Phillips and Yu's approaches, it is required that $T \to \infty$ and that the PAR(1) coefficient takes values less than 1, which may limit their use in practice.

The fourth approach is the maximum likelihood estimation in level suggested by Anderson and Hsiao (1982) and Hsiao and Zhou (2018). They assume exogenous regressors. How to extend Anderson and Hsiao's approach to the case of endogenous regressors does not seem to be obvious. In addition, they do not consider the case where the PAR(1) coefficient takes values greater than or equal to 1, and the type of the initial variable considered in this paper.

The other approaches using the level data include Bai (2013) and Hayakawa (2012). The former proposes MLE utilizing the factor structure of the PAR(1) model and the latter suggests a GMM estimator. Both of them assume a stable PAR(1) coefficient.

The purpose of this paper is to suggest a new approach for the estimation of the PAR(1) model. This approach employs the cross-sectional regression model using the first time series observations as a regressor and the last as a dependent variable. Because the initial observation and the individual effect are correlated, the regressors and errors of the regression model are dependent, making LSEs inappropriate to use. Instead, an MLE is proposed for the regression model, assuming normal distributions. The estimator is called the cross-sectional MLE (CSMLE). This estimator employs a basic result in multivariate analysis for endogeneity correction. The likelihood-based endogeneity correction of this paper turns out to be quite promising and can also be used for other structural econometric models.

MLEs are often subject to finite-sample biases and, so is the CSMLE. Thus, we propose to employ the indirect inference estimator (IIE) of Smith (1993) and Gouriéroux, Monfort and Renault (1993) in conjunction with the CSMLE. Gouriéroux, Phillips and Yu (2010) also employ the indirect inference estimation method for dynamic panels and report good finite sample properties of their estimator. But the parameter space of the PAR(1) coefficient is restricted to (-1,1) in their paper. The IIE improves on the CSMLE according to our simulation results.

The CSMLE makes it feasible to estimate the asymptotic bias of the pooled least squares estimator for the dynamic panel regression model. The estimated bias can be used to construct a bias-corrected pooled LSE that will be called the bias-corrected pooled least squares estimator (BCPLSE).

Asymptotic properties of the CSMLE, IIE and BCPLSE are studied in this paper. The estimators are also extended to the PAR(1) model with endogenous and time-invariant regressors. In addition, their finite sample properties are compared with those of some GMM estimators and LSEs. It is found that the new estimators of this paper perform quite well compared to the other estimators.

There are some advantages of this paper's approach. First, there are no restrictions on the parameter space of the PAR(1) coefficient. It can be any compact subset of the real line. By contrast, most papers require the parameter space to be either (-1,1) or (-1,1]. The parameter space has important implications for the conventional MLEs as discussed above. Moreover, it is possible to estimate the PAR(1) coefficient consistently by using the CSMLE, IIE and BCPLSE even when it is greater than 1. This explosive case is potentially important for applications to financial data. For example, bubbles of financial markets are modelled using an explosive, univariate AR process in Phillips, Wu and Yu (2011).

Second, coefficients of time-invariant, endogenous regressors can be estimated along with other parameters. Because all the procedures discussed so far except Anderson and Hsiao (1982) use differencing or within-group demeaning, it is impossible to estimate the coefficients of time-invariant regressors. Of course, one can use the estimated PAR(1) coefficient for the cross-sectional IV regression that yields a consistent estimator of the coefficient. But the IV regression needs instrumental variables that are sometimes difficult to find or only weakly correlated with the regressors in applications. The CSMLE, IIE and BCPLSE of this paper do not require instrumental variables, which must be a useful aspect of the estimators particularly when instruments are non-existent or weakly correlated with the regressors. Instead, a likelihood-based endogeneity correction is used to obtain consistent estimators of coefficients of the endogenous regressors.

Third, the estimators of this paper require $T \geq 2$, while most of the aforementioned estimators require at least 3 or 4 time series observations. For new panel data sets, this is an important advantage.

There is a difference between this paper's model and that of the papers mentioned above. This paper takes the random effects approach in the sense that individual effects variables are modelled as random variables having a common variance. For the papers mentioned above except Anderson and Hsiao (1982), individual effects variables can be either random or nonrandom, because they are eliminated anyhow by differencing or within-group demeaning. In spite of this difference, the LSE and least-squares-dummy-variables estimator are inconsistent in both the models when T is fixed. Moreover, it does not seem to be important in many applications whether or not the individual effects variables are random. Of course, if there is any compelling reason to believe that individual effects are constants or that they are random variables with heteroskedasticity, methods using the differenced model need to be used. But the method of this paper works quite well when the individual effects variables have only mild degree of heteroskedasticity as will be seen in Section 7 via simulation.

This paper is planned as follows. Section 2 introduces the model and basic assumptions. Section 3 introduces the CSMLE and studies its asymptotic properties. Section 4 studies the IIE. Section 5 proposes the BCPLSE and studies its asymptotic properties. Section 6 extends the CSMLE, IIE and BCPLSE to the PAR(1) model with endogenous, time-invariant regressors. Section 7 introduces various estimators used for simulation and reports simulation results. Section 8 provides summary and further remarks. Proofs are relegated to Appendix.

A few words on our notation. $\mathbb{R}$ and $\mathbb{R}^+$ denote the set of real numbers and the set of positive real numbers, respectively. All the limits are taken as $N \to \infty$. Convergence in probability and weak convergence are denoted by $\xrightarrow{p}$ and $\xrightarrow{d}$, respectively.

2 The model and basic assumptions

We are concerned with the unobserved components model for the panel data $\{y_{it}\}$

$$y_{it} = \mu_i + x_{it}, \ (i = 1, \ldots, N; \ t = 1, \ldots, T; \ T \geq 2), \tag{1}$$

where $\{x_{it}\}$ is unobserved and follows the AR(1) model

$$x_{it} = \alpha x_{i,t-1} + u_{it}. \tag{2}$$

As usual, i and t are indices for individuals and time, respectively, and $\{\mu_i\}$ denotes the unobserved individual effects. In Model (1), $\{x_{it}\}$ brings dynamics to the evolvement of $\{y_{it}\}$. Model (1) can be written in a more familiar format as

$$y_{it} = \mu_i(1 - \alpha) + \alpha y_{i,t-1} + u_{it}. \tag{3}$$

This model has been used in various works for the estimation of dynamic panels. The reader is referred to Hsiao (2003, p.76) for the comparison of Models (1) and (3).

Regarding the individual effects variable μ_i, let

$$\mu_i = \mu + m_i,$$

where μ is a fixed constant and m_i is a random variable. Using this relation, Model (3) can be written as

$$y_{it} = \mu(1 - \alpha) + \alpha y_{i,t-1} + u_{it} + m_i(1 - \alpha) \tag{4}$$

The following assumption introduces the basic characteristics of the individual effects $\{m_i\}$, the error terms $\{u_{it}\}$ and the initial variables $\{x_{i1}\}$.

Assumption 1 *Let* $u_i = [u_{i2}, ..., u_{iT}]'$. *Assume*

$$\begin{pmatrix} m_i \\ u_i \\ x_{i1} \end{pmatrix} \sim \mathrm{i.i.d.} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{\mathrm{m}} & 0 & 0 \\ 0 & \sigma^2_{\mathrm{u}} I_{T-1} & 0 \\ 0 & 0 & \sigma^2_{\mathrm{x}_1} \end{bmatrix} \right),$$

where $\sigma^2_{\mathrm{m}} > 0$, $\sigma^2_{\mathrm{u}} > 0$ *and* $\sigma^2_{\mathrm{x}_1} > 0$.

Under this assumption, the variance of x_{i1} is not assumed to be equal to $\sigma^2_{\mathrm{u}}/(1-\alpha^2)$, which is often assumed in the literature on time series and dynamic panels when $\alpha \in (-1, 1)$. In this paper, we just require that parameter α belongs to a set of real numbers. It is even allowed to be greater than 1. Under Assumption 1, note also that the initial observation y_{i1} and the individual effect variable μ_i are correlated. In fact, $Cov(y_{i1}, \mu_i) = \sigma^2_{\mathrm{m}}$. This shows that there is no need for an extra parameter that signifies the non-zero covariance between y_{i1} and μ_i.

Under Assumption 1, running OLS or Within-OLS on Model (4) does not provide a consistent estimator of the parameter α because $Cov(y_{i,t-1}, u_{it} + m_i(1-\alpha)) = Cov(m_i + x_{i,t-1}, u_{it} + m_i(1-\alpha)) = (1-\alpha)\sigma^2_{\mathrm{m}}$ is not zero unless α is equal to one.

3 Cross-sectional maximum likelihood estimation

This section introduces a cross-sectional maximum likelihood estimator (CSMLE) for Model (1).[3] Model (1) can be written as

$$y_{it} = \mu + \alpha^{t-1} x_{i1} + w_{it} + m_i, \ (t = 2, ..., T), \tag{5}$$

[3] The CSMLE can also be considered for Model (1) in differences. That is, the CSMLE can be derived for the model $\Delta y_{iT} = \alpha^{T-2}\Delta y_{i2} + d_{iT}$, where $d_{iT} = \sum_{j=0}^{T-3} \alpha^j \Delta u_{i,T-j}$. But the resulting information matrix becomes singular at $\alpha = 1$, which bring some complications for inference. In addition, when Model (1) is extended to that with time-invariant regressors, differencing eliminates those variables, making it impossible to estimate their effects on the dependent variable. For these reasons, the CSMLE for Model (1) in differences is not pursued here.

where $w_{it} = \sum_{j=0}^{t-2} \alpha^j u_{i,t-j}$. Assume $\mu = 0$ for simplicity from now on.[4] Because $x_{i1} = y_{i1} - \mu_i$, relation (5) gives for $t = T$

$$\begin{aligned} y_{iT} &= \alpha^{T-1} x_{i1} + w_{iT} + m_i \\ &= \alpha^{T-1} y_{i1} + w_{iT} + (1-\alpha^{T-1}) m_i. \end{aligned} \tag{6}$$

Running OLS on this equation does not yield a consistent estimator of the regression coefficients because y_{i1} and m_i are correlated unless $\alpha = 1$. As an alternative, we consider maximum likelihood estimation in this section. For the maximum likelihood estimation, it is required that $T \geq 2$. Even at $T = 2$, the MLE of α can be obtained. This feature is not shared with extant estimation methods in dynamic panel data analysis.

Let $v_{iT} = w_{iT} + (1-\alpha^{T-1}) m_i$. Then, Assumption 1 gives

$$\begin{pmatrix} v_{iT} \\ y_{i1} \end{pmatrix} \sim \text{i. i. d.} \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{bmatrix} \right), \tag{7}$$

where

$$\omega_{11} = \sigma_{\mathrm{u}}^2 \sum_{j=0}^{T-2} \alpha^{2j} + (1-\alpha^{T-1})^2 \sigma_{\mathrm{m}}^2,\ \omega_{12} = (1-\alpha^{T-1}) \sigma_{\mathrm{m}}^2 \text{ and } \omega_{22} = \sigma_{\mathrm{x}_1}^2 + \sigma_{\mathrm{m}}^2.$$

Suppose that we project $\mathbf{v}_T = [v_{1T}, ..., v_{NT}]'$ onto the space spanned by $\mathbf{y}_1 = [y_{11}, ..., y_{N1}]'$ such that

$$\mathbf{v}_T = \mathbf{y}_1 (\mathbf{y}_1' \mathbf{y}_1)^{-1} \mathbf{y}_1' \mathbf{v}_T + \left(I - \mathbf{y}_1 (\mathbf{y}_1' \mathbf{y}_1)^{-1} \mathbf{y}_1' \right) \mathbf{v}_T.$$

Replacing $(\mathbf{y}_1' \mathbf{y}_1)^{-1} \mathbf{y}_1' \mathbf{v}_T$ with its probability limit $\frac{\omega_{12}}{\omega_{22}}$, this relation gives

$$v_{iT} = \frac{\omega_{12}}{\omega_{22}} y_{i1} + r_{iT}, \tag{8}$$

where $r_{iT} \sim iid\ (0,\ \omega_{11.2})$ with $\omega_{11.2} = \omega_{11} - \frac{\omega_{12}^2}{\omega_{22}}$. This representation bears resemblance to Chamberlain's (1984) approach of modelling individual effects variables as linear combinations of regressors and unknown errors. But relation (8) runs deeper than that in the sense that the coefficients ω_{11} and ω_{12} contain unknown parameters

[4] In practice, we can demean $\{y_{iT}\}$ and $\{y_{1i}\}$ to make this assumption plausible.

of interest, and this aspect will be utilized for the maximum likelihood estimation. Note that $Cov(y_{i1}, r_{iT}) = 0$.

To introduce maximum likelihood estimators, now we assume

Assumption 2

$$r_{iT} \mid y_{i1} \sim \text{i. i. d. } \boldsymbol{N}\left(0,\ \omega_{11.2}\right).$$

Then, we have[5]

$$y_{iT} \mid y_{i1} \sim \text{i. i. d. } \boldsymbol{N}\left(\alpha^{T-1} y_{i1} + \frac{\omega_{12}}{\omega_{22}} y_{i1},\ \omega_{11.2}\right). \tag{9}$$

Thus, we obtain a linear regression model for $\{y_{iT}\}$ having normally distributed errors:

$$y_{iT} = \alpha^{T-1} y_{i1} + \frac{\omega_{12}}{\omega_{22}} y_{i1} + r_{iT}. \tag{10}$$

Let the probability density function of y_{i1} be denoted by $f(\cdot \mid \omega_{22})$. Relation (9) yields the likelihood function of the parameters $\alpha,\ \sigma_{\mathrm{u}}^2,\ \sigma_{\mathrm{m}}^2,\ \sigma_{\mathrm{x}_1}^2$ as

$$\begin{aligned} L(\alpha, \sigma_{\mathrm{u}}^2, \sigma_{\mathrm{m}}^2, \sigma_{\mathrm{x}_1}^2 \quad &| \quad y_{1T}, ..., y_{NT}, y_{11}, ..., y_{N1}) \\ &= \quad \Pi_{i=1}^N \frac{1}{\sqrt{2\pi\omega_{11.2}}} \exp\left(-\frac{(y_{iT} - \alpha^{T-1} y_{i1} - \frac{\omega_{12}}{\omega_{22}} y_{i1})^2}{2\omega_{11.2}}\right) \\ &\qquad \times \Pi_{i=1}^N f(y_{i1} \mid \omega_{22}). \end{aligned}$$

The log-likelihood function is derived from this as

$$\begin{aligned} l(\alpha, \sigma_{\mathrm{u}}^2, \sigma_{\mathrm{m}}^2, \sigma_{\mathrm{x}_1}^2) \quad &= \quad c - \frac{N}{2} \ln\left(\omega_{11.2}\right) - \sum_{i=1}^N \frac{(y_{iT} - \alpha^{T-1} y_{i1} - \frac{\omega_{12}}{\omega_{22}} y_{i1})^2}{2\omega_{11.2}} \\ &\qquad + \sum_{i=1}^N f(y_{i1} \mid \omega_{22}), \end{aligned} \tag{11}$$

where c is a constant.

It is most natural to maximize the likelihood function (11) with respect to the 4 parameters to obtain their MLEs. The MLEs are consistent as the conventional theory of MLEs (cf. Newey and McFadden, 1994) has established. But we will reparametrize the likelihood function using some information from method-of-moments estimators.

[5] This representation can also be obtained once we assume normality in Assumption 1 (see Muirhead, 1982, p.12). But we avoided this because assuming normality for y_{i1} is unnecessary.

There are two reasons for this. First, this reduces the number of parameters, which makes the related nonlinear optimization much simpler. In fact, according to our experiments, the MLEs using the method-of-moments estimators have better finite sample performance than those based on the full likelihood function (11). Second and more importantly, the information matrix of the full MLEs is singular.[6] The singular information matrix makes it very difficult to derive the asymptotic distribution of the MLEs relying on the conventional theory of MLEs and analyze the properties of the MLE of α which are the main focus of this paper.

To reduce the number of parameters in the log-likelihood function (11), assume that $\omega_{22} = \sigma^2_{\mathrm{x}_1} + \sigma^2_{\mathrm{m}}$ is known. The parameter ω_{22} can be estimated consistently by $\frac{1}{N}\sum_{i=1}^{N} y_{i1}^2$. Next, assume $\lambda = Var(\Delta y_{i2}) = (1-\alpha)^2\sigma^2_{\mathrm{x}_1} + \sigma^2_{\mathrm{u}}$ is known.[7] In practice, λ can be estimated consistently by $\frac{1}{N}\sum_{i=1}^{N}(\Delta y_{i2})^2$. Then, because $\sigma^2_{\mathrm{u}} = \lambda - (1-\alpha)^2(\omega_{22} - \sigma^2_{\mathrm{m}})$ and $\sigma^2_{\mathrm{x}_1} = \omega_{22} - \sigma^2_{\mathrm{m}}$, there remain only α and σ^2_{m} in the likelihood function.[8] Once the MLEs of α and σ^2_{m} are obtained, σ^2_{u} and $\sigma^2_{\mathrm{x}_1}$ can be estimated consistently using the estimates of ω_{22} and λ.

Assuming that ω_{22} and λ are known, the likelihood function is now written as

$$l(\alpha, \sigma^2_{\mathrm{m}}) = c - \frac{N}{2}\ln(\omega_{11.2}) - \sum_{i=1}^{N}\frac{(y_{iT} - \alpha^{T-1}y_{i1} - \frac{\omega_{12}}{\omega_{22}}y_{i1})^2}{2\omega_{11.2}}.$$

The MLEs of the parameters α and σ^2_{m} are obtained by maximizing this log-likelihood function. Now, we make an audacious assumption that σ^2_{m} is known for the following

[6] Some laborious calculations yield

$$\frac{\partial l}{\partial \sigma^2_{\mathrm{u}}} = \frac{\sum_{j=0}^{T-2}\alpha^{2j}}{c}\left[(1-\alpha^{T-1})\frac{\partial l}{\partial \alpha} + (T-1)\alpha^{T-2}\sigma^2_{\mathrm{x}_1}\left(\frac{\partial l}{\partial \sigma^2_{\mathrm{x}_1}} - \frac{\partial l}{\partial \sigma^2_{\mathrm{m}}}\right)\right],$$

$c = (1-\alpha^{T-1})a + (T-1)\alpha^{T-2}\sigma^2_{\mathrm{x}_1}b$, $a = 2\left\{\sigma^2_{\mathrm{u}}\sum_{j=0}^{T-2}j\alpha^{2j-1} - (T-1)(1-\alpha^{T-1})\alpha^{T-2}\sigma^2_{\mathrm{m}} + \omega_{12}\left((T-1)\alpha^{T-2}\sigma^2_{\mathrm{m}}\right)/\omega_{22}\right\}$, $b = \frac{\omega^2_{12} - \omega^{dm}_{11.2}\omega^2_{22}}{\omega^2_{22}}$ and $\omega^{dm}_{11.2} = (1-\alpha^{T-1})^2 - \frac{2\omega_{12}\omega_{22}(1-\alpha^{T-1}) - \omega^2_{12}}{\omega^2_{22}}$.

[7] One may wonder immediately why $\sum_{t=2}^{T} Var(\Delta y_{ti})$ is not used instead. But it turns out that this brings higher asymptotic variance of the MLE of α, so we chose to use $Var(\Delta y_{2i})$.

[8] Eliminating σ^2_{u} and σ^2_{m} brings the same asymptotic variance of the MLE of α as that in this section. We can also eliminate $\sigma^2_{\mathrm{x}_1}$ and σ^2_{m} by assuming that $\omega_{22} = \sigma^2_{\mathrm{x}_1} + \sigma^2_{\mathrm{m}}$ and $\alpha^{T-1} + \frac{\omega_{12}}{\omega_{22}}$ are known. The latter can be estimated consistently by the OLS estimator obtained by regressing $\{y_{iT}\}$ on $\{y_{i1}\}$.But the resulting MLEs of α and σ^2_{u} have a singular information matrix and their finite sample properties are worse than those of this section according to our experiments.

reason. When maximizing $l(\alpha, \sigma_{\rm m}^2)$ with respect to the two parameters α and $\sigma_{\rm m}^2$, some complications arise at $\alpha_o = 1$, where α_o denote the true value of α, because for any values of $\sigma_{\rm m}^2$

$$\frac{\partial l}{\partial \sigma_{\rm m}^2} = \frac{\omega_{11.2}^{dm}}{2\omega_{11.2}^2} \sum_{i=1}^{N} \left(r_{iT}^2 - \omega_{11.2} \right) - \frac{1}{\omega_{11.2}} \sum_{i=1}^{N} r_{iT} r_{iT}^{dm} = 0, \tag{12}$$

where $\omega_{11.2}^{dm} = (1-\alpha)^2 \sum_{j=0}^{T-2} \alpha^{2j} + (1-\alpha^{T-1})^2 - 2(1-\alpha^{T-1})\omega_{12}/\omega_{22}$ and $r_{iT}^{dm} = -(1-\alpha^{T-1}) y_{i1}/\omega_{22}$. Relation (12) implies that the objective function drawn in a three-dimensional space has a tunnel-like shape which makes it impossible to estimate $\sigma_{\rm m}^2$ consistently at $\alpha_o = 1$. Figure 1 plots the likelihood function of α and $\sigma_{\rm m}^2$ at $\alpha_o = 1$, which confirms our conjecture. In practice, $\sigma_{\rm m}^2$ can be estimated consistently by the MLE using the likelihood function (11).

Let $\hat{\alpha}_{CSMLE}$ denote the CSMLE of α. The $\hat{\alpha}_{CSMLE}$ is obtained by maximizing $l(\alpha, \sigma_{\rm m}^2)$ with a known value of $\sigma_{\rm m}^2$. The MLE of α is consistent by the standard theory of MLEs (see, e.g., Newey and McFadden, 1994). In Figure 2, we plotted the likelihood function of α for various values of α_o. We find from this figure that the function is well behaved and maximized near α_o.

The limiting distribution of $\hat{\alpha}_{CSMLE}$ is reported in the following theorem.

Theorem 1 *Suppose that Assumptions 1 and 2 hold. Assume that $\alpha \in \Theta$, where Θ is any compact subset of $\mathbb{R}$ and that $\alpha_o \in interior(\Theta)$. Then,*

$$\sqrt{N}(\hat{\alpha}_{CSMLE} - \alpha_o) \xrightarrow{d} \boldsymbol{N} \left(0, \left[\frac{\left(\omega_{11.2}^{d\alpha}\right)^2}{2\left(\omega_{11.2}\right)^2} + \frac{(T-1)^2 \alpha_o^{2T-4} (\omega_{22} - \sigma_{\rm mo}^2)^2}{\omega_{11.2}\omega_{22}} \right]^{-1} \right),$$

where

$$\begin{aligned} \omega_{11.2}^{d\alpha} \;=\;& 2(1-\alpha_o)(\omega_{22} - \sigma_{\rm mo}^2) \sum_{j=0}^{T-2} \alpha_o^{2j} + 2\left\{ \lambda - (1-\alpha_o)^2 (\omega_{22} - \sigma_{\rm mo}^2) \right\} \sum_{j=0}^{T-2} j \alpha_o^{2j-1} \\ & -2(1-\alpha_o^{T-1})(T-1)\alpha_o^{T-2}\sigma_{\rm mo}^2 + 2(T-1)\alpha_o^{T-2}\sigma_{\rm mo}^2 \omega_{12}/\omega_{22}. \end{aligned}$$

Remarkably, this theorem holds for any value of α_o, which makes it possible to construct confidence intervals for α_o[9] and formulate t-ratios using the standard error

[9] Constructing confidence intervals of the AR(1) coefficient for time series and panel data is a

of $\hat{\alpha}_{CSMLE}$

$$se = \sqrt{\frac{1}{\left[\frac{\left(\hat{\omega}_{11.2}^{d\alpha}\right)^2}{2(\hat{\omega}_{11.2})^2} + \frac{(T-1)^2\hat{\alpha}_{CSMLE}^{2T-4}(\hat{\omega}_{22}-\hat{\sigma}_{\mathrm{mo}}^2)^2}{\hat{\omega}_{11.2}\hat{\omega}_{22}}\right]}}, \tag{13}$$

where $\hat{a}$ denotes a consistent estimator of a. The t-ratio

$$(\hat{\alpha}_{CSMLE} - 1)/se$$

can be used as a test statistic for a panel unit root against the alternatives of both stable and explosive AR(1) coefficients. We do not pursue it further here in the presence of many, existing panel unit root tests (see Chapter 7 of Choi (2015) for panel unit root tests).

Note that $\hat{\alpha}_{CSMLE}$ is computed in two steps. First, estimate $\sigma_{\mathrm{m}o}^2$ using the full likelihood function (11). Next, plug the MLE of $\sigma_{\mathrm{m}o}^2$ and the method-of moments estimators of σ_{u}^2 and $\sigma_{\mathrm{x}_1}^2$ into the full likelihood function (11) and maximize it with respect to α.

4 Indirect inference estimation

The MLE is an efficient estimator, but is sometimes subject to bias (cf. Cox and Snell, 1968). Bias-corrections for the MLE of the last section deserve a serious consideration for this reason. There have been various methods for bias-corrections for dynamic panels (e.g., Kiviet, 1995; Bun and Carree, 2005; Hahn and Kuersteiner, 2002; Gouriéroux, Phillips and Yu, 2010). In this section, we apply the principle of indirect inference estimation originally proposed by Smith (1993) and Gouriéroux, Monfort and Renault (1993) to model (10) in conjunction with the CSMLE of the last section. A distinct advantage of indirect inference is that it does not require explicit calculation of bias terms, which is quite complicated to obtain for CSMLE. Resampling methods such as jackknife and bootstrapping have the same advantage, but are not considered here. Median-unbiased estimation also eliminates finite sample

nontrivial task because most estimators undergo changes in their limiting distributions as the value of the AR(1) coefficient changes (see Chapter 5 of Choi, 2015).

bias as shown in Andrews (2003), but it is not pursed here. Gouriéroux, Phillips and Yu (2010) also employ the indirect inference estimation method for dynamic panels and report good finite sample properties of their estimator. But their model is (4) and the parameter space of the PAR(1) coefficient is restricted to (-1,1).

Suppose that we can generate simulated data by using model (10) once the value of α is given, and denote the h-th simulated data as $\{\mathbf{y}_1^h(\alpha), ..., \ \mathbf{y}_T^h(\alpha)\},$ where $\mathbf{y}_t^h(\alpha)$ denotes the vector of $\{y_{it}\}_{i=1,...,N}$ at given α. The details of data generation will be discussed later. Let $\hat{\alpha}_{CSMLE}^h(\alpha)$ be the CSMLE of α using the h-th simulated data $\{\mathbf{y}_1^h(\alpha), ..., \ \mathbf{y}_T^h(\alpha)\}$. The indirect inference estimator of α_o is defined by

$$\hat{\alpha}_{IIE} = \arg\min_{\alpha \in \Lambda} \left\| \hat{\alpha}_{CSMLE} - \frac{1}{H} \sum_{h=1}^{H} \hat{\alpha}_{CSMLE}^h(\alpha) \right\|,$$

where Λ, the parameter space of α, is a compact subset of $\mathbb{R}$ and $\|\cdot\| : \mathbb{R} \to [0, \infty)$ is a norm. In practice, the IIE can be obtained by solving the equation

$$\hat{\alpha}_{CSMLE} - \frac{1}{H} \sum_{h=1}^{H} \hat{\alpha}_{CSMLE}^h(\alpha) = 0 \tag{14}$$

for α. Letting $b_N^H(\alpha) = \frac{1}{H} \sum_{h=1}^{H} \hat{\alpha}_{CSMLE}^h(\alpha)$, assume

Assumption 3 $b_N^H(\alpha) : \Lambda \to \Lambda$ *is an one-to-one, continuous function of* α *for all* N *and* H.

A unique solution of equation (14) exists under this assumption. That is, $\hat{\alpha}_{IIE}^H = b_N^{H-1}(\hat{\alpha}_{CSMLE})$.

How can we know that the IIE alleviates the finite-sample bias problem of $\hat{\alpha}_{CSMLE}$ and what is its limiting distribution? To tackle these questions, assume $H = \infty$. Then, we have by the law of large numbers

$$b_N^\infty(\alpha) = E(\hat{\alpha}_{CSMLE}^1(\alpha)).$$

In addition, it follows that

$$b_N^\infty(\alpha_o) = E(\hat{\alpha}_{CSMLE}^1(\alpha_o)) = E(\hat{\alpha}_{CSMLE}) = E(b_N^\infty(\hat{\alpha}_{IIE}^\infty)),$$

where the last equality holds by the definition of $b_N^\infty(\cdot)$. The standard theory of MLE (cf. Cox and Snell, 1968) yields

$$E(\hat{\alpha}_{CSMLE}) = \alpha_o + N^{-1} c(\alpha_o).$$

Thus, is is reasonable to assume

Assumption 4 $b_N^{\infty}(\alpha) = \alpha + N^{-1}c(\alpha)$, *where* $c(\alpha) : \Lambda \rightarrow R$ *is a continuously differentiable function of* α.

Properties of $\hat{\alpha}_{IIE}^{\infty}$ are reported in the following theorem.

Theorem 2 *Suppose that Assumptions 1-4 hold. Then,*
(i) $\hat{\alpha}_{IIE}^{\infty} \xrightarrow{p} \alpha_o$.
(ii) $\sqrt{N}(\hat{\alpha}_{IIE}^{\infty} - \alpha_o)$ *has the same asymptotic distribution as* $\sqrt{N}(\hat{\alpha}_{CSMLE} - \alpha_o)$.
(iii) $E(\hat{\alpha}_{IIE}^{\infty}) = \alpha_o + E(\zeta_N)$, *where* $\zeta_N = O_p(N^{-3/2})$.

This theorem shows that $\hat{\alpha}_{IIE}^{\infty}$ is consistent for α_o and has the same asymptotic distribution as $\hat{\alpha}_{CSMLE}$. Furthermore, while the bias of CSMLE is $O(N^{-1})$, that of $\hat{\alpha}_{IIE}^{\infty}$ is the mean of an $O_p(N^{-3/2})$ random variable. We conjecture from this that $\hat{\alpha}_{IIE}^{\infty}$ is less biased than $\hat{\alpha}_{CSMLE}$, which is confirmed by simulation in Section 7.

Now, we discuss how $\hat{\alpha}_{CSMLE}^{h}(\alpha)$ is calculated using simulated data $\{\mathbf{y}_1^h(\alpha), ..., \mathbf{y}_T^h(\alpha)\}$.
Step 1: Estimate σ_{m}^2, $\sigma_{\mathrm{x}_1}^2$ and σ_{u}^2 using the observed data.
Step 2: Using a normal number generator along with σ_{m}^2, $\sigma_{\mathrm{x}_1}^2$ and σ_{u}^2 from Step 1, generate $\{\mathbf{y}_t^h(\alpha)\}_{t=1,..,T}$ for a fixed value of α.
Step 3: Calculate $\hat{\lambda}^h(\alpha)$, an estimate of λ given α, using the simulated data $\{\mathbf{y}_1^h(\alpha),\ \mathbf{y}_2^h(\alpha)\}$ from *Step 2*.
Step 4: Calculate $\hat{\alpha}_{CSMLE}^{h}(\alpha)$ using the simulated data $\{\mathbf{y}_1^h(\alpha),\ \mathbf{y}_T^h(\alpha)\}$ from *Step 2* and the estimate of $\lambda^h(\alpha)$ from *Step 3*.

5 Bias-corrected pooled least squares estimator

This section proposes a bias-corrected pooled least squares estimator (BCPLSE) of α for Model (4). We are interested in the PLSE rather than the least squares dummy variable estimator because time-invariant regressors are not eliminated by the PLSE procedure when it is extended to the model with time-invariant regressors.

The PLSE of α is defined as $\hat{\alpha}_{PLSE} = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1}-\bar{y}_{-1})y_{it}}{\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1}-\bar{y}_{-1})^2}$, where $\bar{y}_{-1} =$

$\frac{1}{N(T-1)}\sum_{i=1}^{N}\sum_{t=2}^{T}y_{i,t-1}$. Because

$$\begin{aligned}
& \hat{\alpha}_{PLSE}-\alpha \\
= & \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1}-\bar{y}_{-1})\left(u_{it}+m_i(1-\alpha)\right)}{\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1}-\bar{y}_{-1})^2} \\
= & \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\left\{(x_{i,t-1}-\bar{x}_{-1}+m_i-\bar{m})u_{it}+(x_{i,t-1}-\bar{x}_{-1})m_i(1-\alpha)\right\}}{\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1}-\bar{y}_{-1})^2} \\
& +\frac{(1-\alpha)(T-1)\sum_{i=1}^{N}(m_i-\bar{m})m_i}{\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1}-\bar{y}_{-1})^2}, \qquad (15)
\end{aligned}$$

we have under Assumption 1,

$$\hat{\alpha}_{PLSE}-\alpha\xrightarrow{p}\frac{(1-\alpha)(T-1)\sigma_{\mathrm{m}}^2}{p\lim\frac{1}{N}\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1}-\bar{y}_{-1})^2}.$$

This means that $\hat{\alpha}_{PLSE}$ is inconsistent for α unless $\alpha=1$. The BCPLSE using the CSMLE is defined by

$$\hat{\alpha}_{PLSE}-\frac{(1-\hat{\alpha}_{CSMLE})(T-1)\hat{\sigma}_{\mathrm{m}CSMLE}^2}{\frac{1}{N}\sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1}-\bar{y}_{-1})^2}. \qquad (16)$$

Asymptotic properties of the bias-corrected PLSE are reported in the following theorem.

Theorem 3 *Suppose that assumptions for Theorem 1 hold. Then, as* $N\to\infty$,
(i) $\hat{\alpha}_{BCPLSE}\xrightarrow{p}\alpha$;
(ii) $\sqrt{N}(\hat{\alpha}_{BCPLSE}-\alpha)=O_p(1)$.

In practice, we may use BCPLSE to construct a two-stage BCPLSE using the formula (16) with $\hat{\alpha}_{CSMLE}$ being replaced by $\hat{\alpha}_{BCPLSE}$. The two-stage BCPLSE performs better than BCPLSE according to some unreported simulation results.

The limiting distribution of $\hat{\alpha}_{BCPLSE}$ depends on that of a linear combination of $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\left(m_i^2-\sigma_{\mathrm{m}}^2\right)$, $\sqrt{N}\left(\hat{\sigma}_{\mathrm{m}CSMLE}^2-\sigma_{\mathrm{m}}^2\right)$ and $\sqrt{N}(\hat{\alpha}_{CSMLE}-\alpha)$. Because these are not independent, it is difficult to derive the limiting distribution of $\hat{\alpha}_{BCPLSE}$. Under this circumstance, we can use bootstrapping for interval estimation of α. The interval estimation can also be used for point hypothesis testing in the usual way. The following steps provide bootstrap confidence intervals of α.

Step 1: Let $\mathbf{y}_i = [y_{i1}, ..., y_{iT}]'$. Choose $\mathbf{y}_1^*, ..., \mathbf{y}_N^*$ randomly from $\{\mathbf{y}_i\}_{i=1,...,N}$ with replacements.

Step 2: Calculate BCPLSE of α using $\{\mathbf{y}_i^*\}_{i=1,...,N}$, which is denoted as $\hat{\alpha}^*_{BCPLSE,b}$.

Step 3: Repeat Steps 1 and 2 B times and record the values of $\{\sqrt{N}(\hat{\alpha}^*_{BCPLSE,b} - \hat{\alpha}_{BCPLSE})\}_{b=1,...,B}$.

Step 4: Obtain the $\gamma/2$-th and $(1-\gamma/2)$-th percentiles of $\{\sqrt{N}(\hat{\alpha}^*_{BCPLSE,b} - \hat{\alpha}_{BCPLSE})\}_{b=1,...,B}$, which are denoted as $c_{\gamma/2}$ and $c_{(1-\gamma/2)}$.

The $(1-\gamma)\times 100$ percent bootstrap confidence interval for α is defined as $(\hat{\alpha}_{BCPLSE} - c_{\gamma/2}/\sqrt{N},\ \hat{\alpha}_{BCPLSE} - c_{(1-\gamma/2)}/\sqrt{N})$. Finite sample properties of the bootstrap confidence intervals will be studied in Section 7.

Alternatively, one may resample the data using time series residuals for each i as in the time series literature. According to some experiments the results of which are unreported here, the bootstrap procedure given above provides far better results in finite samples. A plausible reason for this is that the resampled data using the steps above mimic the original data better than those based on resampled residuals.

6 An extension to the dynamic AR(1) model with endogenous regressors

6.1 The model and assumptions

An extended version of Model (1) is

$$y_{it} = \mu_i + \gamma' p_i + x_{it},\ (i = 1, \ldots, N;\ t = 1, \ldots, T), \tag{17}$$

where $\{p_i\}$ is a sequence of observed, time-invariant variables of dimension l_p and $\{x_{it}\}$ is the same as in equation (2). Model (17) yields the conventional dynamic panel data model

$$\begin{aligned} y_{it} &= \mu_i(1-\alpha) + (1-\alpha)\gamma' p_i + \alpha y_{i,t-1} + u_{it} \\ &= \mu(1-\alpha) + (1-\alpha)\gamma' p_i + \alpha y_{i,t-1} + u_{it} + m_i(1-\alpha). \end{aligned} \tag{18}$$

We assume for Model (17)

Assumption 5 *Assume*

$$\begin{pmatrix} p_i \\ m_i \\ u_i \\ x_{i1} \end{pmatrix} \sim \text{i. i. d.} \left(\mathbf{0}, \begin{bmatrix} \phi_{pp} & \phi_{pm} & 0 & 0 \\ \phi'_{pm} & \sigma^2_{\mathrm{m}} & 0 & 0 \\ 0 & 0 & \sigma^2_{\mathrm{u}} I_{T-1} & 0 \\ 0 & 0 & 0 & \sigma^2_{\mathrm{x}_1} \end{bmatrix} \right),$$

where $\sigma^2_{\mathrm{m}} > 0,\ \sigma^2_{\mathrm{u}} > 0,\ \sigma^2_{\mathrm{x}_1} > 0$ *and* $\Psi = \begin{bmatrix} \phi_{pp} & \phi_{pm} \\ \phi'_{pm} & \sigma^2_{\mathrm{m}} \end{bmatrix} > 0.$

This assumption implies that all the regressors of Model (18) are endogenous when $\phi_{pm} \neq 0$ in the sense that they are correlated with the error terms. This feature makes OLS or Within-OLS unusable for Model (18).

6.2 CSMLE

Now, we consider CSMLE for Model (17) as in Section 3. Assume $\mu = 0$ for simplicity from now on. Model (17) can be written as

$$y_{it} = \gamma' p_i + \alpha^{t-1} x_{i1} + w_{it} + m_i,\ (t = 2, ..., T),$$

where $w_{it} = \sum_{j=0}^{t-2} \alpha^j u_{i,t-j}$. Because $y_{i1} = \mu_i + \gamma' p_i + x_{i1}$, we have for $t = T$

$$\begin{aligned} y_{iT} &= \gamma' p_i + \alpha^{T-1}(y_{i1} - \mu_i - \gamma' p_i) + w_{iT} + m_i \\ &= (1 - \alpha^{T-1})\gamma' p_i + \alpha^{T-1} y_{i1} + w_{iT} + (1 - \alpha^{T-1}) m_i. \end{aligned} \tag{19}$$

All the regressors $\{p_i\}$ and $\{y_{i1}\}$ are endogenous if $\phi_{pm} \neq 0$ and $\alpha \neq 1$.

Let $c_{iT} = w_{iT} + (1 - \alpha^{T-1}) m_i$. Then,

$$\begin{pmatrix} c_{iT} \\ p_i \\ y_{i1} \end{pmatrix} \sim \text{i. i. d.} \left(0,\ \Delta\right), \tag{20}$$

where $\Delta = \begin{bmatrix} \delta_{11} & \delta'_{21} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{13} & \delta'_{23} & \delta_{33} \end{bmatrix}$ and

$$\begin{aligned} \delta_{11} &= \sigma_{\mathrm{u}}^2 \sum_{j=0}^{T-2} \alpha^{2j} + (1-\alpha^{T-1})^2 \sigma_{\mathrm{m}}^2; \\ \delta_{21} &= (1-\alpha^{T-1})\phi_{pm};\ \delta_{22} = \phi_{pp}; \\ \delta_{13} &= (1-\alpha^{T-1})\sigma_{\mathrm{m}}^2 + (1-\alpha^{T-1})\gamma'\phi_{pm}; \\ \delta_{23} &= \phi_{pm} + \phi_{pp}\gamma; \\ \delta_{33} &= \sigma_{\mathrm{m}}^2 + \gamma'\phi_{pp}\gamma + +\sigma_{\mathrm{x}_1}^2 + 2\gamma'\phi_{pm}. \end{aligned}$$

Let $\Delta = \begin{bmatrix} \delta_{11} & \delta'_{*1} \\ \delta_{*1} & \delta_{**} \end{bmatrix}$, $q_i = \begin{pmatrix} p'_i & y_{i1} \end{pmatrix}'$ and $\Delta_{11.*} = \delta_{11} - \delta'_{*1}\delta_{**}^{-1}\delta_{*1}$. As in Section 3, write

$$c_{iT} = \delta'_{*1}\delta_{**}^{-1}q_i + r_{iT}$$

and assume

Assumption 6

$$r_{iT} \mid q_i \sim \text{i. i. d.}\, \boldsymbol{N}\left(0,\ \Delta_{11.*}\right).$$

This gives

$$\begin{aligned} y_{iT} \quad &\mid \quad q_i \\ &\sim \quad \text{i. i. d.}\ \boldsymbol{N}\big((1-\alpha^{T-1})\gamma' p_i + \alpha^{T-1} y_{i1} + \delta'_{*1}\delta_{**}^{-1} q_i, \Delta_{11.*}). \end{aligned} \tag{21}$$

Let $s_{iT} = y_{iT} - (1-\alpha^{T-1})\gamma' p_i - \alpha^{T-1} y_{i1} - \delta'_{*1}\delta_{**}^{-1} q_i$. Then, s_{iT} and q_i are independent for every i and T. Furthermore, let the probability density function of q_i be denoted by $f(\cdot \mid \delta_{**})$. The likelihood function of the parameters $\alpha,\ \gamma,\ \beta,\ \sigma_{\mathrm{u}}^2,\ \sigma_{\mathrm{m}}^2,\ \sigma_{\mathrm{x}_1}^2$ are obtained from relation (21) as

$$\begin{aligned} L(\alpha, \gamma, \sigma_{\mathrm{u}}^2, \phi_{pm}, \sigma_{\mathrm{m}}^2, \sigma_{\mathrm{x}_1}^2 \quad &\mid \quad y_{1T}, ..., y_{NT}, q_1, ..., q_N) \\ &= \quad \Pi_{i=1}^{N} \frac{1}{\sqrt{2\pi\Delta_{11.*}}} \exp\left(-\frac{s_{iT}^2}{2\Delta_{11.*}}\right) \\ &\quad\ \times \Pi_{i=1}^{N} f(q_i \mid \delta_{**}). \end{aligned}$$

Assuming that ϕ_{pp} and δ_{**} are known[10], the log-likelihood function is written as

$$l(\alpha, \gamma, \sigma_{\mathrm{u}}^2, \phi_{pm}, \sigma_{\mathrm{m}}^2, \sigma_{\mathrm{x}_1}^2) = c - \frac{N}{2} \ln(\delta_{11} - \delta_{*1}' \delta_{**}^{-1} \delta_{*1}) - \frac{1}{2} \sum_{i=1}^{N} \frac{s_{iT}^2}{\Delta_{11.*}}, \qquad (22)$$

where c is a constant.

The CSMLEs of the unknown parameters are obtained by maximizing the log-likelihood function (22). To reduce the number of parameters to estimate as in Section 3, assume δ_{33}, δ_{23} and $\lambda = Var(y_{i2} - y_{i1}) = \sigma_{\mathrm{u}}^2 + (1-\alpha)^2 \sigma_{\mathrm{x}_1}^2$ are known. Then, because $\sigma_{\mathrm{u}}^2 = \lambda - (1-\alpha)^2 \sigma_{\mathrm{x}_1}^2$, $\phi_{pm} = \delta_{23} - \phi_{pp}\gamma$ and $\sigma_{\mathrm{x}_1}^2 = \delta_{33} - \sigma_{\mathrm{m}}^2 + \gamma' \phi_{pp} \gamma - 2\gamma' \delta_{23}$, only α, γ and σ_{m}^2 need to be estimated by maximizing the log-likelihood function. Since we have

$$\frac{\partial l}{\partial \sigma_{\mathrm{m}}^2} = \left(-\frac{1}{2}\right) \left[\frac{N \Delta_{11.*}^{dm}}{\Delta_{11.*}} + \frac{2 \sum_{i=1}^{N} s_{iT}^{dm} s_{iT}}{\Delta_{11.*}} - \frac{\Delta_{11.*}^{dm} \sum_{i=1}^{N} s_{iT}^2}{(\Delta_{11.*})^2} \right],$$

where $\Delta_{11.*}^{dm} = (1-\alpha)^2 \sum_{j=0}^{T-2} \alpha^{2j} + (1-\alpha^{T-1})^2 - 2 \begin{bmatrix} 0 & (1-\alpha^{T-1}) \end{bmatrix} \delta_{**}^{-1} \delta_{*1}$ and $s_{iT}^{dm} = - \begin{bmatrix} 0 & (1-\alpha^{T-1}) \end{bmatrix} \delta_{**}^{-1} q_i$, $\frac{\partial l}{\partial \sigma_{\mathrm{m}}^2} = 0$ at $\alpha_o = 1$ as in Section 3. Thus, we assume that σ_{m}^2 is known. In practice, σ_{m}^2 can be estimated using the original log-likelihood function.

The limiting distribution of $\hat{\alpha}_{CSMLE}$ is reported in the following theorem.

Theorem 4 *Suppose that Assumptions 5 and 6 hold. Assume that* $(\alpha, \gamma')' \in \Theta$, *where* Θ *is any compact subset of* $\mathbb{R}^{l_p+1}$ *and that* $(\alpha_o, \gamma_o')' \in interior(\Theta)$. *Then,*

$$\sqrt{N} \begin{pmatrix} \hat{\alpha}_{CSMLE} - \alpha_o \\ \hat{\gamma}_{CSMLE} - \gamma_o \end{pmatrix} \xrightarrow{d} \mathbf{N} \left(0, \begin{bmatrix} \frac{\left(\Delta_{11.*}^{d\alpha}\right)^2}{2(\Delta_{11.*})^2} + \frac{c_\alpha' \delta_{**} c_\alpha}{\Delta_{11.*}} & \left(\frac{\Delta_{11.*}^{d\alpha} \Delta_{11.*}^{d\gamma}}{2(\Delta_{11.*})^2} + \frac{B_\gamma' \delta_{**} c_\alpha}{\Delta_{11.*}} \right)' \\ \frac{\Delta_{11.*}^{d\alpha} \Delta_{11.*}^{d\gamma}}{2(\Delta_{11.*})^2} + \frac{B_\gamma' \delta_{**} c_\alpha}{\Delta_{11.*}} & \frac{(\Delta_{11.*}^{d\gamma})(\Delta_{11.*}^{d\gamma})'}{2(\Delta_{11.*})^2} + \frac{B_\gamma' \delta_{**} B_\gamma}{\Delta_{11.*}} \end{bmatrix}^{-1} \right),$$

[10] In practice, ϕ_{pp} and δ_{**} can be estimated consistently using the observations $\{p_i\}$ and $\{(p_i' \;\; y_{i1})\}$, respectively.

where

$$\Delta^{d\alpha}_{11.*} = 2\{(1-\alpha_o)(\delta_{33}-\sigma^2_{\text{mo}}+\gamma_o'\phi_{pp}\gamma_o-2\gamma_o'\delta_{23})\sum_{j=0}^{T-2}\alpha_o^{2j}$$

$$+\left(\lambda-(1-\alpha_o)^2(\delta_{33}-\sigma^2_{\text{mo}}+\gamma_o'\phi_{pp}\gamma_o-2\gamma_o'\delta_{23})\right)\sum_{j=0}^{T-2}j\alpha_o^{2j-1}-(T-1)\alpha_o^{T-2}(1-\alpha_o^{T-1})\sigma^2_{\text{mo}}\}$$

$$-2\delta_{*1}'\delta_{**}^{-1}\begin{bmatrix} -(T-1)\alpha^{T-2}(\delta_{23}-\phi_{pp}\gamma_o) \\ -(T-1)\alpha^{T-2}\sigma^2_{\text{m}}-(T-1)\alpha^{T-2}(\delta_{23}'-\gamma_o'\phi_{pp})\gamma_o \end{bmatrix};$$

$$c_\alpha = (T-1)\alpha^{T-2}\begin{bmatrix} \gamma_o \\ -1 \end{bmatrix} - \delta_{**}^{-1}\begin{bmatrix} -(T-1)\alpha^{T-2}(\delta_{23}-\phi_{pp}\gamma_o) \\ -(T-1)\alpha^{T-2}\sigma^2_{\text{m}}-(T-1)\alpha^{T-2}(\delta_{23}'-\gamma_o'\phi_{pp})\gamma_o \end{bmatrix};$$

$$\Delta^{d\gamma}_{11.*} = -(1-\alpha_o)^2(2\phi_{pp}\gamma_o-2\delta_{23})\sum_{j=0}^{T-2}\alpha_o^{2j}$$

$$-2\begin{bmatrix} -(1-\alpha_o^{T-1})\phi_{pp} & (1-\alpha_o^{T-1})\delta_{23}-2(1-\alpha_o^{T-1})\phi_{pp}\gamma_o \end{bmatrix}\delta_{**}^{-1}\delta_{*1};$$

$$B_\gamma = -(1-\alpha^{T-1})\begin{bmatrix} I_{l_p} \\ 0 \end{bmatrix} - \delta_{**}^{-1}\begin{bmatrix} -(1-\alpha_o^{T-1})\phi_{pp} \\ (1-\alpha_o^{T-1})\delta_{23}'-2(1-\alpha_o^{T-1})\gamma_o'\phi_{pp} \end{bmatrix};$$

These results indicate that the coefficients of time-invariant endogenous regressors can be estimated consistently. This is not possible for conventional methods that rely on differencing. Adding time-variant endogenous variables to Model (18) and deriving similar results are also possible. This requires assumptions on the dynamics of time-variant regressors and correlations between the regressors and individual effect variables $\{\mu_i\}$.

6.3 Indirect inference estimation

The indirect inference estimation for Model (19) proceeds in the same manner as in Section 4. Suppose that we can generate simulated data by using model (19) once the value of α and γ are given, and denote the h-th simulated data as $\{\mathbf{y}_T^h(\alpha,\gamma),\ \mathbf{y}_1^h(\alpha,\gamma)\}$, where $\mathbf{y}_T^h(\alpha,\gamma)$ and $\mathbf{y}_1^h(\alpha,\gamma)$ are vectors of the regressand and regressor, respectively. Let $\theta^h_{CSMLE}(\alpha,\gamma)=\begin{pmatrix}\hat{\alpha}^h_{CSMLE}\\ \hat{\gamma}^h_{CSMLE}\end{pmatrix}$ be the CSMLE of $\theta_o=\begin{pmatrix}\alpha_o\\ \gamma_o\end{pmatrix}$ using the h-th simulated data $\{\mathbf{y}_1^h(\alpha,\gamma),...,\ \mathbf{y}_T^h(\alpha,\gamma)\}$. The indirect inference estimator (IIE) of θ_o

is defined by

$$\hat{\theta}_{IIE} = \arg\min_{\theta\in\Lambda} \left\| \hat{\theta}_{CSMLE} - \frac{1}{H}\sum_{h=1}^{H} \theta^{h}_{CSMLE}(\alpha,\gamma) \right\|,$$

where Λ is a compact subset of $\mathbb{R}^{l_p+1}$ and $\|\cdot\| : \mathbb{R}^{l_p+1} \to [0,\infty)$ is the Euclidian norm. Letting $b^H_N(\alpha,\gamma) = \frac{1}{H}\sum_{h=1}^{H}\theta^h_{CSMLE}(\alpha,\gamma)$, assume as in Section 4

Assumption 7 $b^H_N(\alpha,\gamma) : \Lambda \to \Lambda$ *is an one-to-one, continuous function of* α *and* γ *for all* N *and* H.

Assumption 8 $b^\infty_N(\alpha,\gamma) = \theta + N^{-1}c(\theta)$, *where* $c(\theta) : \Lambda \to \mathbb{R}^{l_p+1}$ *is a continuously differentiable function of* θ.

Now, we have

Theorem 5 *Suppose that Assumptions 5, 6, 7 and 8 hold. Then,*
(i) $\hat{\theta}^\infty_{IIE} \xrightarrow{p} \theta_o$.
(ii) $\sqrt{N}(\hat{\theta}^\infty_{IIE} - \theta_o)$ *has the same asymptotic distribution as* $\sqrt{N}(\hat{\theta}_{CSMLE} - \theta_o)$.
(iii) $E(\hat{\theta}^\infty_{IIE}) = \theta_o + E(\zeta_N)$, *where* $\zeta_N = O_p(N^{-3/2})$.

6.4 Bias-corrected pooled least squares estimator

Let $\psi = [\gamma'_\alpha,\ \alpha]'$ with $\gamma_\alpha = (1-\alpha)\gamma$. Running OLS using Model (18), we obtain the PLSE of ψ as

$$\hat{\psi}_{PLSE} = \begin{bmatrix} \sum_{i=1}^N\sum_{t=2}^T (p_i-\bar{p})^2 & \sum_{i=1}^N\sum_{t=2}^T (p_i-\bar{p})(y_{i,t-1}-\bar{y}_{-1}) \\ \sum_{i=1}^N\sum_{t=2}^T (p_i-\bar{p})(y_{i,t-1}-\bar{y}_{-1}) & \sum_{i=1}^N\sum_{t=2}^T (y_{i,t-1}-\bar{y}_{-1})^2 \end{bmatrix}^{-1} \times \begin{bmatrix} \sum_{i=1}^N\sum_{t=2}^T (p_i-\bar{p})\, y_{it} \\ \sum_{i=1}^N\sum_{t=2}^T (y_{i,t-1}-\bar{y}_{-1}) y_{it} \end{bmatrix}.$$

Under Assumption 5, we obtain

$$\hat{\psi}_{PLSE} - \psi \xrightarrow{p} \begin{bmatrix} (T-1)\phi_{pp} & (T-1)(\phi_{pp}\gamma + \phi_{pm}) \\ (T-1)(\gamma'\phi_{pp} + \phi'_{pm}) & (T-1)(\sigma^2_{\mathrm{m}} + \gamma'\phi_{pp}\gamma + 2\gamma'\phi_{pm}) + \kappa \end{bmatrix}^{-1} \times \begin{bmatrix} (1-\alpha)(T-1)\phi_{pm} \\ (1-\alpha)(T-1)(\sigma^2_{\mathrm{m}} + \gamma'\phi_{pm}) \end{bmatrix},$$

where $\kappa = E\left(\sum_{t=2}^{T}(x_{i,t-1} - \bar{x}_{-1})^2\right)$. This shows that $\hat{\psi}_{PLSE}$ is inconsistent unless $\alpha = 1$. Thus, the BCPLSE of ψ is defined as

$$\begin{aligned}\hat{\psi}_{BCPLSE} &= \hat{\psi}_{PLSE} - \\ &\left[\frac{1}{N}\begin{pmatrix} \sum_{i=1}^{N}\sum_{t=2}^{T}(p_i - \bar{p})^2 & \sum_{i=1}^{N}\sum_{t=2}^{T}(p_i - \bar{p})(y_{i,t-1} - \bar{y}_{-1}) \\ \sum_{i=1}^{N}\sum_{t=2}^{T}(p_i - \bar{p})(y_{i,t-1} - \bar{y}_{-1}) & \sum_{i=1}^{N}\sum_{t=2}^{T}(y_{i,t-1} - \bar{y}_{-1})^2 \end{pmatrix}\right]^{-1} \\ &\times \begin{bmatrix} (1 - \hat{\alpha}_{CSMLE})(T-1)\hat{\phi}_{pmCSMLE} \\ (1 - \hat{\alpha}_{CSMLE})(T-1)(\hat{\sigma}^2_{\mathrm{m}CSMLE} + \hat{\gamma}'_{CSMLE}\hat{\phi}_{pmCSMLE}) \end{bmatrix},\end{aligned}$$

where $\hat{\alpha}_{CSMLE}$, $\hat{\gamma}_{CSMLE}$, $\hat{\phi}_{pmCSMLE}$ and $\hat{\sigma}^2_{\mathrm{m}CSMLE}$ are the CSMLEs of the previous subsection.

Let $\hat{\psi}_{BCPLSE}$ be partitioned as $[\hat{\psi}^{\gamma\prime}_{BCPLSE},\ \hat{\psi}^{\alpha}_{BCPLSE}]'$, where $\hat{\psi}^{\gamma}_{BCPLSE}$ is of dimension l_p, and let $\hat{\gamma}_{BCPLSE} = \hat{\psi}^{\gamma}_{BCPLSE}/(1 - \hat{\psi}^{\alpha}_{BCPLSE})$. Then, we obtain:

Theorem 6 *Under Assumptions 5 and 6,*

(i) $\hat{\psi}^{\alpha}_{BCPLSE} \xrightarrow{p} \alpha$ *and* $\hat{\psi}^{\alpha}_{BCPLSE} = O_p(1)$;

(ii) Assume $\alpha \neq 1$. *Then,* $\hat{\gamma}_{BCPLSE} \xrightarrow{p} \gamma$ *and* $\sqrt{N}(\hat{\gamma}_{BCPLSE} - \gamma) = O_p(1)$.

Part (i) of this theorem shows that $\hat{\psi}^{\alpha}_{BCPLSE}$ estimates α consistently and that $\sqrt{N}$-asymptotics applies to it. The same holds for $\hat{\gamma}_{BCPLSE}$ unless $\alpha = 1$. When $\alpha = 1$, $\hat{\gamma}_{BCPLSE}$ diverges in probability. The bootstrap method of Section 5 can also be used here.

7 Simulation

This section reports simulation results for the following estimators: CSMLE, IIE and BCPLSE of this paper, GMM estimators of Arellano and Bond (1991), Ahn and Schmidt (1995) and Blundell and Bond (1998), Han and Phillips' (2010) first difference least squares estimator and Han, Phillips and Sul's (2014) panel fully aggregated estimator. These estimators are explained briefly in the next subsection in relation to Model (1) and Assumption 1.

7.1 Estimators for the PAR(1) model

7.1.1 GMM estimators

Arellano and Bond (1991) starts from Model (4) without the intercept term. Differencing Model (3) gives

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta u_{it},\ (t = 3, ..., T).$$

Under Assumption 1, we have

$$E(y_{i1} - \mu)u_{it} = 0 \text{ for } t \geq 2$$

and

$$E(m_i u_{it}) = 0 \text{ for every } i \text{ and } t,$$

which provides the $(T-2)(T-1)/2$ moment conditions

$$E\left(\Delta u_{it}(y_{i,t-j} - \mu)\right) = 0 \ (j = 2, ..., t-1; t = 3, ..., T). \tag{23}$$

Letting $\Delta u_i = [\Delta u_{i3}, ..., \Delta u_{iT}]'$, $\tilde{y}_{it} = y_{it} - \mu$, and

$$Z_i = \overset{(T-2)(T-1)/2}{\begin{bmatrix} \tilde{y}_{i1} & & & \mathbf{0} \\ & \tilde{y}_{i1}, \tilde{y}_{i2} & & \\ & & \ddots & \\ \mathbf{0} & & & \tilde{y}_{i1}, ..., \tilde{y}_{i,T-2} \end{bmatrix}} (T-2) \ ,$$

the moment conditions (23) can be written in vector notation as

$$E(Z_i' \Delta u_i) = 0.$$

The GMM estimator using this moment condition is

$$\hat{\alpha}_{ABGMM} = \left((\Delta y_{-1})' Z V_N^{-1} Z' \Delta y_{-1} \right)^{-1} (\Delta y_{-1})' Z V_N^{-1} Z' \Delta y,$$

where $\Delta y_{-1} = [(\Delta y_{1,-1})', ..., (\Delta y_{N,-1})']'$ with $\Delta y_{i,-1} = [\Delta y_{i2}, ..., \Delta y_{i,T-1}]'$, $\Delta y = [(\Delta y_1)', ..., (\Delta y_N)']'$ with $\Delta y_i = [\Delta y_{i3}, ..., \Delta y_{i,T}]'$, $Z = [Z_1', ..., Z_N']'$ and $V_N = \sum_{i=1}^N Z_i' \Delta u_i \Delta u_i' Z_i$. This estimator requires estimating μ and Δu_i. The parameter μ is estimated by

$\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}$, and Δu_i is estimated by the following IV-GLS estimator under the assumption $Var(u_{it}) = \sigma_u^2$ for all i and t

$$\hat{\alpha}_{IV-GLS} = \left((\Delta y_{-1})' Z(Z'(I_N \otimes H)Z)^{-1} Z' \Delta y_{-1}\right)^{-1} (\Delta y_{-1})' Z(Z'(I_N \otimes H)Z)^{-1} Z' \Delta y,$$

where H is a $(T-2)$ square matrix with 2s in the main diagonal, -1s in the first subdiagonals and 0s elsewhere.

In addition to the moment condition (23), Ahn and Schmidt (1995) find another condition

$$E(u_{iT}\Delta u_{it}) = 0, \ (t = 3, ..., T-1),$$

which is equivalent to

$$E(\tilde{y}_{i,t-2}\Delta u_{i,t-1} - \tilde{y}_{i,t-1}\Delta u_{it}) = 0, \ (t = 4, ..., T), \tag{24}$$

if $Var(u_{it}) = \sigma_{\mathrm{u}}^2$ for every i and t. Letting $u_i = [u_{i2}, ..., u_{iT}]'$ and

$$A_i = \begin{bmatrix} -\tilde{y}_{i2} & 0 & \cdots & 0 \\ \tilde{y}_{i2} + \tilde{y}_{i3} & -\tilde{y}_{i3} & \cdots & 0 \\ -\tilde{y}_{i3} & \tilde{y}_{i3} + \tilde{y}_{i4} & \cdots & 0 \\ & -\tilde{y}_{i4} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & -\tilde{y}_{i,T-2} \\ 0 & 0 & \cdots & \tilde{y}_{i,T-2} + \tilde{y}_{i,T-1} \\ 0 & 0 & \cdots & -\tilde{y}_{i,T-1} \end{bmatrix}$$

the moment condition (24) can be written as

$$E(A_i' u_i) = 0.$$

Similarly, Arellano and Bond's moment condition (23) can be written as $E(B_i' u_i) = 0$, where

$$B_i = \begin{bmatrix} -\tilde{y}_{i1} & 0 & \cdots & 0 \\ \tilde{y}_{i1} & -(\tilde{y}_{i1}\ \tilde{y}_{i2}) & \cdots & 0 \\ 0 & (\tilde{y}_{i1}\ \tilde{y}_{i2}) & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & -(\tilde{y}_{i1}\ \tilde{y}_{i2}\ ...\ \tilde{y}_{i,T-2}) \\ 0 & 0 & \cdots & (\tilde{y}_{i1}\ \tilde{y}_{i2}\ ...\ \tilde{y}_{i,T-2}) \end{bmatrix}.$$

Thus, Ahn and Schmidt's estimator we use is the GMM estimator that applies the instrument matrix $[A_i\ B_i]$ to the level-equation (1). Since $A_i'\mathbf{i}_{T-1} = 0$ and $B_i'\mathbf{i}_{T-1} = 0$ where $\mathbf{i}_{T-1}$ is an $(T-1)$-dimension vector of ones, the intercept term of Model (4) cannot be estimated by Ahn and Schmidt's estimator.

Blundell and Bond (1998) consider Model (4) without the intercept term and propose a GMM estimator that employs more moment conditions than Arellano and Bond's (1991). The additional moment conditions of Blundell and Bond are

$$E\left((u_{it} + m_i(1-\alpha))\Delta y_{i,t-1}\right) = 0, (t = 3, ..., T), \tag{25}$$

which hold under Assumption 1. Stacking Model (4) and its differenced version, we have

$$\begin{pmatrix} \Delta y_i \\ \tilde{y}_i \end{pmatrix} = \alpha \begin{pmatrix} \Delta y_{i,-1} \\ \tilde{y}_{i,-1} \end{pmatrix} + \begin{pmatrix} \Delta u_i \\ u_i + m_i\mathbf{i}_{T-2} \end{pmatrix} = \alpha \begin{pmatrix} \Delta y_{i,-1} \\ \tilde{y}_{i,-1} \end{pmatrix} + \xi_i,$$

where $y_i = [y_{i3}, ..., y_{iT}]'$ and $y_{i,-1} = [y_{i2}, ..., y_{i,T-1}]'$, The instrument matrix for this model is

$$Z_i^+ = \begin{bmatrix} Z_i & & & \mathbf{0} \\ & \Delta y_{i2} & & \\ & & \ddots & \\ \mathbf{0} & & & \Delta y_{i,T-1} \end{bmatrix},$$

which satisfies the condition $E(Z_i^{+\prime}\xi_i) = 0$ under Assumption 1. The GMM estimator of α using the instrument Z_i^+ is defined in the same way as in Arellano and Bond. This GMM estimator is called the system GMM estimator.

Arellano and Bond's, Ahn and Schmidt's and Blundell and Bonds's GMM estimators require $T \geq 3$. All of them were constructed under the assumption $|\alpha| < 1$. When the value α is in the vicinity of 1, Arellano and Bond's and Ahn and Schmidt's GMM estimators suffer from the problem of weak instruments as discussed in Blundell and Bond.

7.1.2 Least squares estimators

First-differencing of Model (3) gives

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta u_{it},$$

which Han and Phillips (2010) transform further such that[11]

$$2\Delta y_{it} + \Delta y_{i,t-1} = \alpha \Delta y_{i,t-1} + \eta_{it},\ \eta_{it} = 2\Delta y_{it} + (1-\alpha)\Delta y_{i,t-1}. \tag{26}$$

Because $E(\Delta y_{i,t-1}\eta_{it}) = 0$ for every $\alpha \in (-1, 1]$ when $E(x_{it}^2) = \sigma_{\mathrm{u}}^2/(1-\alpha^2)$ and $E(x_{it}x_{i,t+1}) = \alpha E(x_{it}^2)$, Han and Phillips suggest using the pooled LSE for Model (26). This estimator is called the first difference least squares estimator (FDLSE). However, their requirements needed for the validity of their method do not hold under Assumption 1. Consider the simplest case $T = 3$. Then, $\Delta y_{i2} = (\alpha - 1)x_{i1} + u_{i2}$ and $\eta_{i3} = (\alpha^2 - 1)x_{i1} + (\alpha - 1)u_{i2} + 2u_{i3}$. Thus, $E(\Delta y_{i2}\eta_{i3}) = (\alpha-1)^2(\alpha+1)\sigma_{\mathrm{x}_1}^2 + (\alpha-1)\sigma_{\mathrm{u}}^2$ is not always equal to zero.

Han, Phillips and Sul (2014) employ Model (3) and the forward-looking regression equation

$$y_{is} = \mu_i(1-\alpha) + \alpha y_{i,s+1} + u_{is}^*,$$

where $u_{is}^* = u_{is} - \alpha(y_{i,s+1} - y_{i,s-1})$. Subtracting this equation from equation (3), the new regression equation

$$y_{it} - y_{is} = \alpha(y_{i,t-1} - y_{i,s+1}) + u_{it} - u_{is}^* \tag{27}$$

is obtained. When $\{u_{it}\}$ are serially uncorrelated, so are the regressors and errors for all $s < t-1$ and $-1 < \alpha \leq 1$ under Han, Phillips and Sul's model set-up. Using

[11] This transformation dates back to Phillips and Han (2008), who use the transformation to estimate the univariate AR(1) model.

the regression equation (27) for $s = 1, 2, ..., t-3$, Han, Phillips and Sul propose the estimator

$$\hat{\alpha}_{PFAE} = \frac{\sum_{i=1}^{N}\sum_{t=4}^{T}\sum_{s=1}^{t-3}(y_{i,t-1} - y_{i,s+1})(y_{it} - y_{is})}{\sum_{i=1}^{N}\sum_{t=4}^{T}\sum_{s=1}^{t-3}(y_{i,t-1} - y_{i,s+1})^2},$$

which is called the panel fully aggregated estimator (PFAE). This looks similar to Hahn, Hausman and Kuersteiner's (2007) estimator, but they are different as explained in Han, Phillips and Sul's (2014, p.207). PFAE requires $T \geq 4$ for its implementation. In addition, the regressors and errors are correlated under Assumption 1. In the simplest case $T = 4$ and $s = 1$, we have $y_{i,3} - y_{i,2} = \alpha(\alpha-1)x_{i1} + (\alpha-1)u_{i2} + u_{i3},\ u_{i4} - u_{i1}^* = \alpha(\alpha-1)x_{i1} + u_{i4} + \alpha u_{i2} - u_{i1}$. Therefore, $E(y_{i,3} - y_{i,2})(u_{i4} - u_{i1}^*) = \alpha^2(\alpha-1)^2\sigma^2_{\mathrm{x}_1} + \alpha(\alpha-1)\sigma^2_{\mathrm{u}} - \alpha(\alpha-1)E(x_{i1}u_{i1})$, which is not always equal to zero.

7.2 Efficiency comparison

Data were generated by Model (1) under Assumption 1 with normality. For simulation, values of the parameters $\alpha,\ \mu,\ \sigma^2_{\mathrm{u}},\ \sigma^2_{\mathrm{m}}$ and $\sigma^2_{\mathrm{x}_1}$, and sample sizes N and T should be selected. First of all, we set $\mu = 0$ and $\sigma^2_{\mathrm{u}} = 1$. For the initial variable x_{i1}, we set

$$\sigma^2_{\mathrm{x}_1} = \begin{cases} \sigma^2_{\mathrm{u}}/(1-\alpha^2), & \text{if } \alpha < 1 \\ 5, & \text{if } \alpha \geq 1 \end{cases} \tag{28}$$

and let $\sigma^2_{\mathrm{m}}/\sigma^2_{\mathrm{u}} = k$. The set-up for the initial variable follows previous studies (e.g., Blundell and Bond, 1998) when $\alpha < 1$. But when $\alpha \geq 1$, the conventional set-up cannot be used. Thus, we chose $\sigma^2_{\mathrm{x}_1} = 5$[12], which is larger than the variance for the stationary case. It is reasonable to do so, because the variance of observations $\{x_{it}\}$ becomes larger as the value of α increases. The parameter values considered are $\alpha = 0.5.0.8, 1.0, 1.1;\ k = 1,\ 2$, and the sample sizes are $N = 100,\ 500$ and $T = 4,\ 6$. The number of iterations for simulation is 1,000, except the case of IIE for

[12] When $\sigma^2_{\mathrm{x}_1} = 10$, it is found that the RMSEs of CSMLE and BCPLSE decrease and those of FDLSE and PFAE increase. CSMLE and BCPLSE continue to show better performance than the rest in this case.

which it is 500. For IIE, we set $H = 10$. All the computation was done by Matlab, and CSMLE was calculated by a Matlab procedure fmincon.

Table 1 reports empirical biases, variances and RMSEs of IIE, CSMLE, the two-stage BCPLSE, Arellano and Bond's (1991) GMM (GMM1), Ahn and Schmidt's (1995) GMM (GMM2) and Blundell and Bond's (1998) GMM (GMM3) estimators, FDLSE and PFAE. The results reported in Table 1 can be summarized as follows.

(i) IIE tends to be less biased than CSMLE as is expected from Theorem 2. But in some cases it has slightly higher variance than CSMLE.

(ii) Comparing the 8 estimators in terms of RMSEs, BCPLSE performs best: its RMSE is the lowest in 42 cases out of 48. PFAE performs best in 6 cases at $\alpha = 0.5$. IIE and CSMLE also show good performance and their RMSEs are quite close to those of BCPLSE in most cases. When PFAE performs best, the differences between BCPLSE and PFAE are marginal. In contrast, when BCPLSE performs best, its RMSE is sometimes less than 3% of that of PFAE (see the case of $(T, \alpha, k) = (6, 1.1, , 1.0)$ at $N = 1000$).

(iii) At fixed N, RMSEs of most of the estimators tend to decrease as T increases.

(iv) At fixed T, RMSEs of most of the estimators decrease as N increases. Exceptions are sometimes observed for GMM1 and GMM2 at $\alpha = 1$. This occurs due to the problem of weak instruments. At the other values of α, their RMSEs also decrease as N increases.

(v) At $\alpha = 1$ and $\alpha = 1.1$, GMM1 and GMM2 perform poorly compared with IIE, CSMLE, BCPLSE and GMM3. This stems from the problem of weak instruments that these estimators share at these values of α.

(vi) At $\alpha = 1.1$, FDLSE and PFAE are strongly biased, causing high RMSEs.

(vii) As the value of k increases, most of the estimators perform poorer. FDLSE and PFAE are invariant to the value of k.

(viii) As the value of α grows, the RMSEs of CSMLE and BCPLSE decrease, while those of FDLSE and PFAE show the opposite behavior.

IIE, CSMLE and BCPLSE are based on the assumption of normality. Naturally, one wonders how these estimators behave when the assumption is violated. To investigate this issue, we generated $\{m_i\}$, $\{u_{it}\}$ and $\{x_{i1}\}$ by using the chi-square distribution with 2 degrees of freedom. Except this, all the features of Assumption 1 are satisfied. The distribution's skewness and kurtosis are 2 and 6, respectively, demonstrating that the distribution is quite skewed and have thick tails compared to the standard normal distribution. The results of this robustness check are reported in Table 2, which is summarized as follows.

(i) RMSE's of IIE, CSMLE and BCPLSE worsen only slightly relative to Table 1. This implies that the procedures of this paper work reasonably well even under the chi-square distribution.

(ii) All the statements made for Table 1 continue to hold in Table 2.

In the experimental design for Tables 1 and 2, $\sigma_{\mathrm{m}}^2 = k$. What would happen if this assumption is violated? To study this issue, we performed the same experiments as in Table 1, assuming that one half of individual effects have variance $0.5\times k$, and the other half $1.5\times k$. The results are reported in Table 3, which we summarize as follows.

(i) RMSE's of IIE, CSMLE and BCPLSE worsen only slightly relative to Table 1. This indicates that the procedures of this paper work reasonably well even when the assumption of homogenous individual effects are violated.

(ii) All the other statements made for Table 1 continue to hold in Table 3.

To summarize the results in Tables 1-3, Table 4 reports frequencies of the 8 estimators performing as best and second-best players in terms of RMSEs. BCPLSE performs best in 121 cases out of 144, GMM3 in 7 cases, and PFAE in 16 cases. Those 7+16 cases are from $\alpha = 0.5,\ 0.8$. In the spots for the second-best players, CSMLE comes 82 times and IIE 21. One last comment to add is that IIE, CSMLE and BCPLSE show quite similar performance, although BCPLSE performs best in a strict sense. But an advantage of the former two is that their standard errors can be computed easily relative to BCPLSE. For BCPLSE, we can use bootstrapping for the construction of confidence intervals as will be in the next subsection.

7.3 Bootstrap confidence intervals

This subsection reports finite-sample performance of the bootstrap confidence interval of the PAR(1) coefficient explained in Section 5. Data were generated as for Table 1. We considered only the case $N = 100$ to save space and computation time. The number of bootstrap iterations (B) is set at 1,000. Empirical coverage ratios of the 95% and 90% confidence intervals based on 300 iterations are reported in Table 5. The results in Table 5 show the coverage ratios are reasonably close to the nominal coverage ratios. It is expected that they can be improved further by increasing the number of bootstrap iterations, although this incurs longer computation time.

8 Summary and further remarks

We have proposed three new estimators for the PAR(1) models with short T and large N. These estimators are based on the cross-sectional regression model using the first time series observations as a regressor and the last as a dependent variable. The regressors and errors of this regression model are dependent. The first estimator is the cross-sectional MLE under the assumption of normal distributions that are consistent in the presence of the regressor-error dependency of the cross-sectional regression model. Using the cross-sectional MLE, we constructed the indirect inference estimator and the bias-corrected pooled least squares estimator. These three estimators were also extended to the PAR model with endogenous, time-invariant regressors. The estimators of this paper are consistent for the PAR coefficients for stationary, unit root and explosive PAR models, estimate the coefficients of time-invariant regressors consistently and can be computed as long as $T \geq 2$. The estimators were shown to perform quite well in finite samples relative to existing estimators.

This paper is focused only on the PAR(1) model, but it is possible to extend the methods of this paper to higher-order PAR models and panel vector autoregressive model. These are deemed to be meaningful extensions of the idea of this paper and await further research in the future. In addition, if individual effects are heterogenous, the idea of grouped-fixed effects method (cf. Bonhomme, Lamadon and Manresa,

2016) may be applied.

Figure 1:Likelihood function of α & σ_m^2

α_o=1, N=100,T=4

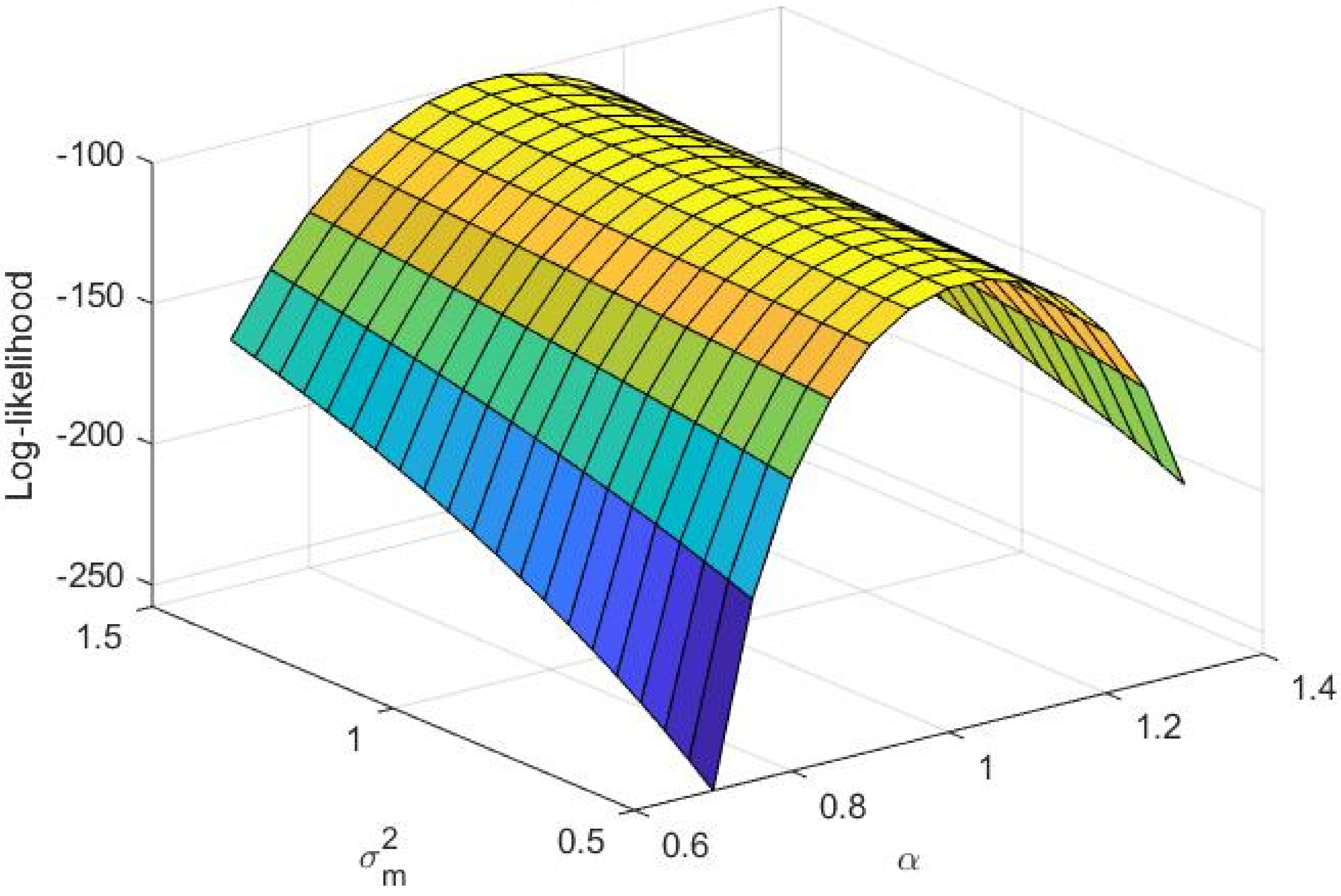
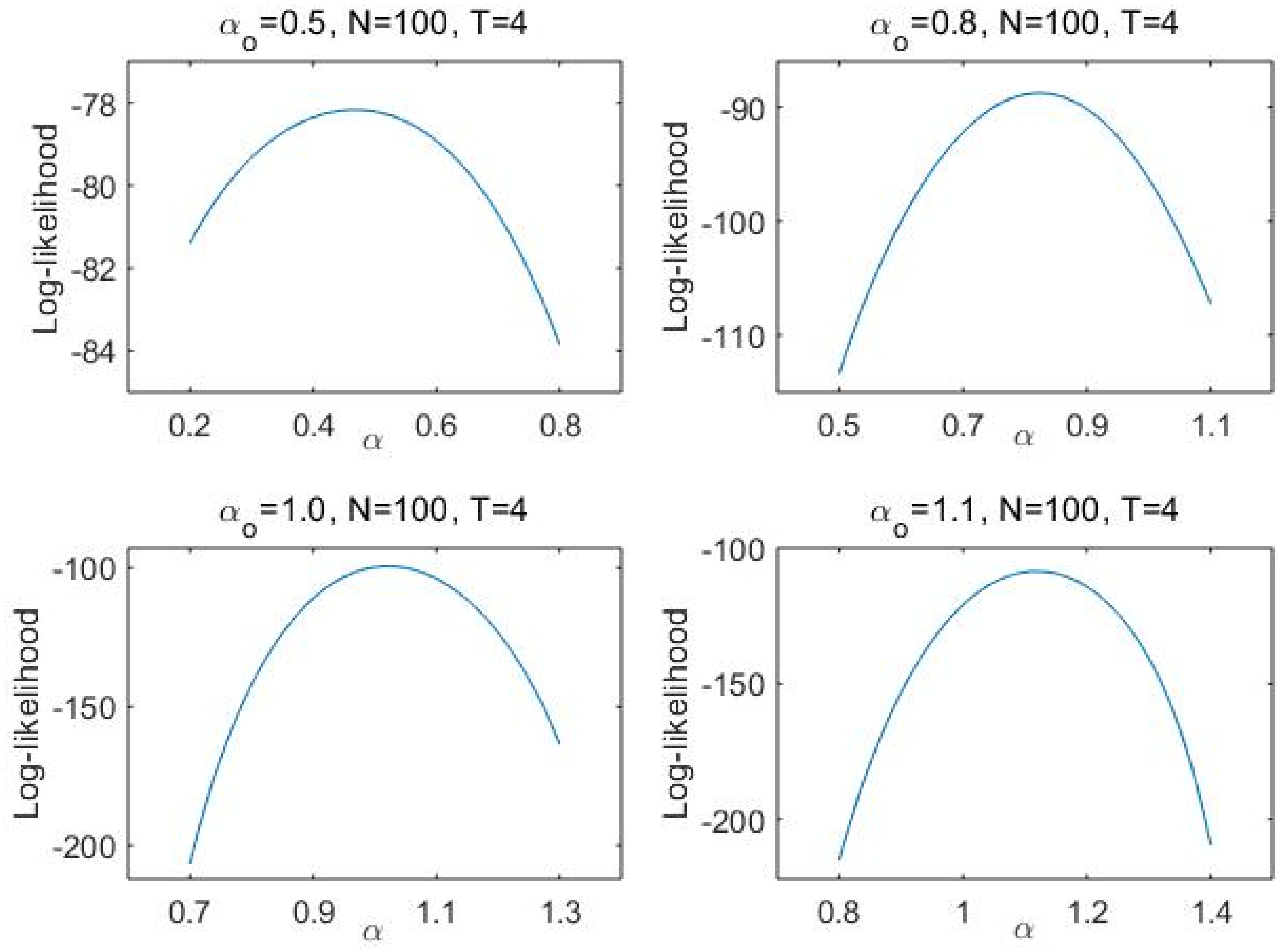


Figure 2: Likelihood function of α



Appendix I Proofs

Proof of Theorem 1. This follows from the standard theory of MLEs, noting that

$$\frac{\partial l_i}{\partial \alpha} = \left(-\frac{1}{2}\right)\left[\frac{\omega_{11.2}^{d\alpha}}{\omega_{11.2}} + \frac{2r_{iT}r_{iT}^{d\alpha}}{\omega_{11.2}} - \frac{r_{iT}^2\omega_{11.2}^{d\alpha}}{\omega_{11.2}^2}\right]$$

and

$$E(\frac{\partial l_i}{\partial \alpha})^2 = \frac{\left(\omega_{11.2}^{d\alpha}\right)^2}{2\left(\omega_{11.2}\right)^2} + \frac{(T-1)^2\alpha^{2T-4}(\omega_{22}-\sigma_{\mathrm{m}}^2)^2}{\omega_{11.2}\omega_{22}},$$

where $r_{iT}^{d\alpha} = y_{i1}(T-1)\alpha^{T-2}\left(\sigma_{\mathrm{m}}^2/\omega_{22}-1\right)$.

Proof of Theorem 2: (i) Under the given assumptions, $\hat{\alpha}_{CSMLE} = \hat{\alpha}_{IIE}^{\infty} + N^{-1}c(\hat{\alpha}_{IIE}^{\infty})$. Since $\hat{\alpha}_{CSMLE} \xrightarrow{p} \alpha_o$ and $\hat{\alpha}_{IIE}^{\infty} = O_p(1)$, $N^{-1}c(\hat{\alpha}_{IIE}^{\infty}) = o_p(1)$ and $\hat{\alpha}_{IIE}^{\infty} \xrightarrow{p} \alpha_o$.

(ii) $\sqrt{N}(\hat{\alpha}_{CSMLE}-\alpha_o) - \sqrt{N}(\hat{\alpha}_{IIE}^{\infty}-\alpha_o) = N^{-1/2}c(\hat{\alpha}_{IIE}^{\infty}) = O_p(N^{-1/2})$. Thus, the stated result follows.

(iii) We obtain by the mean-value theorem

$$\begin{aligned}
E(\hat{\alpha}_{IIE}^{\infty}) &= E(\hat{\alpha}_{CSMLE}) - Ec_N(\hat{\alpha}_{IIE}^{\infty}) \\
&= \alpha_o + N^{-1}c(\alpha_o) - N^{-1}E\left[c(\hat{\alpha}_{IIE}^{\infty})\right] \\
&= \alpha_o - E\left[N^{-3/2}c'(\alpha^*)\sqrt{N}(\hat{\alpha}_{IIE}^{\infty}-\alpha_o)\right],
\end{aligned}$$

where α^* lies in between $\hat{\alpha}_{IIE}^{\infty}$ and α_o. The result follows because $N^{-3/2}c'(\alpha^*)\sqrt{N}(\hat{\alpha}_{IIE}^{\infty} - \alpha_o) = O_p(N^{-3/2})$.

Proof of Theorem 3: (i) This follows from Assumption 1 and Theorem 2.
(ii) We have

$$\begin{aligned}
&\sqrt{N}\left(\hat{\alpha}_{BCPLSE}-\alpha\right) \\
= \;&\sqrt{N}\left(\hat{\alpha}_{PLSE} - \alpha - \frac{(1-\alpha)(T-1)\sum_{i=1}^N\sum_{t=2}^T(m_i-\bar{m})m_i}{\sum_{i=1}^N\sum_{t=2}^T(y_{i,t-1}-\bar{y}_{-1})^2}\right) \\
&-\sqrt{N}\left(\frac{(1-\alpha)(T-1)\sum_{i=1}^N\sum_{t=2}^T(m_i-\bar{m})m_i/N}{\sum_{i=1}^N\sum_{t=2}^T(y_{i,t-1}-\bar{y}_{-1})^2/N} - \frac{(1-\hat{\alpha}_{CSMLE})(T-1)\hat{\sigma}_{\mathrm{m}CSMLE}^2}{\sum_{i=1}^N\sum_{t=2}^T(y_{i,t-1}-\bar{y}_{-1})^2/N}\right) \\
= \;&A_N - B_N, \text{ say.}
\end{aligned}$$

Equation (15) indicates $A_N = O_p(1)$ under Assumption 1. In addition,

$$\begin{aligned}
&\sqrt{N}\left[(1-\alpha)(T-1)\sum_{i=1}^{N}\sum_{t=2}^{T}(m_i-\bar{m})m_i/N-(1-\hat{\alpha}_{CSMLE})(T-1)\hat{\sigma}^2_{\mathrm{m}CSMLE}\right]\\
=\;& (T-1)\sqrt{N}\left[\sum_{i=1}^{N}(m_i-\bar{m})m_i/N-\hat{\sigma}^2_{\mathrm{m}CSMLE}\right]\\
&+(T-1)\sqrt{N}\left[\hat{\alpha}_{CSMLE}\hat{\sigma}^2_{\mathrm{m}CSMLE}-\alpha\sum_{i=1}^{N}(m_i-\bar{m})m_i/N\right]\\
=\;& (T-1)\sqrt{N}\left[\frac{1}{N}\sum_{i=1}^{N}\left((m_i-\bar{m})m_i-\sigma^2_{\mathrm{m}}\right)-\left(\hat{\sigma}^2_{\mathrm{m}CSMLE}-\sigma^2_{\mathrm{m}}\right)\right]\\
&+(T-1)\sqrt{N}\left[(\hat{\alpha}_{CSMLE}-\alpha)\hat{\sigma}^2_{\mathrm{m}CSMLE}-\alpha\left(\sum_{i=1}^{N}(m_i-\bar{m})m_i/N-\hat{\sigma}^2_{\mathrm{m}CSMLE}\right)\right]\\
=\;& (T-1)\sqrt{N}\left[\frac{1}{N}\sum_{i=1}^{N}\left((m_i-\bar{m})m_i-\sigma^2_{\mathrm{m}}\right)-\left(\hat{\sigma}^2_{\mathrm{m}CSMLE}-\sigma^2_{\mathrm{m}}\right)\right]\\
&+\hat{\sigma}^2_{\mathrm{m}CSMLE}(T-1)\sqrt{N}\left[(\hat{\alpha}_{CSMLE}-\alpha)\right]\\
&-\alpha(T-1)\sqrt{N}\left[\frac{1}{N}\sum_{i=1}^{N}\left((m_i-\bar{m})m_i-\sigma^2_{\mathrm{m}}\right)-\left(\hat{\sigma}^2_{\mathrm{m}CSMLE}-\sigma^2_{\mathrm{m}}\right)\right]\\
=\;& (1-\alpha)(T-1)\sqrt{N}\left[\frac{1}{N}\sum_{i=1}^{N}\left((m_i-\bar{m})m_i-\sigma^2_{\mathrm{m}}\right)-\left(\hat{\sigma}^2_{\mathrm{m}CSMLE}-\sigma^2_{\mathrm{m}}\right)\right]\\
&+\hat{\sigma}^2_{\mathrm{m}CSMLE}(T-1)\sqrt{N}(\hat{\alpha}_{CSMLE}-\alpha)\\
=\;& O_p(1),
\end{aligned}$$

which implies $B_N = O_p(1)$. Thus, the result follows.

Proof of Theorem 4: The stated results require lengthy and tedious calculations. The reader is refereed to http://inchoi.sogang.ac.kr/papers/list_hi.php for details.

Proof of Theorem 5: The result can shown in the same way as for Theorem 2**.**

Proof of Theorem 6: As in Section 5, it is straightforward to show under Assumption 5 (i) $\hat{\psi}_{BCPLSE} \xrightarrow{p} \psi$ and (ii) $\sqrt{N}(\hat{\psi}_{BCPLSE} - \psi) = O_p(1)$. The results follow from these.

References

Ahn, S. C., & Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. Journal of Econometrics, 68(1), 5-27.

Anderson, T. W., & Hsiao, C. (1981). Estimation of dynamic models with error components. Journal of the American Statistical Association, 76(375), 598-606.

Anderson, T. W., & Hsiao, C. (1982). Formulation and estimation of dynamic models using panel data. Journal of Econometrics, 18(1), 47-82.

Andrews, D. W. K. (1993). Exactly median-unbiased estimation of first order autoregressive/unit root models. Econometrica, 61(1), 139–165.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The Review of Economic Studies, 58(2), 277-297.

Ashley, R. A., & Sun, X. (2016). Subset-Continuous-Updating GMM Estimators for Dynamic Panel Data Models. Econometrics, 4(4), 47.

Bai, J. (2013). Fixed-effects dynamic panel models, a factor analytical method. Econometrica, 81(1), 285-314.

Baltagi, B. (2008). Econometric analysis of panel data. John Wiley & Sons.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. Journal of Econometrics, 87(1), 115-143.

Bonhomme, S., Lamadon, T. & Manresa, E. (2016). Discretizing unobserved heterogeneity: approximate clustering methods for dimension reduction. Unpublished manuscript, University of Chicago.

Bun, M. J. & Carree, M. A. (2005) Bias-corrected estimation in dynamic panel data models. Journal of Business & Economic Statistics, 23(2), 200–210.

Bun, M. J., & Sarafidis, V. (2015). Dynamic panel data models. Oxford Handbook of Panel Data Econometrics, 76-110, Oxford University Press.

Bun, M. J., & Windmeijer, F. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. The Econometrics Journal, 13(1), 95-126.

Chamberlain, G. (1984). Panel data. Handbook of Econometrics, Vol. 2, 1247-1318, Elsevier.

Choi, I. (2015). Almost all about unit roots: foundations, developments and applications. Cambridge University Press.

Cox, D. R. and Snell, E. J. (1968). A general definition of residuals. Journal of the Royal Statistical Society, Series B, 30(2), 248–275.

Gourieroux, C., Monfort, A., & Renault, E. (1993). Indirect inference. Journal of Applied Econometrics, 8(S1), S85-S118.

Gourieroux, C., Phillips, P. C., & Yu, J. (2010). Indirect inference for dynamic panel models. Journal of Econometrics, 157(1), 68-77.

Hahn, J. (1999). How informative is the initial condition in the dynamic panel model with fixed effects?. Journal of Econometrics, 93(2), 309-326.

Hahn, J., Hausman, J., & Kuersteiner, G. (2007). Long difference instrumental variables estimation for dynamic panel models with fixed effects. Journal of Econometrics, 140(2), 574-617.

Hahn, J., & Kuersteiner, G. (2002). Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large. Econometrica, 70(4), 1639-1657.

Han, C., & Phillips, P. C. (2010). GMM estimation for dynamic panels with fixed effects and strong instruments at unity. Econometric Theory, 26(01), 119-151

Han, C., & Phillips, P. C. (2013). First difference maximum likelihood and dynamic panel estimation. Journal of Econometrics, 175(1), 35-45.

Han, C., Phillips, P. C., & Sul, D. (2014). X-differencing and dynamic panel model estimation. Econometric Theory, 30(01), 201-251.

Hansen, L. P., Heaton, J., & Yaron, A. (1996). Finite-sample properties of some alternative GMM estimators. Journal of Business & Economic Statistics, 14(3), 262-280.

Hayakawa, K. (2007). Small sample bias properties of the system GMM estimator in dynamic panel data models. Economics Letters, 95(1), 32-38.

Hayakawa, K. (2012). GMM estimation of short dynamic panel data models with interactive fixed effects. Journal of the Japan Statistical Society, 42(2), 109-123.

Hsiao, C. (2003). Analysis of panel data. Second Edition, Cambridge University Press.

Hsiao, C., Pesaran, M. H., & Tahmiscioglu, A. K. (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. Journal of Econometrics, 109(1), 107-150.

Hsiao, C. and Q. Zhou (2018). Incidental parameters, initial conditions and sample size in statistical inference for dynamic panel data models. Journal of Econometrics, 207(1), 114-128.

Kiviet, J. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models, Journal of Econometrics, 68(1), 53-78.

Kruiniger, H. (2008). Maximum likelihood estimation and inference methods for the covariance stationary panel AR (1)/unit root model. Journal of Econometrics, 144(2), 447-464.

Muirhead, R. J. (1982). Aspects of multivariate statistical theory, Wiley.

Newey, W.K. and D.McFadden (1994). Large sample estimation and hypothesis testing, Handbook of Econometrics, Vol. 4, 2111-2245, Elsevier.

Phillips, P. C., & Han, C. (2008). Gaussian inference in AR (1) time series with or without a unit root. Econometric Theory, 24(03), 631-650.

Phillips, P. C., Wu, Y., & Yu, J. (2011). Explosive Behavior In The 1990S Nasdaq: When Did Exuberance Escalate Asset Values?. International Economic Review, 52(1), 201-226.

Smith, A. A. (1993). Estimating nonlinear time-series models using simulated vector autoregressions, Journal of Applied Econometrics, 8(S1), S63–S84.

Table 1: Efficiency comparison under normal distributions

Note: Data were generated by Models (1) and (2) under Assumption 1 with $\sigma_{\mathrm{u}}^2 = 1$. Formula (28) provides the value of $\sigma_{\mathrm{x}_1}^2$. The number of iterations is 1000 for all estimators except IIE. The number of iterations is 500 for IIE.

Part A: $N = 100$

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(4,0.5,1.0)	Bias	-0.0266	-0.0365	-0.0389	-0.0124	0.0155	-0.0062	0.0043	0.0032
	Var	0.0110	0.0144	0.0076	0.0355	0.0326	0.0142	0.0146	0.0148
	RMSE	0.1084	0.1252	0.0953	0.1888	0.1813	0.1192	0.1211	0.1218
(4,0.5,2.0)	Bias	-0.0182	-0.0264	-0.0320	-0.0183	0.0234	0.0011	0.0043	0.0032
	Var	0.0136	0.0174	0.0101	0.0498	0.0443	0.0165	0.0146	0.0148
	RMSE	0.1180	0.1346	0.1057	0.2239	0.2117	0.1283	0.1211	0.1218
(4,0.8,1.0)	Bias	-0.0214	-0.0259	-0.0247	-0.0329	-0.0236	-0.0274	0.0014	0.0006
	Var	0.0033	0.0035	0.0023	0.0774	0.0573	0.0201	0.0170	0.0171
	RMSE	0.0617	0.0643	0.0538	0.2801	0.2405	0.1443	0.1303	0.1308
(4,0.8,2.0)	Bias	-0.0199	-0.0243	-0.0232	-0.0462	-0.0293	-0.0271	0.0014	0.0006
	Var	0.0044	0.0045	0.0032	0.1038	0.0691	0.0224	0.0170	0.0171
	RMSE	0.0695	0.0714	0.0608	0.3255	0.2644	0.1520	0.1303	0.1308
(4,1.0,1.0)	Bias	0.0010	-0.0013	-0.0011	-0.9061	-0.5763	-0.0193	-0.0046	-0.0044
	Var	0.0010	0.0009	0.0008	0.8369	0.4535	0.0151	0.0193	0.0197
	RMSE	0.0314	0.0304	0.0291	1.2876	0.8863	0.1244	0.1391	0.1405
(4,1.0,2.0)	Bias	0.0003	-0.0013	-0.0011	-0.9106	-0.5733	-0.0207	-0.0046	-0.0044
	Var	0.0011	0.0010	0.0010	0.7852	0.4636	0.0170	0.0193	0.0197
	RMSE	0.0328	0.0319	0.0309	1.2706	0.8901	0.1321	0.1391	0.1405
(4,1.1,1.0)	Bias	0.0035	0.0019	0.0019	-0.1018	-0.2159	-0.0046	0.2047	0.2039
	Var	0.0007	0.0006	0.0006	0.1587	0.1753	0.0061	0.0227	0.0231
	RMSE	0.0266	0.0251	0.0245	0.4112	0.4711	0.0780	0.2542	0.2544
(4,1.1,2.0)	Bias	0.0023	0.0007	0.0007	-0.1223	-0.2300	-0.0039	0.2047	0.2039
	Var	0.0007	0.0007	0.0006	0.1867	0.2065	0.0071	0.0227	0.0231
	RMSE	0.0262	0.0259	0.0250	0.4491	0.5093	0.0843	0.2542	0.2544

Part A: $N = 100$ (continued)

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(6,0.5,1.0)	Bias	-0.0023	-0.0148	-0.0154	-0.0423	0.0317	0.0091	0.0042	0.0009
	Var	0.0101	0.0131	0.0070	0.0115	0.0178	0.0089	0.0075	0.0041
	RMSE	0.1005	0.1156	0.0851	0.1152	0.1372	0.0950	0.0869	0.0641
(6,0.5,2.0)	Bias	0.0037	-0.0116	-0.0139	-0.0520	0.0511	0.0287	0.0042	0.0009
	Var	0.0114	0.0150	0.0090	0.0145	0.0254	0.0136	0.0075	0.0041
	RMSE	0.1071	0.1232	0.0959	0.1312	0.1673	0.1199	0.0869	0.0641
(6,0.8,1.0)	Bias	-0.0248	-0.0322	-0.0297	-0.0820	-0.0180	-0.0147	0.0043	0.0014
	Var	0.0032	0.0037	0.0019	0.0192	0.0205	0.0103	0.0087	0.0044
	RMSE	0.0617	0.0692	0.0525	0.1611	0.1444	0.1028	0.0935	0.0665
(6,0.8,2.0)	Bias	-0.0219	-0.0334	-0.0316	-0.1021	-0.0182	-0.0094	0.0043	0.0014
	Var	0.0042	0.0050	0.0027	0.0241	0.0254	0.0119	0.0087	0.0044
	RMSE	0.0686	0.0781	0.0608	0.1857	0.1603	0.1096	0.0935	0.0665
(6,1.0,1.0)	Bias	-0.0004	-0.0019	-0.0023	-0.8195	-0.3595	-0.0197	0.0034	-0.0002
	Var	0.0005	0.0005	0.0004	0.1954	0.2242	0.0091	0.0104	0.0051
	RMSE	0.0226	0.0216	0.0195	0.9311	0.5945	0.0974	0.1020	0.0715
(6,1.0,2.0)	Bias	-0.0006	-0.0020	-0.0024	-0.8192	-0.3441	-0.0204	0.0034	-0.0002
	Var	0.0005	0.0005	0.0004	0.1942	0.2172	0.0099	0.0104	0.0051
	RMSE	0.0228	0.0226	0.0204	0.9303	0.5793	0.1018	0.1020	0.0715
(6,1.1,1.0)	Bias	0.0018	0.0004	-0.0001	-0.0855	-0.0756	-0.0069	0.2578	0.2005
	Var	0.0003	0.0003	0.0002	0.0247	0.0284	0.0039	0.0133	0.0051
	RMSE	0.0168	0.0161	0.0143	0.1789	0.1846	0.0627	0.2824	0.2129
(6,1.1,2.0)	Bias	0.0010	-0.0002	-0.0008	-0.0932	-0.0820	-0.0066	0.2578	0.2005
	Var	0.0003	0.0003	0.0002	0.0270	0.0283	0.0041	0.0133	0.0051
	RMSE	0.0172	0.0167	0.0146	0.1888	0.1872	0.0643	0.2824	0.2129

Part B: $N = 500$

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(4,0.5,1.0)	Bias	-0.0114	-0.0154	-0.0207	0.0019	0.0074	-0.0020	0.0009	0.0005
	Var	0.0033	0.0032	0.0019	0.0069	0.0067	0.0027	0.0032	0.0031
	RMSE	0.0587	0.0589	0.0485	0.0831	0.0822	0.0524	0.0567	0.0560
(4,0.5,2.0)	Bias	-0.0090	-0.0094	-0.0162	0.0014	0.0094	-0.0023	0.0009	0.0005
	Var	0.0045	0.0042	0.0028	0.0095	0.0091	0.0032	0.0032	0.0031
	RMSE	0.0675	0.0656	0.0551	0.0973	0.0958	0.0563	0.0567	0.0560
(4,0.8,1.0)	Bias	-0.0127	-0.0126	-0.0132	0.0010	0.0018	-0.0042	0.0006	0.0004
	Var	0.0007	0.0007	0.0005	0.0142	0.0134	0.0031	0.0038	0.0038
	RMSE	0.0291	0.0285	0.0268	0.1190	0.1157	0.0562	0.0613	0.0613
(4,0.8,2.0)	Bias	-0.0105	-0.0108	-0.0115	0.0002	0.0011	-0.0061	0.0006	0.0004
	Var	0.0008	0.0008	0.0006	0.0179	0.0166	0.0035	0.0038	0.0038
	RMSE	0.0309	0.0297	0.0276	0.1340	0.1287	0.0596	0.0613	0.0613
(4,1.0,1.0)	Bias	-0.0008	0.0003	0.0001	-0.8893	-0.7471	0.0006	0.0006	0.0006
	Var	0.0002	0.0002	0.0001	0.7894	0.4243	0.0009	0.0041	0.0041
	RMSE	0.0136	0.0127	0.0121	1.2571	0.9912	0.0300	0.0639	0.0641
(4,1.0,2.0)	Bias	-0.0009	0.0004	0.0002	-0.8863	-0.7308	0.0007	0.0006	0.0006
	Var	0.0002	0.0002	0.0002	0.7228	0.4241	0.0009	0.0041	0.0041
	RMSE	0.0143	0.0134	0.0128	1.2281	0.9789	0.0306	0.0639	0.0641
(4,1.1,1.0)	Bias	0.0030	0.0034	0.0031	-0.0201	-0.0451	0.0008	0.2141	0.2115
	Var	0.0001	0.0001	0.0001	0.0217	0.0259	0.0004	0.0045	0.0045
	RMSE	0.0115	0.0109	0.0103	0.1485	0.1672	0.0190	0.2244	0.2218
(4,1.1,2.0)	Bias	0.0023	0.0027	0.0025	-0.0233	-0.0502	0.0011	0.2141	0.2115
	Var	0.0001	0.0001	0.0001	0.0250	0.0301	0.0004	0.0045	0.0045
	RMSE	0.0119	0.0110	0.0105	0.1598	0.1805	0.0194	0.2244	0.2218

Part B: $N = 500$ (continued)

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(6,0.5,1.0)	Bias	-0.0026	-0.0031	-0.0050	-0.0060	0.0069	-0.0011	0.0038	0.0014
	Var	0.0025	0.0024	0.0014	0.0021	0.0024	0.0010	0.0016	0.0009
	RMSE	0.0504	0.0489	0.0377	0.0458	0.0491	0.0323	0.0396	0.0296
(6,0.5,2.0)	Bias	-0.0002	-0.0010	-0.0035	-0.0071	0.0118	0.0003	0.0038	0.0014
	Var	0.0029	0.0028	0.0019	0.0026	0.0033	0.0011	0.0016	0.0009
	RMSE	0.0537	0.0532	0.0435	0.0512	0.0583	0.0338	0.0396	0.0296
(6,0.8,1.0)	Bias	-0.0102	-0.0112	-0.0111	-0.0146	-0.0043	-0.0065	0.0035	0.0007
	Var	0.0006	0.0007	0.0005	0.0035	0.0036	0.0012	0.0018	0.0010
	RMSE	0.0272	0.0289	0.0248	0.0609	0.0601	0.0355	0.0427	0.0311
(6,0.8,2.0)	Bias	-0.0090	-0.0105	-0.0108	-0.0173	-0.0039	-0.0076	0.0035	0.0007
	Var	0.0008	0.0009	0.0006	0.0044	0.0045	0.0014	0.0018	0.0010
	RMSE	0.0294	0.0312	0.0263	0.0683	0.0672	0.0384	0.0427	0.0311
(6,1.0,1.0)	Bias	-0.0000	0.0002	-0.0000	-0.8341	-0.5166	-0.0021	0.0025	-0.0004
	Var	0.0001	0.0001	0.0001	0.1976	0.1862	0.0006	0.0020	0.0010
	RMSE	0.0097	0.0095	0.0086	0.9452	0.6732	0.0241	0.0449	0.0312
(6,1.0,2.0)	Bias	0.0002	0.0002	-0.0000	-0.8350	-0.5019	-0.0021	0.0025	-0.0004
	Var	0.0001	0.0001	0.0001	0.1977	0.1835	0.0007	0.0020	0.0010
	RMSE	0.0101	0.0101	0.0091	0.9459	0.6599	0.0256	0.0449	0.0312
(6,1.1,1.0)	Bias	0.0016	0.0016	0.0014	-0.0106	-0.0250	-0.0001	0.2588	0.2027
	Var	0.0001	0.0000	0.0000	0.0034	0.0042	0.0002	0.0025	0.0009
	RMSE	0.0077	0.0072	0.0065	0.0592	0.0695	0.0141	0.2635	0.2049
(6,1.1,2.0)	Bias	0.0014	0.0010	0.0009	-0.0125	-0.0267	0.0004	0.2588	0.2027
	Var	0.0001	0.0001	0.0000	0.0038	0.0045	0.0002	0.0025	0.0009
	RMSE	0.0082	0.0074	0.0067	0.0628	0.0719	0.0147	0.2635	0.2049

Part C: $N = 1000$

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(4,0.5,1.0)	Bias	-0.0097	-0.0141	-0.0160	-0.0005	0.0022	-0.0016	0.0014	0.0010
	Var	0.0016	0.0014	0.0009	0.0034	0.0034	0.0014	0.0016	0.0015
	RMSE	0.0416	0.0401	0.0344	0.0587	0.0583	0.0373	0.0404	0.0391
(4,0.5,2.0)	Bias	-0.0050	-0.0101	-0.0132	-0.0009	0.0031	-0.0017	0.0014	0.0010
	Var	0.0022	0.0018	0.0012	0.0047	0.0046	0.0016	0.0016	0.0015
	RMSE	0.0470	0.0442	0.0377	0.0684	0.0678	0.0399	0.0404	0.0391
(4,0.8,1.0)	Bias	-0.0081	-0.0104	-0.0100	-0.0025	-0.0021	-0.0027	0.0012	0.0011
	Var	0.0004	0.0004	0.0003	0.0072	0.0070	0.0016	0.0019	0.0019
	RMSE	0.0213	0.0219	0.0199	0.0848	0.0836	0.0399	0.0439	0.0437
(4,0.8,2.0)	Bias	-0.0059	-0.0093	-0.0090	-0.0033	-0.0028	-0.0038	0.0012	0.0011
	Var	0.0005	0.0005	0.0004	0.0090	0.0086	0.0017	0.0019	0.0019
	RMSE	0.0228	0.0248	0.0230	0.0948	0.0929	0.0418	0.0439	0.0437
(4,1.0,1.0)	Bias	0.0008	-0.0002	-0.0001	-0.9119	-0.8158	0.0002	0.0008	0.0009
	Var	0.0001	0.0001	0.0001	1.2067	0.4797	0.0004	0.0021	0.0021
	RMSE	0.0097	0.0091	0.0087	1.4277	1.0701	0.0208	0.0455	0.0456
(4,1.0,2.0)	Bias	0.0009	-0.0003	-0.0002	-0.9247	-0.8163	0.0001	0.0008	0.0009
	Var	0.0001	0.0001	0.0001	1.5444	0.4659	0.0004	0.0021	0.0021
	RMSE	0.0102	0.0097	0.0093	1.5490	1.0641	0.0210	0.0455	0.0456
(4,1.1,1.0)	Bias	0.0034	0.0027	0.0026	-0.0077	-0.0201	0.0003	0.2142	0.2114
	Var	0.0001	0.0001	0.0001	0.0111	0.0119	0.0002	0.0023	0.0023
	RMSE	0.0088	0.0081	0.0077	0.1055	0.1111	0.0132	0.2195	0.2167
(4,1.1,2.0)	Bias	0.0027	0.0023	0.0023	-0.0095	-0.0227	0.0004	0.2142	0.2114
	Var	0.0001	0.0001	0.0001	0.0126	0.0137	0.0002	0.0023	0.0023
	RMSE	0.0091	0.0082	0.0079	0.1128	0.1191	0.0134	0.2195	0.2167

Part C: $N = 1000$ (continued)

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(6,0.5,1.0)	Bias	-0.0015	-0.0040	-0.0069	-0.0024	0.0038	-0.0007	0.0014	0.0007
	Var	0.0013	0.0012	0.0007	0.0011	0.0011	0.0005	0.0008	0.0004
	RMSE	0.0361	0.0353	0.0265	0.0327	0.0334	0.0222	0.0276	0.0209
(6,0.5,2.0)	Bias	-0.0002	-0.0022	-0.0056	-0.0030	0.0060	-0.0003	0.0014	0.0007
	Var	0.0015	0.0014	0.0008	0.0013	0.0015	0.0005	0.0008	0.0004
	RMSE	0.0392	0.0377	0.0296	0.0368	0.0390	0.0232	0.0276	0.0209
(6,0.8,1.0)	Bias	-0.0089	-0.0088	-0.0088	-0.0062	-0.0012	-0.0036	0.0013	0.0000
	Var	0.0004	0.0004	0.0003	0.0017	0.0017	0.0005	0.0009	0.0005
	RMSE	0.0223	0.0215	0.0185	0.0422	0.0414	0.0235	0.0296	0.0216
(6,0.8,2.0)	Bias	-0.0086	-0.0080	-0.0083	-0.0076	-0.0012	-0.0047	0.0013	0.0000
	Var	0.0006	0.0005	0.0004	0.0022	0.0022	0.0006	0.0009	0.0005
	RMSE	0.0252	0.0235	0.0206	0.0474	0.0464	0.0252	0.0296	0.0216
(6,1.0,1.0)	Bias	-0.0005	0.0001	0.0000	-0.8135	-0.5888	-0.0009	0.0013	-0.0001
	Var	0.0001	0.0000	0.0000	0.1886	0.1835	0.0002	0.0010	0.0005
	RMSE	0.0076	0.0063	0.0056	0.9222	0.7281	0.0133	0.0317	0.0224
(6,1.0,2.0)	Bias	-0.0006	0.0001	-0.0000	-0.8113	-0.5758	-0.0009	0.0013	-0.0001
	Var	0.0001	0.0000	0.0000	0.1877	0.1838	0.0002	0.0010	0.0005
	RMSE	0.0082	0.0067	0.0060	0.9197	0.7179	0.0137	0.0317	0.0224
(6,1.1,1.0)	Bias	0.0011	0.0012	0.0012	-0.0061	-0.0142	0.0000	0.2588	0.2038
	Var	0.0000	0.0000	0.0000	0.0016	0.0018	0.0001	0.0013	0.0005
	RMSE	0.0062	0.0054	0.0050	0.0405	0.0447	0.0079	0.2614	0.2050
(6,1.1,2.0)	Bias	0.0005	0.0005	0.0006	-0.0071	-0.0148	0.0002	0.2588	0.2038
	Var	0.0000	0.0000	0.0000	0.0018	0.0020	0.0001	0.0013	0.0005
	RMSE	0.0066	0.0058	0.0053	0.0430	0.0470	0.0081	0.2614	0.2050

Table 2: Efficiency comparison under nonnormal distributions

Note: Note for Table 1 also applies here except that $(\chi^2(2) - 2)/2$ is used instead of standard normal variates.

Part A: $N = 100$

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(4,0.5,1.0)	Bias	-0.0184	-0.0416	-0.0391	-0.0110	-0.0081	-0.0104	0.0059	0.0018
	Var	0.0169	0.0321	0.0086	0.0340	0.0349	0.0149	0.0225	0.0214
	RMSE	0.1314	0.1840	0.1005	0.1847	0.1871	0.1224	0.1501	0.1463
(4,0.5,2.0)	Bias	-0.0125	-0.0326	-0.0330	-0.0161	-0.0067	-0.0042	0.0059	0.0018
	Var	0.0181	0.0371	0.0133	0.0446	0.0472	0.0176	0.0225	0.0214
	RMSE	0.1353	0.1952	0.1199	0.2117	0.2174	0.1326	0.1501	0.1463
(4,0.8,1.0)	Bias	-0.0349	-0.0416	-0.0387	-0.0437	-0.0789	-0.0266	0.0032	-0.0014
	Var	0.0075	0.0093	0.0042	0.0912	0.0757	0.0197	0.0220	0.0246
	RMSE	0.0935	0.1050	0.0755	0.3051	0.2862	0.1428	0.1484	0.1570
(4,0.8,2.0)	Bias	-0.0281	-0.0414	-0.0381	-0.0373	-0.0939	-0.0255	0.0032	-0.0014
	Var	0.0083	0.0124	0.0059	0.6916	0.0903	0.0220	0.0220	0.0246
	RMSE	0.0951	0.1186	0.0856	0.8324	0.3149	0.1505	0.1484	0.1570
(4,1.0,1.0)	Bias	0.0015	-0.0010	-0.0010	-0.9639	-0.6930	-0.0222	-0.0022	-0.0027
	Var	0.0015	0.0016	0.0009	1.1333	0.4539	0.0138	0.0214	0.0235
	RMSE	0.0389	0.0395	0.0304	1.4361	0.9665	0.1194	0.1463	0.1535
(4,1.0,2.0)	Bias	0.0018	-0.0008	-0.0009	-0.9224	-0.6730	-0.0249	-0.0022	-0.0027
	Var	0.0016	0.0016	0.0009	0.8872	0.4262	0.0157	0.0214	0.0235
	RMSE	0.0402	0.0394	0.0307	1.3183	0.9376	0.1278	0.1463	0.1535
(4,1.1,1.0)	Bias	0.0069	0.0051	0.0052	-0.1042	-0.3218	-0.0101	0.2134	0.2212
	Var	0.0010	0.0009	0.0006	0.1956	0.2176	0.0059	0.0240	0.0284
	RMSE	0.0323	0.0308	0.0253	0.4544	0.5667	0.0772	0.2638	0.2781
(4,1.1,2.0)	Bias	0.0044	0.0031	0.0031	-0.1229	-0.3367	-0.0102	0.2134	0.2212
	Var	0.0011	0.0010	0.0007	0.2216	0.2408	0.0067	0.0240	0.0284
	RMSE	0.0330	0.0319	0.0258	0.4865	0.5951	0.0823	0.2638	0.2781

Part A: $N = 100$, (continued)

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(6,0.5,1.0)	Bias	-0.0087	-0.0244	-0.0200	-0.0274	0.0159	0.0073	0.0043	0.0008
	Var	0.0168	0.0269	0.0076	0.0103	0.0187	0.0096	0.0087	0.0055
	RMSE	0.1300	0.1659	0.0893	0.1053	0.1375	0.0984	0.0934	0.0739
(6,0.5,2.0)	Bias	-0.0027	-0.0206	-0.0195	-0.0336	0.0303	0.0256	0.0043	0.0008
	Var	0.0175	0.0291	0.0110	0.0128	0.0264	0.0137	0.0087	0.0055
	RMSE	0.1324	0.1719	0.1067	0.1180	0.1653	0.1197	0.0934	0.0739
(6,0.8,1.0)	Bias	-0.0369	-0.0469	-0.0420	-0.0645	-0.0467	-0.0207	0.0028	-0.0011
	Var	0.0068	0.0091	0.0030	0.0184	0.0238	0.0099	0.0095	0.0060
	RMSE	0.0905	0.1065	0.0692	0.1501	0.1613	0.1018	0.0974	0.0772
(6,0.8,2.0)	Bias	-0.0345	-0.0474	-0.0435	-0.0780	-0.0519	-0.0117	0.0028	-0.0011
	Var	0.0082	0.0110	0.0041	0.0226	0.0311	0.0126	0.0095	0.0060
	RMSE	0.0968	0.1152	0.0776	0.1693	0.1839	0.1129	0.0974	0.0772
(6,1.0,1.0)	Bias	-0.0009	-0.0015	-0.0023	-0.7656	-0.4026	-0.0308	-0.0020	-0.0050
	Var	0.0010	0.0009	0.0004	0.1884	0.2119	0.0086	0.0099	0.0049
	RMSE	0.0320	0.0300	0.0213	0.8800	0.6116	0.0979	0.0995	0.0703
(6,1.0,2.0)	Bias	-0.0004	-0.0014	-0.0024	-0.7659	-0.3873	-0.0312	-0.0020	-0.0050
	Var	0.0011	0.0010	0.0005	0.1907	0.2067	0.0100	0.0099	0.0049
	RMSE	0.0334	0.0311	0.0221	0.8816	0.5972	0.1049	0.0995	0.0703
(6,1.1,1.0)	Bias	0.0031	0.0020	0.0017	-0.0820	-0.1187	-0.0144	0.2552	0.1999
	Var	0.0005	0.0004	0.0003	0.0243	0.0372	0.0039	0.0138	0.0058
	RMSE	0.0226	0.0211	0.0161	0.1763	0.2264	0.0641	0.2809	0.2140
(6,1.1,2.0)	Bias	0.0018	0.0012	0.0006	-0.0899	-0.1264	-0.0151	0.2552	0.1999
	Var	0.0006	0.0005	0.0003	0.0263	0.0393	0.0042	0.0138	0.0058
	RMSE	0.0239	0.0226	0.0168	0.1855	0.2351	0.0662	0.2809	0.2140

Part B: $N = 500$

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(4,0.5,1.0)	Bias	-0.0189	-0.0220	-0.0254	-0.0039	0.0002	-0.0062	0.0038	0.0032
	Var	0.0059	0.0064	0.0019	0.0068	0.0069	0.0028	0.0043	0.0042
	RMSE	0.0794	0.0830	0.0507	0.0823	0.0830	0.0531	0.0658	0.0652
(4,0.5,2.0)	Bias	-0.0091	-0.0136	-0.0185	-0.0060	0.0003	-0.0067	0.0038	0.0032
	Var	0.0077	0.0080	0.0034	0.0090	0.0093	0.0032	0.0043	0.0042
	RMSE	0.0883	0.0905	0.0608	0.0950	0.0964	0.0570	0.0658	0.0652
(4,0.8,1.0)	Bias	-0.0162	-0.0181	-0.0181	-0.0076	-0.0112	-0.0118	0.0028	0.0018
	Var	0.0014	0.0014	0.0008	0.0144	0.0144	0.0033	0.0042	0.0045
	RMSE	0.0414	0.0415	0.0338	0.1203	0.1204	0.0587	0.0651	0.0675
(4,0.8,2.0)	Bias	-0.0146	-0.0159	-0.0160	-0.0107	-0.0155	-0.0145	0.0028	0.0018
	Var	0.0017	0.0017	0.0010	0.0181	0.0180	0.0037	0.0042	0.0045
	RMSE	0.0434	0.0437	0.0352	0.1349	0.1351	0.0626	0.0651	0.0675
(4,1.0,1.0)	Bias	-0.0011	-0.0007	-0.0008	-0.8922	-0.7708	-0.0050	-0.0004	-0.0011
	Var	0.0003	0.0002	0.0002	0.7861	0.3515	0.0010	0.0042	0.0043
	RMSE	0.0170	0.0157	0.0131	1.2578	0.9725	0.0323	0.0652	0.0655
(4,1.0,2.0)	Bias	-0.0009	-0.0007	-0.0008	-0.8855	-0.7764	-0.0051	-0.0004	-0.0011
	Var	0.0003	0.0003	0.0002	0.7978	0.3250	0.0011	0.0042	0.0043
	RMSE	0.0173	0.0163	0.0135	1.2577	0.9632	0.0333	0.0652	0.0655
(4,1.1,1.0)	Bias	0.0046	0.0044	0.0043	-0.0055	-0.0532	-0.0030	0.2128	0.2110
	Var	0.0002	0.0002	0.0001	0.0257	0.0305	0.0004	0.0051	0.0051
	RMSE	0.0141	0.0131	0.0119	0.1604	0.1827	0.0203	0.2244	0.2228
(4,1.1,2.0)	Bias	0.0031	0.0033	0.0032	-0.0063	-0.0600	-0.0026	0.2128	0.2110
	Var	0.0002	0.0002	0.0001	0.0295	0.0362	0.0004	0.0051	0.0051
	RMSE	0.0146	0.0138	0.0120	0.1719	0.1996	0.0208	0.2244	0.2228

Part B:$N = 500$ (continued)

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(6,0.5,1.0)	Bias	-0.0018	-0.0062	-0.0104	-0.0040	0.0076	-0.0046	0.0001	-0.0007
	Var	0.0055	0.0056	0.0014	0.0020	0.0024	0.0010	0.0019	0.0012
	RMSE	0.0740	0.0750	0.0392	0.0448	0.0499	0.0314	0.0433	0.0340
(6,0.5,2.0)	Bias	0.0017	-0.0032	-0.0076	-0.0048	0.0122	-0.0032	0.0001	-0.0007
	Var	0.0059	0.0063	0.0022	0.0025	0.0034	0.0011	0.0019	0.0012
	RMSE	0.0768	0.0792	0.0478	0.0501	0.0599	0.0333	0.0433	0.0340
(6,0.8,1.0)	Bias	-0.0159	-0.0169	-0.0169	-0.0116	-0.0054	-0.0127	-0.0001	-0.0009
	Var	0.0016	0.0016	0.0007	0.0034	0.0035	0.0012	0.0020	0.0012
	RMSE	0.0427	0.0429	0.0307	0.0593	0.0597	0.0371	0.0442	0.0352
(6,0.8,2.0)	Bias	-0.0136	-0.0148	-0.0154	-0.0138	-0.0059	-0.0143	-0.0001	-0.0009
	Var	0.0018	0.0018	0.0008	0.0041	0.0044	0.0014	0.0020	0.0012
	RMSE	0.0447	0.0451	0.0328	0.0657	0.0667	0.0395	0.0442	0.0352
(6,1.0,1.0)	Bias	0.0006	0.0000	-0.0003	-0.7941	-0.6217	-0.0075	0.0003	-0.0001
	Var	0.0002	0.0002	0.0001	0.1827	0.1723	0.0007	0.0020	0.0010
	RMSE	0.0124	0.0125	0.0086	0.9018	0.7475	0.0268	0.0450	0.0323
(6,1.0,2.0)	Bias	0.0003	-0.0000	-0.0003	-0.7901	-0.6046	-0.0079	0.0003	-0.0001
	Var	0.0002	0.0002	0.0001	0.1808	0.1751	0.0007	0.0020	0.0010
	RMSE	0.0132	0.0131	0.0090	0.8972	0.7354	0.0283	0.0450	0.0323
(6,1.1,1.0)	Bias	0.0018	0.0018	0.0016	-0.0112	-0.0359	-0.0030	0.2590	0.2044
	Var	0.0001	0.0001	0.0000	0.0033	0.0045	0.0002	0.0029	0.0011
	RMSE	0.0092	0.0090	0.0066	0.0582	0.0760	0.0148	0.2645	0.2072
(6,1.1,2.0)	Bias	0.0017	0.0014	0.0013	-0.0135	-0.0386	-0.0025	0.2590	0.2044
	Var	0.0001	0.0001	0.0000	0.0036	0.0049	0.0002	0.0029	0.0011
	RMSE	0.0098	0.0097	0.0070	0.0613	0.0798	0.0152	0.2645	0.2072

Part C: $N = 1000$

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(4,0.5,1.0)	Bias	-0.0218	-0.0208	-0.0221	-0.0028	0.0000	-0.0037	-0.0003	-0.0003
	Var	0.0039	0.0037	0.0013	0.0032	0.0032	0.0014	0.0022	0.0020
	RMSE	0.0659	0.0640	0.0417	0.0565	0.0563	0.0377	0.0466	0.0442
(4,0.5,2.0)	Bias	-0.0149	-0.0141	-0.0166	-0.0038	0.0002	-0.0043	-0.0003	-0.0003
	Var	0.0048	0.0045	0.0019	0.0043	0.0043	0.0016	0.0022	0.0020
	RMSE	0.0709	0.0687	0.0470	0.0659	0.0655	0.0408	0.0466	0.0442
(4,0.8,1.0)	Bias	-0.0128	-0.0126	-0.0116	-0.0042	-0.0051	-0.0056	0.0000	-0.0002
	Var	0.0009	0.0008	0.0005	0.0066	0.0065	0.0015	0.0021	0.0021
	RMSE	0.0327	0.0307	0.0244	0.0815	0.0809	0.0395	0.0458	0.0460
(4,0.8,2.0)	Bias	-0.0123	-0.0122	-0.0115	-0.0054	-0.0067	-0.0071	0.0000	-0.0002
	Var	0.0011	0.0010	0.0007	0.0083	0.0080	0.0017	0.0021	0.0021
	RMSE	0.0347	0.0336	0.0280	0.0910	0.0899	0.0418	0.0458	0.0460
(4,1.0,1.0)	Bias	-0.0007	-0.0005	-0.0002	-0.8818	-0.8520	-0.0016	-0.0006	-0.0006
	Var	0.0001	0.0001	0.0001	0.9145	0.3348	0.0004	0.0021	0.0021
	RMSE	0.0114	0.0107	0.0086	1.3008	1.0299	0.0203	0.0464	0.0463
(4,1.0,2.0)	Bias	-0.0008	-0.0006	-0.0003	-0.8819	-0.8474	-0.0016	-0.0006	-0.0006
	Var	0.0002	0.0001	0.0001	0.9095	0.3615	0.0004	0.0021	0.0021
	RMSE	0.0123	0.0114	0.0091	1.2989	1.0390	0.0206	0.0464	0.0463
(4,1.1,1.0)	Bias	0.0032	0.0036	0.0036	-0.0044	-0.0219	-0.0008	0.2134	0.2107
	Var	0.0001	0.0001	0.0001	0.0107	0.0123	0.0002	0.0025	0.0027
	RMSE	0.0101	0.0096	0.0088	0.1035	0.1130	0.0129	0.2192	0.2171
(4,1.1,2.0)	Bias	0.0020	0.0026	0.0028	-0.0054	-0.0246	-0.0006	0.2134	0.2107
	Var	0.0001	0.0001	0.0001	0.0121	0.0143	0.0002	0.0025	0.0027
	RMSE	0.0107	0.0100	0.0089	0.1102	0.1220	0.0131	0.2192	0.2171

Part C: $N = 1000$ (continued)

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(6,0.5,1.0)	Bias	-0.0075	-0.0059	-0.0073	-0.0013	0.0044	-0.0018	0.0000	0.0004
	Var	0.0030	0.0027	0.0007	0.0010	0.0011	0.0005	0.0009	0.0005
	RMSE	0.0556	0.0524	0.0278	0.0312	0.0327	0.0218	0.0294	0.0231
(6,0.5,2.0)	Bias	-0.0040	-0.0034	-0.0055	-0.0022	0.0062	-0.0016	0.0000	0.0004
	Var	0.0035	0.0031	0.0011	0.0012	0.0014	0.0005	0.0009	0.0005
	RMSE	0.0596	0.0555	0.0340	0.0351	0.0383	0.0228	0.0294	0.0231
(6,0.8,1.0)	Bias	-0.0123	-0.0124	-0.0118	-0.0032	-0.0005	-0.0060	0.0003	0.0005
	Var	0.0008	0.0007	0.0003	0.0017	0.0016	0.0005	0.0009	0.0006
	RMSE	0.0303	0.0294	0.0221	0.0413	0.0406	0.0241	0.0303	0.0240
(6,0.8,2.0)	Bias	-0.0114	-0.0114	-0.0114	-0.0046	-0.0008	-0.0078	0.0003	0.0005
	Var	0.0009	0.0009	0.0005	0.0021	0.0020	0.0006	0.0009	0.0006
	RMSE	0.0322	0.0320	0.0248	0.0458	0.0452	0.0262	0.0303	0.0240
(6,1.0,1.0)	Bias	-0.0007	-0.0006	-0.0004	-0.8059	-0.6949	-0.0035	-0.0001	0.0001
	Var	0.0001	0.0001	0.0000	0.1849	0.1577	0.0002	0.0010	0.0005
	RMSE	0.0096	0.0090	0.0064	0.9134	0.8004	0.0150	0.0313	0.0221
(6,1.0,2.0)	Bias	-0.0009	-0.0006	-0.0004	-0.8058	-0.6770	-0.0036	-0.0001	0.0001
	Var	0.0001	0.0001	0.0000	0.1866	0.1603	0.0002	0.0010	0.0005
	RMSE	0.0100	0.0095	0.0066	0.9143	0.7865	0.0157	0.0313	0.0221
(6,1.1,1.0)	Bias	0.0010	0.0010	0.0012	-0.0087	-0.0199	-0.0017	0.2566	0.2033
	Var	0.0001	0.0000	0.0000	0.0017	0.0020	0.0001	0.0014	0.0005
	RMSE	0.0071	0.0068	0.0055	0.0419	0.0487	0.0088	0.2593	0.2046
(6,1.1,2.0)	Bias	0.0001	0.0002	0.0005	-0.0097	-0.0209	-0.0014	0.2566	0.2033
	Var	0.0001	0.0001	0.0000	0.0019	0.0022	0.0001	0.0014	0.0005
	RMSE	0.0080	0.0074	0.0058	0.0441	0.0510	0.0090	0.2593	0.2046

Table 3: Efficiency comparison under heterogenous individual effects

Note: Note for Table 1 applies here except that one half of random effect variables have variance $0.5k$, and the other half $1.5k$.

Part A: $N = 100$

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(4,0.5,1.0)	Bias	-0.0158	-0.0310	-0.0364	-0.0174	0.0134	-0.0041	0.0051	0.0039
	Var	0.0112	0.0144	0.0080	0.0368	0.0343	0.0144	0.0151	0.0149
	RMSE	0.1068	0.1239	0.0966	0.1927	0.1857	0.1201	0.1232	0.1219
(4,0.5,2.0)	Bias	-0.0075	-0.0218	-0.0290	-0.0229	0.0208	0.0036	0.0051	0.0039
	Var	0.0140	0.0176	0.0104	0.0506	0.0455	0.0171	0.0151	0.0149
	RMSE	0.1185	0.1343	0.1061	0.2261	0.2143	0.1307	0.1232	0.1219
(4,0.8,1.0)	Bias	-0.0314	-0.0324	-0.0327	-0.0454	-0.0212	-0.0219	0.0025	0.0015
	Var	0.0043	0.0041	0.0030	0.0815	0.0599	0.0215	0.0172	0.0174
	RMSE	0.0727	0.0719	0.0640	0.2891	0.2456	0.1483	0.1310	0.1318
(4,0.8,2.0)	Bias	-0.0232	-0.0304	-0.0302	-0.0601	-0.0281	-0.0203	0.0025	0.0015
	Var	0.0049	0.0052	0.0038	0.1054	0.0728	0.0243	0.0172	0.0174
	RMSE	0.0738	0.0781	0.0690	0.3302	0.2713	0.1571	0.1310	0.1318
(4,1.0,1.0)	Bias	-0.0015	-0.0040	-0.0044	-0.8814	-0.4670	-0.0180	-0.0046	-0.0044
	Var	0.0011	0.0011	0.0010	0.8468	0.5386	0.0190	0.0193	0.0197
	RMSE	0.0337	0.0327	0.0313	1.2743	0.8699	0.1391	0.1391	0.1405
(4,1.0,2.0)	Bias	-0.0032	-0.0042	-0.0046	-0.8945	-0.4467	-0.0182	-0.0046	-0.0044
	Var	0.0013	0.0012	0.0011	1.0443	0.5032	0.0201	0.0193	0.0197
	RMSE	0.0360	0.0344	0.0329	1.3581	0.8383	0.1428	0.1391	0.1405
(4,1.1,1.0)	Bias	0.0028	0.0011	0.0005	-0.1279	-0.1720	-0.0077	0.2039	0.2030
	Var	0.0008	0.0007	0.0006	0.8555	0.1636	0.0097	0.0230	0.0235
	RMSE	0.0275	0.0261	0.0250	0.9337	0.4395	0.0989	0.2541	0.2543
(4,1.1,2.0)	Bias	0.0012	-0.0006	-0.0012	-0.1332	-0.1817	-0.0072	0.2039	0.2030
	Var	0.0008	0.0007	0.0007	0.3281	0.1990	0.0105	0.0230	0.0235
	RMSE	0.0285	0.0272	0.0261	0.5880	0.4817	0.1028	0.2541	0.2543

Part A: $N = 100$, (continued)

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(6,0.5,1.0)	Bias	0.0063	-0.0048	-0.0156	-0.0408	0.0329	0.0129	0.0042	0.0014
	Var	0.0115	0.0118	0.0070	0.0115	0.0168	0.0103	0.0079	0.0044
	RMSE	0.1076	0.1087	0.0852	0.1147	0.1337	0.1021	0.0888	0.0666
(6,0.5,2.0)	Bias	0.0096	-0.0009	-0.0132	-0.0498	0.0525	0.0314	0.0042	0.0014
	Var	0.0122	0.0135	0.0089	0.0142	0.0237	0.0148	0.0079	0.0044
	RMSE	0.1108	0.1161	0.0954	0.1292	0.1627	0.1258	0.0888	0.0666
(6,0.8,1.0)	Bias	-0.0225	-0.0289	-0.0304	-0.0865	-0.0144	-0.0090	0.0042	0.0012
	Var	0.0034	0.0035	0.0020	0.0224	0.0189	0.0110	0.0088	0.0047
	RMSE	0.0628	0.0655	0.0537	0.1727	0.1384	0.1052	0.0942	0.0684
(6,0.8,2.0)	Bias	-0.0179	-0.0292	-0.0322	-0.1055	-0.0143	-0.0030	0.0042	0.0012
	Var	0.0044	0.0046	0.0028	0.0278	0.0227	0.0133	0.0088	0.0047
	RMSE	0.0687	0.0740	0.0622	0.1973	0.1515	0.1153	0.0942	0.0684
(6,1.0,1.0)	Bias	0.0003	-0.0008	-0.0013	-0.8234	-0.2980	-0.0194	0.0034	-0.0002
	Var	0.0006	0.0005	0.0004	0.2023	0.2043	0.0097	0.0104	0.0051
	RMSE	0.0244	0.0224	0.0205	0.9383	0.5414	0.1004	0.1020	0.0715
(6,1.0,2.0)	Bias	0.0013	-0.0008	-0.0014	-0.8201	-0.2853	-0.0203	0.0034	-0.0002
	Var	0.0006	0.0005	0.0005	0.2016	0.1914	0.0102	0.0104	0.0051
	RMSE	0.0246	0.0232	0.0214	0.9350	0.5223	0.1033	0.1020	0.0715
(6,1.1,1.0)	Bias	0.0030	0.0017	0.0015	-0.0926	-0.0600	-0.0066	0.2590	0.2015
	Var	0.0003	0.0003	0.0002	0.0270	0.0232	0.0044	0.0137	0.0052
	RMSE	0.0173	0.0165	0.0152	0.1885	0.1637	0.0665	0.2843	0.2139
(6,1.1,2.0)	Bias	0.0021	0.0007	0.0004	-0.0995	-0.0668	-0.0079	0.2590	0.2015
	Var	0.0003	0.0003	0.0002	0.0282	0.0246	0.0048	0.0137	0.0052
	RMSE	0.0179	0.0169	0.0154	0.1953	0.1705	0.0694	0.2843	0.2139

Part B: $N = 500$

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(4,0.5,1.0)	Bias	-0.0205	-0.0208	-0.0220	-0.0011	0.0048	-0.0028	0.0002	-0.0007
	Var	0.0033	0.0030	0.0018	0.0067	0.0065	0.0027	0.0033	0.0032
	RMSE	0.0607	0.0582	0.0479	0.0821	0.0811	0.0522	0.0577	0.0567
(4,0.5,2.0)	Bias	-0.0139	-0.0152	-0.0178	-0.0008	0.0077	-0.0027	0.0002	-0.0007
	Var	0.0043	0.0039	0.0025	0.0093	0.0090	0.0031	0.0033	0.0032
	RMSE	0.0673	0.0641	0.0533	0.0967	0.0954	0.0561	0.0577	0.0567
(4,0.8,1.0)	Bias	-0.0137	-0.0149	-0.0141	-0.0059	-0.0042	-0.0055	0.0003	0.0000
	Var	0.0008	0.0008	0.0006	0.0138	0.0126	0.0033	0.0038	0.0038
	RMSE	0.0318	0.0315	0.0278	0.1175	0.1121	0.0574	0.0614	0.0616
(4,0.8,2.0)	Bias	-0.0130	-0.0133	-0.0127	-0.0064	-0.0043	-0.0070	0.0003	0.0000
	Var	0.0011	0.0009	0.0007	0.0177	0.0158	0.0036	0.0038	0.0038
	RMSE	0.0350	0.0327	0.0294	0.1334	0.1258	0.0607	0.0614	0.0616
(4,1.0,1.0)	Bias	-0.0001	-0.0006	-0.0005	-0.8904	-0.6759	0.0006	0.0006	0.0006
	Var	0.0002	0.0002	0.0002	0.8617	0.4098	0.0011	0.0041	0.0041
	RMSE	0.0143	0.0136	0.0127	1.2863	0.9310	0.0331	0.0639	0.0641
(4,1.0,2.0)	Bias	-0.0002	-0.0007	-0.0006	-0.8626	-0.6674	0.0006	0.0006	0.0006
	Var	0.0002	0.0002	0.0002	1.5980	0.4218	0.0012	0.0041	0.0041
	RMSE	0.0152	0.0144	0.0135	1.5304	0.9313	0.0339	0.0639	0.0641
(4,1.1,1.0)	Bias	0.0042	0.0035	0.0033	-0.0110	-0.0378	0.0009	0.2137	0.2110
	Var	0.0001	0.0001	0.0001	0.0226	0.0296	0.0004	0.0046	0.0046
	RMSE	0.0120	0.0113	0.0107	0.1507	0.1761	0.0212	0.2243	0.2216
(4,1.1,2.0)	Bias	0.0030	0.0026	0.0026	-0.0147	-0.0434	0.0013	0.2137	0.2110
	Var	0.0001	0.0001	0.0001	0.0262	0.0335	0.0005	0.0046	0.0046
	RMSE	0.0124	0.0119	0.0113	0.1624	0.1880	0.0217	0.2243	0.2216

Part B:$N = 500$ (continued)

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(6,0.5,1.0)	Bias	-0.0069	-0.0046	-0.0078	-0.0046	0.0083	0.0001	0.0034	0.0018
	Var	0.0027	0.0025	0.0013	0.0020	0.0024	0.0010	0.0016	0.0009
	RMSE	0.0526	0.0504	0.0369	0.0455	0.0493	0.0318	0.0405	0.0307
(6,0.5,2.0)	Bias	-0.0052	-0.0025	-0.0062	-0.0061	0.0130	0.0015	0.0034	0.0018
	Var	0.0032	0.0029	0.0017	0.0026	0.0033	0.0011	0.0016	0.0009
	RMSE	0.0572	0.0543	0.0420	0.0510	0.0589	0.0335	0.0405	0.0307
(6,0.8,1.0)	Bias	-0.0146	-0.0141	-0.0140	-0.0118	-0.0017	-0.0048	0.0032	0.0007
	Var	0.0007	0.0007	0.0005	0.0035	0.0035	0.0012	0.0019	0.0010
	RMSE	0.0304	0.0298	0.0260	0.0607	0.0596	0.0355	0.0432	0.0319
(6,0.8,2.0)	Bias	-0.0139	-0.0125	-0.0129	-0.0147	-0.0016	-0.0058	0.0032	0.0007
	Var	0.0010	0.0009	0.0006	0.0044	0.0044	0.0014	0.0019	0.0010
	RMSE	0.0338	0.0322	0.0282	0.0679	0.0664	0.0384	0.0432	0.0319
(6,1.0,1.0)	Bias	-0.0012	-0.0006	-0.0006	-0.8205	-0.4350	-0.0022	0.0025	-0.0004
	Var	0.0001	0.0001	0.0001	0.1940	0.1943	0.0008	0.0020	0.0010
	RMSE	0.0099	0.0097	0.0087	0.9313	0.6193	0.0286	0.0449	0.0312
(6,1.0,2.0)	Bias	-0.0014	-0.0007	-0.0007	-0.8207	-0.4271	-0.0022	0.0025	-0.0004
	Var	0.0001	0.0001	0.0001	0.1972	0.1950	0.0009	0.0020	0.0010
	RMSE	0.0107	0.0101	0.0091	0.9331	0.6143	0.0301	0.0449	0.0312
(6,1.1,1.0)	Bias	0.0013	0.0013	0.0013	-0.0145	-0.0263	-0.0003	0.2588	0.2028
	Var	0.0001	0.0000	0.0000	0.0035	0.0043	0.0003	0.0026	0.0009
	RMSE	0.0073	0.0071	0.0064	0.0609	0.0706	0.0169	0.2638	0.2051
(6,1.1,2.0)	Bias	0.0006	0.0007	0.0008	-0.0162	-0.0282	0.0002	0.2588	0.2028
	Var	0.0001	0.0001	0.0000	0.0039	0.0046	0.0003	0.0026	0.0009
	RMSE	0.0079	0.0075	0.0067	0.0641	0.0736	0.0177	0.2638	0.2051

Part C: $N = 1000$

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(4,0.5,1.0)	Bias	-0.0126	-0.0133	-0.0159	-0.0007	0.0020	-0.0010	0.0019	0.0015
	Var	0.0019	0.0016	0.0011	0.0033	0.0033	0.0014	0.0018	0.0016
	RMSE	0.0449	0.0426	0.0361	0.0578	0.0571	0.0375	0.0421	0.0406
(4,0.5,2.0)	Bias	-0.0070	-0.0090	-0.0130	-0.0012	0.0027	-0.0011	0.0019	0.0015
	Var	0.0026	0.0022	0.0015	0.0045	0.0044	0.0016	0.0018	0.0016
	RMSE	0.0511	0.0476	0.0413	0.0672	0.0665	0.0401	0.0421	0.0406
(4,0.8,1.0)	Bias	-0.0081	-0.0096	-0.0094	-0.0024	-0.0018	-0.0027	0.0017	0.0016
	Var	0.0005	0.0004	0.0003	0.0070	0.0066	0.0016	0.0020	0.0020
	RMSE	0.0230	0.0222	0.0199	0.0837	0.0814	0.0404	0.0446	0.0445
(4,0.8,2.0)	Bias	-0.0073	-0.0086	-0.0086	-0.0034	-0.0028	-0.0038	0.0017	0.0016
	Var	0.0006	0.0005	0.0004	0.0087	0.0082	0.0018	0.0020	0.0020
	RMSE	0.0247	0.0243	0.0221	0.0935	0.0904	0.0422	0.0446	0.0445
(4,1.0,1.0)	Bias	-0.0001	-0.0005	-0.0005	-0.9323	-0.7714	-0.0003	0.0008	0.0009
	Var	0.0001	0.0001	0.0001	0.8102	0.4474	0.0005	0.0021	0.0021
	RMSE	0.0095	0.0092	0.0087	1.2959	1.0211	0.0217	0.0455	0.0456
(4,1.0,2.0)	Bias	-0.0002	-0.0006	-0.0005	-0.9289	-0.7675	-0.0003	0.0008	0.0009
	Var	0.0001	0.0001	0.0001	0.8938	0.4459	0.0005	0.0021	0.0021
	RMSE	0.0100	0.0097	0.0092	1.3254	1.0174	0.0219	0.0455	0.0456
(4,1.1,1.0)	Bias	0.0029	0.0029	0.0027	-0.0074	-0.0206	0.0000	0.2138	0.2110
	Var	0.0001	0.0001	0.0001	0.0112	0.0131	0.0002	0.0023	0.0023
	RMSE	0.0083	0.0081	0.0078	0.1061	0.1162	0.0139	0.2192	0.2164
(4,1.1,2.0)	Bias	0.0024	0.0021	0.0020	-0.0090	-0.0231	0.0001	0.2138	0.2110
	Var	0.0001	0.0001	0.0001	0.0127	0.0149	0.0002	0.0023	0.0023
	RMSE	0.0090	0.0083	0.0079	0.1131	0.1242	0.0140	0.2192	0.2164

Part C: $N = 1000$ (continued)

(T, α, k)		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE
(6,0.5,1.0)	Bias	-0.0043	-0.0030	-0.0059	-0.0022	0.0041	-0.0003	0.0016	0.0009
	Var	0.0014	0.0013	0.0006	0.0010	0.0011	0.0005	0.0008	0.0005
	RMSE	0.0380	0.0362	0.0260	0.0322	0.0335	0.0225	0.0285	0.0221
(6,0.5,2.0)	Bias	-0.0025	-0.0012	-0.0045	-0.0028	0.0064	0.0002	0.0016	0.0009
	Var	0.0016	0.0015	0.0009	0.0013	0.0015	0.0006	0.0008	0.0005
	RMSE	0.0403	0.0392	0.0301	0.0363	0.0392	0.0235	0.0285	0.0221
(6,0.8,1.0)	Bias	-0.0106	-0.0094	-0.0094	-0.0057	-0.0005	-0.0033	0.0015	0.0003
	Var	0.0004	0.0004	0.0003	0.0017	0.0017	0.0006	0.0009	0.0005
	RMSE	0.0235	0.0218	0.0187	0.0418	0.0411	0.0244	0.0305	0.0227
(6,0.8,2.0)	Bias	-0.0087	-0.0079	-0.0082	-0.0071	-0.0006	-0.0043	0.0015	0.0003
	Var	0.0006	0.0005	0.0004	0.0021	0.0021	0.0007	0.0009	0.0005
	RMSE	0.0262	0.0240	0.0210	0.0467	0.0459	0.0259	0.0305	0.0227
(6,1.0,1.0)	Bias	-0.0003	-0.0001	-0.0001	-0.8187	-0.4923	-0.0014	0.0013	-0.0001
	Var	0.0000	0.0000	0.0000	0.1903	0.1889	0.0003	0.0010	0.0005
	RMSE	0.0070	0.0067	0.0058	0.9277	0.6567	0.0164	0.0317	0.0224
(6,1.0,2.0)	Bias	-0.0004	-0.0001	-0.0001	-0.8172	-0.4867	-0.0015	0.0013	-0.0001
	Var	0.0001	0.0000	0.0000	0.1880	0.1869	0.0003	0.0010	0.0005
	RMSE	0.0073	0.0070	0.0061	0.9251	0.6510	0.0169	0.0317	0.0224
(6,1.1,1.0)	Bias	0.0011	0.0010	0.0011	-0.0072	-0.0142	-0.0003	0.2587	0.2037
	Var	0.0000	0.0000	0.0000	0.0016	0.0019	0.0001	0.0013	0.0005
	RMSE	0.0056	0.0054	0.0050	0.0405	0.0461	0.0096	0.2612	0.2049
(6,1.1,2.0)	Bias	0.0004	0.0005	0.0006	-0.0082	-0.0150	-0.0001	0.2587	0.2037
	Var	0.0000	0.0000	0.0000	0.0018	0.0021	0.0001	0.0013	0.0005
	RMSE	0.0060	0.0058	0.0053	0.0427	0.0481	0.0097	0.2612	0.2049

Table 4: Frequencies of best and second-best performers in terms of RMSE

		IIE	CSMLE	BCPLSE	GMM1	GMM2	GMM3	FDLSE	PFAE	Total
Tbl 1	Best	0	0	42	0	0	0	0	6	48
	2nd-best	11	26	1	0	0	5	1	4	48
Tbl 2	Best	0	0	38	0	0	6	0	4	48
	2nd-best	6	25	4	0	0	3	1	9	48
Tbl 3	Best	0	0	41	0	0	1	0	6	48
	2nd-best	4	31	1	0	0	6	1	5	48
All	Best	0	0	121	0	0	7	0	16	144
	2nd-best	21	82	6	0	0	14	3	18	144

Table 5: Empirical coverage ratios of bootstrap confidence intervals of the PAR(1) coefficient at $N = 100$

Note: Data were generated as for Table 1. Empirical coverage ratios are based on 300 iterations. The number of bootstrap iterations is set at 1000.

(T, α, k)	95%	90%
(4,0.5,1.0)	0.97	0.91
(4,0.5,2.0)	0.95	0.88
(4,0.8,1.0)	0.94	0.88
(4,0.8,2.0)	0.95	0.90
(4,1.0,1.0)	0.93	0.87
(4,1.0,2.0)	0.94	0.88
(4,1.1,1.0)	0.93	0.87
(4,1.1,2.0)	0.94	0.86
(6,0.5,1.0)	0.93	0.88
(6,0.5,2.0)	0.94	0.88
(6,0.8,1.0)	0.93	0.86
(6,0.8,2.0)	0.92	0.89
(6,1.0,1.0)	0.93	0.89
(6,1.0,2.0)	0.94	0.89
(6,1.1,1.0)	0.94	0.89
(6,1.1,2.0)	0.94	0.89