Hacettepe Journal of Mathematics and Statistics
Volume 40 (1) (2011), 105 – 114

BIAS CORRECTION ESTIMATOR FOR A DYNAMIC PANEL DATA MODEL WITH FIXED EFFECTS USING AN ITERATED BOOTSTRAP

Gang Yu*†‡, Wei Gao* and Ning-Zhong Shi*

Received 12 : 07 : 2009 : Accepted 21 : 06 : 2010

Abstract

A bias correction estimator (BCE) for a dynamic panel data model with fixed effects is given, based on the alternating iterative maximum likelihood estimator (AIMLE). The new estimator is asymptotically unbiased and consistent. Monte Carlo studies are conducted to evaluate the finite sample properties of the MLE, AIMLE and BCE. It is shown that the BCE based on AIMLE appears to dominate the AIMLE approach both in terms of the median bias (Bias) and median absolute error (MAE) of the estimators.

Keywords: Bias correction estimator (BCE), Panel data, Fixed effects, Maximum likelihood estimator (MLE), Alternating iterative maximum likelihood estimator (AIMLE), Asymptotic normality.

2000 AMS Classification: 62 F 10, 62 F 12, 62 F 40.

1. Introduction

In this paper, we consider the estimation of a dynamic panel data model with fixed effects. The model is in the following form:

$$y_{it} = \gamma y_{it-1} + \eta_i + \varepsilon_{it},\ i = 1, \dots, N;\quad t = 1, \dots, T. \tag{1}$$

Here, y_{it} is an observed dependent variable, ε_{it} is an unobservable error term which is iid across units and time periods with normal distribution $N(0, \sigma^2)$, η_i are fixed effects, as parameters to be estimated (nuisance parameters) and γ is the parameter of interest to

*Key Laboratory for Applied Statistics of MOE and School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin, 130024, P. R. China. E-mail: (G. Yu) `yugang543@163.com` (W. Gao) `gaow@nenu.edu.cn` (N.-Z. Shi) `shinz@nenu.edu.cn`

†School of Mathematics and Quantitative Economics, Dongbei University of Finance and Economics, Dalian, Liaoning, 116025, P. R. China.

‡Corresponding Author.

be estimated. We assume that the initial observations, y_{i0}, are observable. The number of parameters increases with the increase in the number of individuals N, the maximum likelihood estimator (MLE) would lead to inconsistent estimates of the parameter γ. This is the well-known incidental parameter problem proposed by Neyman and Scott [12].

The inconsistency of the MLE in dynamic panel data models has led to the development of a range of new estimators, see e.g., Nickell [13], Arellano and Bond [2], Arellano and Bover [3], Kiviet [10], Ahn and Schmidt [1], Blundell and Bond [4], Hansen [8], Hahn and Kuersteiner [7], Hsiao *et al.* [9], Bun and Carree [5], Bun and Carree [6].

In this paper, we give the definition of the alternating iterative maximum likelihood estimator (AIMLE), based on the alternating iterative method (AIM) proposed by Shi *et al.* [15] and derive the asymptotic normality of the AIMLE. In fact, the AIMLE is just an algorithm to compute the MLE, which can also be applied to the binary panel data model. Also, a new bias correction estimator of the AIMLE using iterative bootstrap bias correction is proposed, which is asymptotically unbiased and consistent, though the approach we use may be all that new, and was adopted in Kuk [11]. The obtained simulation results are desirable. The bias correction estimator based on AIMLE can potentially be extended to the binary panel data model.

The remainder of the paper is organized as follows. The MLE, AIMLE and the bias correction estimator (BCE) based on AIMLE are introduced in Section 2. In Section 3, we report some simulation results. Section 4 concludes the paper.

2. Bias correction estimator

The traditional approach to the bias correction of estimating model (1) is based on MLE (Hahn and Kuersteiner [7]). In this section, we will firstly give another method to compute the MLE, i.e. AIMLE, which is obtained by using the alternating iterative algorithm, and derive the asymptotic normality of the AIMLE. Secondly, we adjust the biased estimator AIMLE, and propose a bias correction estimator based on AIMLE.

The MLE $\hat{\gamma}_{MLE}$ and $\hat{\eta}_{MLE} = (\hat{\eta}_1, \ldots, \hat{\eta}_N)$ maximize the following log-likelihood function

$$l(\gamma,\eta) = \sum_{i=1}^{N} l_i(\gamma,\eta_i) = \sum_{i=1}^{N}\sum_{t=1}^{T}\{-\frac{(y_{it}-\gamma y_{it-1}-\eta_i)^2}{2\sigma^2} + \frac{1}{2}\log\frac{1}{\sigma^2}\} + c, \tag{2}$$

where c is any constant. It is well known that the MLE of γ and η_i have an explicit expression

$$\hat{\gamma}_{\mathrm{MLE}} = \frac{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} y_{it}y_{it-1} - \frac{1}{T}\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} y_{it-1}\left(\sum\limits_{t=1}^{T} y_{it}\right)}{\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} y_{it-1}^2 - \frac{1}{T}\sum\limits_{i=1}^{N}\sum\limits_{t=1}^{T} y_{it-1}\left(\sum\limits_{t=1}^{T} y_{it-1}\right)}, \tag{3}$$

$$\hat{\eta}_{i\mathrm{MLE}} = \frac{1}{T}\sum_{t=1}^{T}(y_{it} - \hat{\gamma}_{\mathrm{MLE}}y_{it-1}),\ i = 1,\ldots,N. \tag{4}$$

If $T \to \infty$, $\hat{\gamma}_{\mathrm{MLE}}$ is consistent. If T is finite and $|\gamma| < 1$, then as N tends to infinity, we know that $\hat{\gamma}_{\mathrm{MLE}}$ is inconsistent. The inconsistency of the MLE of the common parameter, $\hat{\gamma}_{\mathrm{MLE}}$, is due to the classical incidental parameters problem in which the number of parameters increases with the number of observations (Neyman and Scott [12]). The introduction of exogenous variables does not help to solve this incidental parameters problem.

Next, we will use another method to obtain the MLE of γ and η, which is not an explicit expression. If the parameter γ is given, from the first order condition of η_i, $\dfrac{\partial l_i(\gamma,\eta_i)}{\partial \eta_i} = \sum_{t=1}^{T}(y_{it} - \gamma y_{it-1} - \eta_i)$, the MLE of η_i, solves

$$\sum_{t=1}^{T} y_{it} = \sum_{t=1}^{T}(\gamma y_{it-1} + \eta_i). \tag{5}$$

If the fixed effects η_i are given, the MLE of γ can be obtained by the first condition of γ

$$\frac{\partial l(\gamma,\eta)}{\partial \gamma} = \sum_{i=1}^{N}\sum_{t=1}^{T} y_{it-1}(y_{it} - \gamma y_{it-1} - \eta_i) = 0. \tag{6}$$

According to (5) and (6), we can obtain the MLE of γ and η by the following alternating iterative algorithm, which is similar to AIM proposed by Shi *et al.* [15]:

Algorithm:

Step 0 For any fixed $\gamma^{(0)}$, find $\eta^{(0)}$, the maximum point of $l\big(\gamma^{(0)},\eta\big)$ on R^N, i.e. solve Equation (5);

Step (n,1) Fix $\eta^{(n-1)}$, find $\gamma^{(n)}$, the maximum point of $l\big(\gamma,\eta^{(n-1)}\big)$ on R, i.e. solve Equation (6);

Step (n,2) Fix $\gamma^{(n)}$, find $\eta^{(n)}$, the maximum point of $l\big(\gamma^{(n)},\eta\big)$ on R^N, i.e. solve Equation (5).

We stop the AIM at Step(n,1) if $\big|\gamma^{(n+1)} - \gamma^{(n)}\big| \le 10^{-m}$ for some integers $m \ge 1$.

From the above algorithm, we can obtain two point estimation sequences $\{\gamma^{(n)}\} \subset R$ and $\{\eta^{(n)}\} \subset R^N$. It can also be seen that, for $n \ge 1$

$$l\big(\gamma^{(n)},\eta^{(n)}\big) \le l\big(\gamma^{(n+1)},\eta^{(n)}\big) \le l\big(\gamma^{(n+1)},\eta^{(n+1)}\big). \tag{7}$$

The proof of the following Theorem 2.1 is similar to those of Shi *et al.* [15], and is given in the Appendix.

2.1. Theorem. *The point estimation sequences $\{\gamma^{(n)},\ \eta^{(n)}\}$ given in the above algorithm converge to $\hat\gamma$, $\hat\eta$, respectively.*

Since the sequences $\{\gamma^{(n)}\}$ and $\{\eta^{(n)}\}$ are obtained alternately, the estimators $\hat\gamma$ and $\hat\eta$ are also called the *alternating iterative maximum likelihood estimator* (AIMLE) which converges to the MLE. The AIMLE can be used to obtain the MLE when the MLE does not have an explicit expression.

It is well known that the profile likelihood function $l(\gamma,\hat\eta)$ is not a true likelihood function, for example $\mathrm{E}_\gamma\left\{\dfrac{\partial l(\gamma,\hat\eta)}{\partial\gamma}\right\} \ne 0$, i.e. the estimating equation obtained by using the profile likelihood function $l(\gamma,\hat\eta)$ is biased.

From a first order expansion of the concentrated score $\dfrac{\partial l(\gamma,\hat\eta)}{\partial\gamma}$ around the value $\gamma^\star$, which is not the true value γ and is called a *pseudo true value*, we obtain the usual expression for the AIMLE $\hat\gamma$,

$$H_N\sqrt{N}(\hat\gamma - \gamma^\star) = -\frac{1}{\sqrt{N}}\frac{\partial l(\gamma,\hat\eta)}{\partial\gamma}\bigg|_{\gamma=\gamma^\star} + O_p\left(\frac{1}{\sqrt{N}}\right), \tag{8}$$

where $\gamma^\star = g(\gamma)$ is contained in the following Equation (9) for given $\hat\eta$,

$$\mathrm{E}_\gamma\left\{\frac{\partial l(\gamma,\hat\eta)}{\partial\gamma}\bigg|_{\gamma=\gamma^\star}\right\} = 0 \tag{9}$$

and

$$H_N = \frac{1}{N}\frac{\partial^2 l(\gamma,\hat{\eta})}{\partial\gamma^2}\bigg|_{\gamma=\gamma^\star}. \tag{10}$$

A standard central limit theorem applies to the concentrated score $\frac{\partial l(\gamma,\hat{\eta})}{\partial\gamma}\Big|_{\gamma=\gamma^\star}$, given (9) we have

$$\frac{1}{\sqrt{N}}\frac{\partial l(\gamma,\hat{\eta})}{\partial\gamma}\bigg|_{\gamma=\gamma^\star} \xrightarrow{d} N(0,V_N), \tag{11}$$

where

$$V_N = \frac{1}{N}E_\gamma\{(\frac{\partial l(\gamma,\hat{\eta})}{\partial\gamma}|_{\gamma=\gamma^\star})^2\}. \tag{12}$$

Finally, combining (8) and (11), we give the following Theorem 2.2:

2.2. Theorem. (Asymptotic normality of AIMLE) *When the sample size $N\to\infty$,*

$$\sqrt{N}\big(\hat{\gamma}_{\text{AIMLE}}-\gamma^\star\big) \xrightarrow{d} N(0,\Lambda), \tag{13}$$

where

$$\Lambda = V_N/H_N^2. \tag{14}$$

Our approach is not to modify the biased estimating equation to make it unbiased. We will adjust the $\hat{\gamma}_{\text{AIMLE}}$, which is a biased estimator, using the iterative bootstrap bias correction method as in Kuk [11], to obtain asymptotically unbiased and consistent estimators. The estimator $\hat{\gamma}_{\text{AIMLE}}$ has an asymptotic bias given by

$$b(\gamma) = \gamma^\star - \gamma = g(\gamma) - \gamma. \tag{15}$$

The following is the bias correction procedure, proposed by Kuk [11]. Let $b^{(0)}$ be an initial estimate of the bias of $\hat{\gamma}$. In the $k+1$ step, the updated estimate of the bias of $\hat{\gamma}_{\text{AIMLE}}$ can be written as

$$b^{(k+1)} = g\big(\hat{\gamma}_{\text{AIMLE}} - b^{(k)}\big) - \big(\hat{\gamma}_{\text{AIMLE}} - b^{(k)}\big). \tag{16}$$

The $k+1$ step updated bias corrected estimate of γ can be denoted by

$$\tilde{\gamma}_{\text{BCE}}^{(k+1)} = \hat{\gamma}_{\text{AIMLE}} - b^{(k+1)}. \tag{17}$$

Assuming that the limit of $b^{(k)}$ exists, we can let $k\to\infty$ in equation (16) to obtain

$$b = g(\tilde{\gamma}_{\text{BCE}}) - (\hat{\gamma}_{\text{AIMLE}} - b), \tag{18}$$

so that

$$\tilde{\gamma}_{\text{BCE}} = g^{-1}(\hat{\gamma}_{\text{AIMLE}}). \tag{19}$$

Assuming that $g(\cdot)$ is one to one and differentiable, from the above expression (13) and Slutsky's Theorem, we can obtain the following Theorem 2.3:

2.3. Theorem. (Bias correction of AIMLE) *When $N\to\infty$,*

$$\sqrt{N}\big(\tilde{\gamma}_{\text{BCE}}-\gamma\big) \xrightarrow{d} N\big(0,\Lambda D^2\big), \tag{20}$$

where $D = \frac{dg^{-1}(\gamma)}{d\gamma}\Big|_{\gamma=\gamma^\star}$. Thus the estimator $\tilde{\gamma}_{\text{BCE}}$ defined by equation (19) *is asymptotically unbiased and consistent.*

The equations (16) and (17) are at the very core of the subject about iterative bias correction of $\hat{\gamma}_{\text{AIMLE}}$. Except for very simple problems, the function $\gamma^\star = g(\gamma)$ does not have an explicit expression, and is usually a complicated integral equation. Just because of the complexity of the function $g(\cdot)$, the implementation of iterative bias correction is difficult. From (13), we can find that $g(\gamma) = \gamma^*$ is the asymptotic mean of $\hat{\gamma}_{\text{AIMLE}}$. We propose to approximate $g(\gamma)$ by $g_M(\gamma)$, which is the average of $\hat{\gamma}_{\text{AIMLE}}$ over simulated samples. $g_M(\gamma)$ can be written as the following:

$$g_M(\gamma) = \frac{1}{M}\sum_{i=1}^{M} \hat{\gamma}_{\text{AIMLE}}(\mathbf{y}_i), \tag{21}$$

where $\mathbf{y}_1, \ldots, \mathbf{y}_M$ are simulated from the model with the parameters set at γ, $\hat{\eta}(\gamma)$ and $\hat{\sigma^2}(\gamma)$, where $\hat{\sigma^2}(\gamma) = \sum_{i=1}^{N}\sum_{t=1}^{T}\big(y_{it} - \gamma y_{it-1} - \hat{\eta}_i(\gamma)\big)^2/NT$. Replacing $g_M(\gamma)$ in equations (16) and (17), we have

$$b_M^{(k+1)} = g_M\big(\hat{\gamma}_{\text{AIMLE}} - b_M^{(k)}\big) - \big(\hat{\gamma}_{\text{AIMLE}} - b_M^{(k)}\big) \tag{22}$$

as the bootstrap estimate of the bias of $\hat{\gamma}_{\text{AIMLE}}$ at the $k+1$th iteration and

$$\tilde{\gamma}_{\text{BCE}}^{(k+1)} = \hat{\gamma}_{\text{AIMLE}} - b_M^{(k+1)} \tag{23}$$

as the updated bootstrap iterative bias correction estimator of γ.

3. A simulation study

In this section we present a small Monte Carlo study to illustrate the usefulness of our bias correction estimator. The simulation is based on the following dynamic panel data model

$$y_{it} = \gamma y_{it-1} + \eta_i + \varepsilon_{it}, \tag{24}$$

where $y_{it} \in R, \gamma \in \{0.0, 0.3, 0.6, 0.9\}$, $\eta_i \sim N(0,1)$ independent across i, and $\varepsilon_{it} \sim N(0,1)$ is independent across i and t. We generate η_i and ε_{it} such that they are independent of each other.

As in Hahn and Kuersteiner [7], the initial observations y_{i0} are assumed to be generated by the normal distribution $N\left(\frac{\eta_i}{1-\gamma}, \frac{1}{1-\gamma^2}\right)$. For each simulated sample we have reported the results for the MLE $\hat{\gamma}_{\text{MLE}}$, the AIMLE $\hat{\gamma}_{\text{AIMLE}}$ and the bias correction estimator $\tilde{\gamma}_{\text{BCE}}$. Throughout, we show the mean, median, SD (standard deviation), the median bias (Bias) and median absolute error (MAE) of these three estimator based on 100 replications for each design with different T. We compute the MLE $\hat{\gamma}_{\text{MLE}}$ based on formulation (3). Using the AIM, we can obtain the AIMLE $\hat{\gamma}_{\text{AIMLE}}$. Both MLE and AIMLE are biased estimators. Here we stop the AIM if $|\gamma^{(n+1)} - \gamma^{(n)}| \le 10^{-3}$.

The bootstrap iterative bias corrected estimator $\tilde{\gamma}_{\text{BCE}}$ is obtained from $\hat{\gamma}_{\text{AIMLE}}$ iteratively by using equations (22) and (23), with starting value $b^{(0)} = 0$, so that $\tilde{\gamma}_{\text{BCE}}^{(0)} = \hat{\gamma}_{\text{AIMLE}}$. To save computation, we set $M = 10$. Table 1 reports results for $T = 5$, 10 and 20 (to conserve space, we consider only $N = 100$). Finite sample properties of these three estimators obtained by 100 Monte Carlo runs are summarized in Table 1. The most striking feature of these results is that the $\hat{\gamma}_{\text{AIMLE}}$ is inferior in terms of median bias (Bias) and median absolute error (MAE) to the $\tilde{\gamma}_{\text{BCE}}$. We have simulated four different values of γ, namely $\gamma = 0,\ 0.3,\ 0.6,\ 0.9$. As expected, we find that as γ increases, the bias increases dramatically.

Table 1. Various estimators of γ with different T values ($N = 100$)

		T	Mean	Median	SD	Bias	MAE
True value $\gamma = 0.0$							
MLE	$\hat{\gamma}_{\mathrm{MLE}}$	5	-0.197	-0.196	0.045	-0.196	0.196
		10	-0.101	-0.104	0.030	-0.104	0.104
		20	-0.051	-0.051	0.022	-0.051	0.051
AIMLE	$\hat{\gamma}_{\mathrm{AIMLE}}$	5	-0.195	-0.195	0.045	-0.195	0.195
		10	-0.100	-0.103	0.030	-0.103	0.103
		20	-0.051	-0.050	0.022	-0.050	0.050
BCE	$\hat{\gamma}_{\mathrm{BCE}}$	5	-0.024	-0.023	0.058	-0.023	0.048
		10	-0.004	-0.006	0.034	-0.006	0.024
		20	-0.002	-0.003	0.024	-0.003	0.018
True value $\gamma = 0.3$							
MLE	$\hat{\gamma}_{\mathrm{MLE}}$	5	0.025	0.025	0.048	-0.274	0.274
		10	0.164	0.166	0.027	-0.133	0.133
		20	0.232	0.233	0.021	-0.066	0.066
AIMLE	$\hat{\gamma}_{\mathrm{AIMLE}}$	5	0.027	0.028	0.047	-0.271	0.271
		10	0.162	0.164	0.027	-0.135	0.135
		20	0.230	0.231	0.021	-0.068	0.068
BCE	$\hat{\gamma}_{\mathrm{BCE}}$	5	0.234	0.238	0.058	-0.061	0.062
		10	0.288	0.288	0.033	-0.011	0.027
		20	0.293	0.292	0.023	-0.007	0.014
True value $\gamma = 0.6$							
MLE	$\hat{\gamma}_{\mathrm{MLE}}$	5	0.234	0.234	0.052	-0.365	0.365
		10	0.420	0.422	0.030	-0.177	0.177
		20	0.513	0.511	0.019	-0.088	0.088
AIMLE	$\hat{\gamma}_{\mathrm{AIMLE}}$	5	0.225	0.225	0.053	-0.374	0.374
		10	0.414	0.416	0.031	-0.183	0.183
		20	0.508	0.505	0.019	-0.094	0.094
BCE	$\hat{\gamma}_{\mathrm{BCE}}$	5	0.418	0.417	0.059	-0.182	0.182
		10	0.551	0.551	0.034	-0.048	0.048
		20	0.582	0.579	0.022	-0.020	0.023
True value $\gamma = 0.9$							
MLE	$\hat{\gamma}_{\mathrm{MLE}}$	5	0.443	0.440	0.045	-0.459	0.459
		10	0.657	0.655	0.029	-0.244	0.244
		20	0.776	0.777	0.016	-0.122	0.122
AIMLE	$\hat{\gamma}_{\mathrm{AIMLE}}$	5	0.310	0.303	0.055	-0.596	0.596
		10	0.585	0.586	0.034	-0.313	0.313
		20	0.733	0.733	0.017	-0.166	0.166
BCE	$\hat{\gamma}_{\mathrm{BCE}}$	5	0.346	0.343	0.065	-0.556	0.556
		10	0.618	0.622	0.041	-0.277	0.277
		20	0.753	0.752	0.021	-0.147	0.147

From Table 1, we can also find that the two estimators MLE and AIMLE perform similarly, i.e. the AIMLE will converge to the MLE.

4. Conclusion

In this paper, we apply the bias correction estimator (BCE) proposed by Kuk [11] to a dynamic panel data model with fixed effects, using iterative bootstrap. It is shown that the BCE is asymptotically unbiased and consistent. Moreover, we give a definition of the alternating iterative maximum likelihood estimator (AIMLE), which is a useful algorithm to compute the MLE, and the algorithm can be applied to binary panel data. A Monte Carlo Simulation shows that the bias correction estimator (BCE) proposed here performs better than the AIMLE.

5. Appendix

The following Lemma 5.1 is well known:

5.1. Lemma. *Let $\{y_n\}$ be a bounded infinite sequence in R^k. If every convergent subsequence of $\{y_n\}$ has a common limit point $\hat{y}$, then $\{y_n\}$ converges.* □

5.2. Lemma. *For the iterative sequence $\{\gamma^{(n)}\}$, when $n \to \infty$, $|\gamma^{(n+1)} - \gamma^{(n)}| \to 0$.*

Proof. We need to prove first that $l'(\gamma^{(n+1)}, \eta^{(n)}) \times (\gamma^{(n)} - \gamma^{(n+1)}) \leq 0$ for all $n \geq 1$. If there exists $n_0 \geq 1$ such that $l'(\gamma^{(n_0+1)}, \eta^{(n_0)}) \times (\gamma^{(n_0)} - \gamma^{(n_0+1)}) > 0$, then

$$\begin{aligned} l\big(\gamma^{(n_0+1)} + t(\gamma^{(n_0)} - \gamma^{(n_0+1)}), \eta^{(n_0)}\big) &\\ = l\big(\gamma^{(n_0+1)}, \eta^{(n_0)}\big) + l'\big(\gamma^{(n_0+1)}, \eta^{(n_0)}\big) \times \big(\gamma^{(n_0)} - \gamma^{(n_0+1)}\big) + o(t) &\\ > l\big(\gamma^{(n_0+1)}, \eta^{(n_0)}\big). & \end{aligned}$$

This contradicts $l\big(\gamma^{(n_0+1)}, \eta^{(n_0)}\big) = \max_{\gamma \in R} l\big(\gamma, \eta^{(n_0)}\big)$, so the above claim is true.

Secondly, from (2), we know that $l''(\gamma, \eta_0) = -\sum_{i=1}^{N}\sum_{t=1}^{T} \frac{y_{it-1}^2}{\sigma^2} < M < 0$. So we can prove that there exists $M < 0$ such that

$$(\gamma_1 - \gamma_2)^2 \times l''(\gamma, \eta_0) \leq M \times |\gamma_1 - \gamma_2|^2,$$

for any $\gamma, \gamma_1, \gamma_1 \in R$.

Since $l(\gamma, \eta)$ has continuous second-order partial derivatives on some open set containing R^{N+1}, then

$$\begin{aligned} l\big(\gamma^{(n)}, \eta^{(n)}\big) = l\big(\gamma^{(n+1)}, \eta^{(n)}\big) + l'\big(\gamma^{(n+1)}, \eta^{(n)}\big) \times \big(\gamma^{(n)} - \gamma^{(n+1)}\big) &\\ + \frac{1}{2} l''\big(\xi^{(n+1)}, \eta^{(n)}\big) \times \big(\gamma^{(n)} - \gamma^{(n+1)}\big)^2, & \end{aligned}$$

where $\xi^{(n)} = t^{(n)}\gamma^{(n)} + \big(1 - t^{(n)}\big)\gamma^{(n+1)}$, for $0 < t^{(n)} < 1$. Hence

$$\begin{aligned} \frac{M}{2}\Big|\gamma^{(n)} - \gamma^{(n+1)}\Big|^2 &\geq \frac{1}{2} l''\big(\xi^{(n+1)}, \eta^{(n)}\big) \times \big(\gamma^{(n)} - \gamma^{(n+1)}\big)^2 \\ &= l\big(\gamma^{(n)}, \eta^{(n)}\big) - l\big(\gamma^{(n+1)}, \eta^{(n)}\big) - l'\big(\gamma^{(n+1)}, \eta^{(n)}\big) \\ &\qquad \times \big(\gamma^{(n)} - \gamma^{(n+1)}\big) \\ &\geq l\big(\gamma^{(n)}, \eta^{(n)}\big) - l\big(\gamma^{(n+1)}, \eta^{(n)}\big) \\ &= f_{2n} - f_{2n+1}. \end{aligned}$$

From (7), the sequence $\{f_n\}$ is a monotone increasing sequence and $f(\gamma^*, \eta^*)$ is its upper bound, so $\{f_n\}$ is convergent. Because $f_{2n} - f_{2n+1} \to 0$ as $n \to \infty$, the desired result is obtained. □

5.3. Lemma. *Suppose that* $\{\gamma^{(n_k)},\ \eta^{(n_k)}\}$ *is an arbitrary convergent subsequence of* $\{\gamma^{(n)},\ \eta^{(n)}\}$*, and* $(\gamma^{(n_k)},\ \eta^{(n_k)}) \to (\gamma', \eta')$ *when* $k \to \infty$*, then*

$$l(\gamma', \eta') = \max_{\gamma \in R} l(\gamma, \eta') = \max_{\eta \in R^n} l(\gamma', \eta).$$

Proof. For any $\eta \in R^N$, the algorithm shows that

$$l\big(\gamma^{(n_k)}, \eta^{(n_k)}\big) \geq l\big(\gamma^{(n_k)}, \eta\big).$$

Therefore,

$$\lim_{k\to\infty} l\big(\gamma^{(n_k)}, \eta^{(n_k)}\big) \geq \lim_{k\to\infty} l\big(\gamma^{(n_k)}, \eta\big),$$

i.e.

$$(\gamma', \eta') \geq l(\gamma', \eta)$$

holds for any $\eta \in R^N$. Similarly, for any $\gamma \in R$ we have

$$l(\gamma', \eta') \geq l(\gamma, \eta').$$

This completes the proof of Lemma 5.3. □

By assumption, for any $\gamma \in R$, $l(\gamma, \cdot\,)$ is strictly concave on R^N. Thus, we establish a continuous mapping relation $\eta(\,\cdot\,)$ from R to R^N which satisfies

$$l(\gamma, \eta(\gamma)) = \max_{\eta \in R^N} l(\gamma, \eta).$$

5.4. Lemma. *The mapping* $\eta(\,\cdot\,)$ *is continuous from* R *to* R^N*.*

Proof. Suppose that the sequence $\{\gamma^{(n)}\}$ converges to $\gamma_0 \in R$, and let $\{\eta(\gamma^{(n_k)})\}$ be a convergent subsequence of $\{\eta(\gamma^{(n)})\}$ with $\eta_0 = \lim_{k\to\infty} \{\eta(\gamma^{(n_k)})\}$. From the definition of $\eta(\,\cdot\,)$, for any $\eta \in R^N$ we have

$$l\big(\gamma^{(n_k)}, \eta(\gamma^{(n_k)})\big) \geq l\big(\gamma^{(n_k)}, \eta\big).$$

By the continuity of $l(\gamma, \eta)$, as $k \to \infty$ we know that

$$l\big(\gamma_0, \eta_0\big) \geq l\big(\gamma_0, \eta\big).$$

Namely $l\big(\gamma_0, \eta_0\big) = \max_{\eta \in R^N} l(\gamma_0, \eta)$ or $\eta_0 = \eta(\gamma_0)$. By Lemma 5.1, we have

$$\lim_{n\to\infty} \{\eta(\gamma^{(n)}\} = \eta(\gamma_0).$$

This completes the proof of Lemma 5.4. □

From the alternating iterative process in Section 2, the sequence $\{\gamma^{(n)},\ \eta^{(n)}\}$ obtained by the two-step iterative algorithm is the same as the sequence $\{\gamma^{(n)},\ \eta(\gamma^{(n)})\}$. According to Lemma 5.4, we have the following Corollary 5.5.

5.5. Corollary. *The alternating iterative sequence* $\{\gamma^{(n)},\ \eta^{(n)}\}$ *obtained by the two-step iterative algorithm converges if and only if* $\{\gamma^{(n)}\}$ *converges.* □

The following Lemma 5.6 is given in Shi and Jiang [14]:

5.6. Lemma. *Let $\{y_n\}$ be a uniformly bounded sequence in R^k. If $\|y_{n+1} - y_n\| \to 0$ as $n \to \infty$, and the sequence is not convergent, then there are infinitely many accumulation point of the sequence $\{y_n\}$. In particular, for $k = 1$, y' is an accumulation point of $\{y_n\}$ for any $y' \in [\underline{y}, \bar{y}]$, where $\underline{y} = \underline{\lim}_{n\to\infty} y_n$ and $\overline{y} = \overline{\lim}_{n\to\infty} y_n$.* □

Similarly to $\eta(\,\cdot\,)$, we can also establish a continuous mapping $\gamma(\,\cdot\,)$ from R^N to R such that $l(\gamma(\eta), \eta) = \max_{\gamma \in R} l(\gamma, \eta)$. According to the two continuous mapping, $\gamma(\,\cdot\,)$ and $\eta(\,\cdot\,)$, a continuous composite mapping $\gamma \circ \eta(\,\cdot\,)$ from R to R can be obtained. Hence the alternating iterative formula can be written as

$$\gamma^{(n+1)} = \gamma \circ \eta\big(\gamma^{(n)}\big) = \gamma\big(\eta\big(\gamma^{(n)}\big)\big), \ (n \geq 1).$$

Since R is a convex and compact set, and $\gamma \circ \eta(\,\cdot\,)$ is continuous from R to R, there exists $\gamma_0 \in R$ such that $\gamma \circ \eta(\gamma_0) = \gamma_0$ (see Smart [16]). We have the following Lemma 5.7.

5.7. Lemma. *Any accumulation point of the iterative sequence $\{\gamma^{(n)}\}$ is a fixed point of the composed mapping $\gamma \circ \eta(\,\cdot\,)$ in R.*

Proof. Let γ' be an accumulation point of $\{\gamma^{(n)}\}$, that is, there exists a subsequence $\{\gamma^{(n_k)}\} \subset \{\gamma^{(n)}\}$ satisfying $\lim_{k\to\infty} \gamma^{(n_k)} = \gamma'$. By Corollary 5.5, $\{\eta\big(\gamma^{(n_k)}\big)\}$ converges, i.e. $\{\eta^{(n_k)}\}$ converges, and $\lim_{k\to\infty} \eta^{(n_k)} = \lim_{k\to\infty} \eta(\gamma^{(n_k)}) = \eta(\gamma')$. By Lemma 5.3, $l(\gamma', \eta(\gamma')) = \max_{\gamma\in R} l(\gamma, \eta(\gamma'))$ or $\gamma' = \gamma \circ \eta(\gamma')$. This completes the proof of Lemma 5.7. □

Proof of Theorem 2.1. By Corollary 5.5, it is sufficient to prove that $\{\gamma^{(n)}\}$ is convergent. We have $\big|\gamma^{(n+1)} - \gamma^{(n)}\big| \to 0$, as $n \to \infty$ by Lemma 5.2. Let $\underline{\gamma} = \underline{\lim}_{n\to\infty} \gamma^{(n)}$ and $\overline{\gamma} = \overline{\lim}_{n\to\infty} \gamma^{(n)}$. If $\{\gamma^{(n)}\}$ is not convergent, then the conditions of Lemma 5.6 hold. Therefore, $\underline{\gamma} < \overline{\gamma}$, and γ' is an accumulation point of $\{\gamma^{(n)}\}$ for any $\gamma' \in [\underline{\gamma}, \overline{\gamma}]$ in view of Lemma 5.6. By Lemma 5.7, γ' is a fixed point of the composite mapping $\gamma \circ \eta(\,\cdot\,)$ satisfying $\gamma \circ \eta(\gamma') = \gamma'$.

On the other hand, it can be shown that $\gamma^{(n)}$ is not element of $[\underline{\gamma}, \overline{\gamma}]$ for any $n \geq 1$. If it is not true, then we assume that there exists n_0 such that $\gamma^{(n_0)} \in [\underline{\gamma}, \overline{\gamma}]$. By Lemma 5.7, $\gamma^{(n_0)} = \gamma \circ \eta\big(\gamma^{(n_0)}\big) = \gamma^{(n_0+1)}$. Then $l\big(\gamma^{(n_0)},\ \eta^{(n_0)}\big) = l\big(\gamma^{(n_0+1)}, \eta^{(n_0)}\big)$, which contradicts the inequality $l\big(\gamma^{(n_0)}, \eta^{(n_0)}\big) > l\big(\gamma^{(n_0+1)}, \eta^{(n_0)}\big)$ for all $n \geq 1$.

So, for any $\gamma' \in \big(\underline{\gamma}, \overline{\gamma}\big)$ there exist $\epsilon_{\gamma'} > 0$ and a neighborhood of γ', $B(\gamma', \epsilon_{\gamma'}) = (\gamma' - \epsilon_{\gamma'}, \gamma' - \epsilon_{\gamma'})$, such that $B(\gamma', \epsilon_{\gamma'}) \bigcap \{\gamma^{(n)}\}$ is the empty set, which contradicts the fact that γ' is an accumulation point of $\{\gamma^{(n)}\}$. This contradiction implies that Theorem 2.1 is true. □

5.8. Remark. Theorem 2.1 only considers the case where the inequalities

$$l\big(\gamma^{(n)}, \eta^{(n)}\big) < l\big(\gamma^{(n+1)},\ \eta^{(n)}\big) < l\big(\gamma^{(n+1)},\ \eta^{(n+1)}\big)$$

hold for all $n \in N$.

5.9. Remark. If there exists a positive n_0 such that $l\big(\gamma^{(n_0)}, \eta^{(n_0)}\big) = l\big(\gamma^{(n_0+1)}, \eta^{(n_0)}\big)$, or $l\big(\gamma^{(n_0)}, \eta^{(n_0-1)}\big) = l\big(\gamma^{(n_0)}, \eta^{(n_0)}\big)$, then the sequence $\{\gamma^{(n)},\ \eta^{(n)}\}$ converges. $l\big(\gamma, \eta^{(n_0)}\big)$ is strictly concave on R, hence its maximum point is unique. So

$$l\big(\gamma^{(n_0)},\ \eta^{(n_0)}\big) = l\big(\gamma^{(n_0+1)},\ \eta^{(n_0)}\big) = \max_{\gamma\in R} l\big(\gamma,\ \eta^{(n_0)}\big).$$

It is obvious that $\gamma^{(n_0+1)} = \gamma^{(n_0)} \Rightarrow \eta^{(n_0+1)} = \eta^{(n_0)} \Rightarrow \gamma^{(n_0+2)} = \gamma^{(n_0+1)} \Rightarrow \cdots$. Namely, for any $n \geq n_0$, the above sequence implies that $\big(\gamma^{(n)},\ \eta^{(n)}\big) = (\gamma^{(n_0)}, \eta^{(n_0)})$. Therefore, the sequence $\{\gamma^{(n)}, \eta^{(n)}\}$ converges.

Acknowledgments. The authors would like to thank the anonymous referees and the editor for helpful suggestions and insightful comments that greatly improved the paper.

This research is supported by the National Natural Science Foundation of China under Grants No. 10431010, 10501005; and Innovation Group Program of Liaoning Educational Committee No.2007T050.

References

[1] Ahn, S. C. and Schmidt, P. *Efficient estimation of models for dynamic panel data*, Journal of Econometrics **68**, 5–27, 1995.

[2] Arellano, M. and Bond, S. R. *Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations*, Review of Economic Studies **58**, 277–297, 1991.

[3] Arellano, M. and Bover, O. *Another look at the instrumental variable estimator of error-components models*, Journal of Econometrics **68**, 29–51, 1995.

[4] Blundell, R. and Bond, S. *Initial conditions and moment restrictions in dynamic panel data models*, Journal of Econometrics **87**, 115–143, 1998.

[5] Bun, M. J. G. and Carree, M. A. *Bias-corrected estimation in dynamic panel data models*, Journal of Business and Economic Statistics **23**, 200–210, 2005.

[6] Bun, M. J. G. and Carree, M. A. *Bias-corrected estimation in dynamic panel data models with heteroscedasticity*, Economics Letters **92**, 220–227, 2006.

[7] Hahn, J. and Kuersteiner, G. *Asymptotically unbiased inference for a dynamic panel model with fixed effects when both n and T are large*, Econometrica **70**, 1639–1657, 2002.

[8] Hansen, G. *A bias-corrected least squares estimator of dynamic panel models*, Allgemeines Statistisches Archiv **85**, 127–140, 2001.

[9] Hsiao, C., Pesaran, M. H. and Tahmiscioglu, A. K. *Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods*, Journal of Econometrics **109**, 107–150, 2002.

[10] Kiviet, J. F. *On bias, inconsistency, and efficiency of various estimators in dynamic panel data models*, Journal of Econometrics **68**, 53–78, 1995.

[11] Kuk, A. Y. C. *Asymptotically unbiased estimation in generalized linear models with random effects*, J. R. Statist. Soc. B **57**, 395–407, 1995.

[12] Neyman, J. and Scott, E. L. *Consistent estimates based on partially consistent observations*, Econometrica **16**, 1–32, 1948.

[13] Nickell, S. *Biases in dynamic models with fixed effects*, Econometrica **49**, 1417–1426, 1981.

[14] Shi, N. Z. and Jiang, H. *Maximum likelihood estimation of isotonic normal means with unknown variance*, J. Multivariate Anal **64**, 183–195, 1998.

[15] Shi, N. Z., Hu, G. R. and Cui, Q. *An alternating iterative method and its application in statistical inference*, Acta Mathematica Sinica **24**, 843–856, 2008.

[16] Smart, D. R. *Fixed Point Theorems* (Cambridge University Press, London, 1974).