School of Business, Economics and Law
GÖTEBORG UNIVERSITY

WORKING PAPERS IN ECONOMICS

No 278

Monte Carlo Investigation of the Initial Values Problem in Censored Dynamic Random-Effects Panel Data Models

by

Alpaslan Akay

December 2007

ISSN 1403-2473 (print)
ISSN 1403-2465 (online)

SCHOOL OF BUSINESS, ECONOMICS AND LAW, GÖTEBORG UNIVERSITY

Department of Economics
Visiting adress Vasagatan 1,
Postal adress P.O.Box 640, SE 405 30 Göteborg, Sweden
Phone + 46 (0) 31 786 0000

Monte Carlo Investigation of the Initial Values Problem in Censored Dynamic Random-Effects Panel Data Models*

Alpaslan Akay[†]

Department of Economics, Göteborg University

December 3, 2007

Abstract

Three designs of Monte Carlo experiments are used to investigate the initial-value problem in censored dynamic random-effects (Tobit type 1) models. We compared three widely used solution methods: naive method based on exogenous initial values assumption; Heckman's approximation; and the simple method of Wooldridge. The results suggest that the initial values problem is a serious issue: using a method which misspecifies the conditional distribution of initial values can cause misleading results on the magnitude of true (structural) and spurious state-dependence. The naive exogenous method is substantially biased for panels of short duration. Heckman's approximation works well. The simple method of Wooldridge works better than naive exogenous method in short panels, but it is not as good as Heckman's approximation. It is also observed that these methods performs equally well for panels of long duration.

Keywords: *Initial value problem, Dynamic Tobit model, Monte Carlo experiment, Heckman's approximation, Simple method of Wooldridge*

J.E.L Classification: C23, C25.

*I thank Lennart Flood, Konstantin Tatsiramos, Roger Wahlberg, Elias Tsakas, Peter Martinsson, and seminar participants in Göteborg for their valuable comments.

[†]Department of Economics, Göteborg University, Box 600 SE 405 30 Göteborg, Sweden. Tel: +46-(31) 773 5304 Email: Alpaslan.Akay@Economics.gu.se

1 Introduction

Censored dynamic panel data models have been widely analyzed by many authors (Honore, 1993; Arellano and Bover, 1997; Arellano, Bover and Labeaga, 1999; Honore and Hu, 2001; Hu, 2002). Given the goal of disentangling the *true* (*structural*) *state-dependence* from *spurious state-dependence*, one of the crucial issues is the initial values problem (Heckman, 1981; Blundell and Smith, 1991; An and Liu, 1997; Blundell and Bond, 1998; Lee, 1999; Arellano and Honore, 2001; Honore, 2002; Hsiao, 2003; Arellano and Carrasco, 2003; Honore and Hu, 2004; Arellano and Hahn, 2005; Honore and Tamer, 2006). The aim of this paper is to compare some widely used solution methods of the initial values problem in censored dynamic random-effects panel data models using various designs of Monte Carlo experiments (MCE).

The initial values problem can appear if the history of the stochastic process underlying the model is not fully observed. If the process is operated before the sample data is observed and if the initial (*sample*) values have been affected by the unobserved past, then the initial values problem can emerge since the initial values have possibly been created by the evolution of the strictly exogenous variables in interaction with unobserved individual-effects. The solution of the problem is to specify a distribution of initial values which is conditioned on strictly exogenous variables and unobserved individual-effects. Ad hoc treatments of this problem can produce bias and inconsistency in the estimators of the censored dynamic random-effects model as it would also cause in similar probit, logit or Poisson models (Heckman, 1981; Honore, 2002; Hsiao, 2003; Honore and Tamer, 2006).

Besides the initial values problem, the random-effect approach has some other limitations. It requires an assumption about the conditional distribution of unobserved individual-effects. To avoid these problems a fixed-effects approach can be used, which can be attractive as a way to ensure that the conditional distribution of unobserved individual-effects does not play a role in the estimation of the parameters. However, it can also be seriously biased since it suffers from the *incidental parameters problem* (Neyman and Scott, 1948; Greene, 2004). Alternatively, some other estimators based on semiparametric methods or combinations of these methods with the fixed-effects approach (such as censored least absolute deviation estimator suggested by Hu (2002) or the fixed-effects approach developed by Honore (1993)) can be used for estimating a censored dynamic panel data model (see also Honore and Hu, 2001). However, these estimators are still subject to the incidental parameters problem and in these estimators time-invariant exogenous variables are swept away, which can also be a serious problem in the practice. Thus, the random-effects approach is still attractive, and if it is preferred, a proper solution for the initial values problem is necessary.

The aim of this paper is to compare some widely used solution methods of the initial values problem in censored dynamic random-effects panel data models. To do this, various designs of MCE are provided. We designed cases in which a solution for the initial values problem is necessary, and three solution methods are investigated: The first is the naive approach in which the initial values are considered as *exogeneous* variables, independent from unobserved individual-effects and strictly exogenous variables. The other two consider the initial values as *endogenous* variables. Thus, the second is the Heckman's (1981) method, which uses a reduced-form approximation for the conditional distribution of initial values based on available pre-sample information. The third method is the simple method of Wooldridge (2005), which uses an auxiliary distribution of unobserved individual-effects conditioned on initial values and strictly exogenous variables.

The results suggest that the initial values problem is a serious issue which can lead to substantial bias if the conditional distribution of initial values is misspecified. The naive exogenous method can highly overstate (understate) the size of the true state-dependence (spurious state-dependence), if it is wrong. It is found that Heckman's reduced-form approximation works well for all durations of panels. The simple method of Wooldridge works much better than naive exogenous method, and it is as successful as Heckman's approximation with moderately long panels. It is also found that these methods tend to perform equally well for panels of long durations.

The paper is organized as follows; the next section will give the model, description of the initial values problem and three solution methods. Section 3 presents our Monte Carlo designs and results. Section 4 concludes.

2 The model and three solution methods of the initial values problem

Consider the following censored dynamic random-effects model with one lag of *censored* dependent variable:[1]

$$y_{i0} = \max(0, x_{i0}'\beta + \epsilon_{i0}) \tag{1}$$

$$y_{it} = \max(0, x_{it}'\beta + \gamma y_{i,t-1} + \epsilon_{it}) \tag{2}$$

[1] The other alternative is to consider that the lagged values of the dependent variable is also latent. Considering the lagged dependent variable as *observed* or latent lead to different implications in both economic and estimation terms. See Honore (1993), Hu (2002) and Hsiao (2003) for useful discussions.

where $\epsilon_{it} = \alpha_i + u_{it}$ is the composite error terms; x_{it} is a vector of *strictly* exogenous variables in a sense that they are independent from all past, current and future values of the disturbance $u_{it} \sim iidN(0, \sigma_u^2)$; α_i is time-persistent unobserved individual-effects (unobserved heterogeneity) with a conditional probability distribution $f(\alpha_i|x_{it})$. In this paper, we assume that the distribution of the random-effects is $\alpha_i \sim iidN(0, \sigma_\alpha^2)$, and they are orthogonal to exogenous variables following the standard random-effects assumption. Throughout the paper, the number of individuals N $(i = 1, ..., N)$ is considered to be large relative to the number of periods T $(t = 1, ..., T)$. Covariance structure of the model is assumed as

$$E[\epsilon_{it}\epsilon_{i,t-s}|\,\{x\}_{t=1}^{T}] = \begin{Bmatrix} \sigma_\epsilon^2 & s = 0 \\ \rho\sigma_\epsilon^2 & s \neq 0 \end{Bmatrix} \tag{3}$$

The composite variance is written as $\sigma_\epsilon^2 = \sigma_\alpha^2 + \sigma_u^2$ and ρ is the fraction of the variation explained by the unobserved individual-effects. The likelihood at time t for an individual i is given by

$$f_{it}(y_{it}|y_{i,t-1}, x_{it}, \alpha_i, \theta) = \begin{Bmatrix} 1 - \Phi\left[(x_{it}'\beta + \gamma y_{i,t-1} + \sigma_\alpha\alpha_i)/\sigma_u\right] & y_{it} = 0 \\ (1/\sigma_u)\phi\left[(y_{it} - x_{it}'\beta - \gamma y_{i,t-1} - \sigma_\alpha\alpha_i)/\sigma_u\right] & y_{it} > 0 \end{Bmatrix} \tag{4}$$

where Φ denotes the distribution function and ϕ denotes the density function of standard normal random variable; and $\theta = \begin{bmatrix} \beta & \gamma & \sigma_u \end{bmatrix}$. The full log-likelihood function is given as

$$\ln \mathcal{L} = \sum_{i=1}^{N} \ln \left[\int_{-\infty}^{\infty} \begin{bmatrix} f_0\left(y_{i0}|\,\{x_{it}\}_{t=0}^{T}, \alpha_i; \theta\right) \times \\ \prod_{t=1}^{T} f_{it[y_{it}=0]} \times \prod_{t=1}^{T} f_{it[y_{it}>0]} \end{bmatrix} f(\alpha_i)d\alpha_i \right] \tag{5}$$

where $f_0\left(y_{i0}|\,\{x_{it}\}_{t=0}^{T}, \alpha_i; \theta\right) = \left\{f_{0[y_{it}=0]}, f_{0[y_{it}>0]}\right\}$ is the probability distribution of initial values which is conditioned on strictly exogenous variables and the unobserved individual-effects.

There are two alternatives; either *logical* starting point of the stochastic process underlying the model (2) and the observed sample data is the same or the sample data are observed after the process is operated many periods.[2] For the first case, initial values y_{i0} may be known constants and therefore there is no reason to specify a probability distribution for initial values. Thus, $f_0\left(y_{i0}|\,\{x_{it}\}_{t=0}^{T}, \alpha_i; \theta\right)$ can be taken out from the likelihood function (Heckman, 1981; Honore, 2002; Hsiao, 2003). However, if observed

[2] Considering the complex associations between variables in economics it is not easy to determine an objective starting point for a process. For example, let us consider the relative earnings of immigrants in a host country. We can start to observe them upon arrival and *logically* the starting point of the earnings generating process can be assumed as started upon arrival. However, this assumption will ignore earnings experiences and accumulated human-capital acquired in county of origin which can also be considered as a part of the process.

sample data start after the process has been operated through many periods, the initial values (the first period in the observed sample data, $t = 1$) cannot be constant since they have possibly been created by the evolution of exogenous variables interacting with unobserved individual-effects. Thus, in this case a probability distribution of initial values ($f_{i1}\left(y_{i1}|\left\{x_{it}\right\}_{t=1}^{T}, \alpha_i; \theta\right)$) must be specified.

In general, researchers can follow two alternative ways to solve the initial values problem in practice. The first is to naively *forget* the problem and assume that the initial values have not been affected by unobserved past, even if it may not be true. It means that the initial values are *exogenous* variables, independent from unobserved individual-effects. Thus the conditional distribution of the initial values would be equal to their marginal distributions $f_{i1}(y_{i1})$ and it can be taken outside the maximization procedure of the likelihood function. If the data have not been observed at the beginning of the process, and if the disturbances that generate the process is serially correlated (which is inevitable in the presence of unobserved individual-effects), then this assumption is too strong and causes serious consequences such as bias and inconsistency in the estimators (Heckman, 1981; Hyslop, 1998; Honore, 2002).[3]

The second and more realistic approach is to assume *endogenous* initial values and specify the conditional distribution. However, it is not a easy task to find a closed-form expression for this distribution.[4] Heckman (1981) suggested a reduced-form approximation for the conditional distribution of initial values, based on available pre-sample information. Heckman's approximation can provide flexible specifications for the relationship between initial values, unobserved individual-effects and exogenous variables. Consider the following *reduced-form* equation for initial values:

$$y_{i1} = \max(0, z_{i1}'\pi + \epsilon_{i1}) \tag{6}$$

$$\epsilon_{i1} = \lambda\alpha_i + u_{i1} \tag{7}$$

where z_{i1} is a vector of available strictly exogenous instruments which will constitute the pre-sample information. This vector can also contain the first observations of exogenous variables in the observed sample; π and λ are the nuisance parameters to be estimated; ϵ_{i1}

3 It is assumed that the actual disturbance process is serially uncorrelated (such as first order autocorrelation $AR(1)$) and the dynamic feature of the model is obtained by including a lagged dependent variable. However, it does not mean that the disturbances are serially uncorrelated. It is possible only if the variance of the unobserved individual-effects is zero, meaning that the model has no panel data characteristics.

4 One possibility is to assume that the conditional distribution of initial values to be at the steady state. However, it is still difficult to find a closed-form expression for the distribution even for the simplest case where there is no explanatory variable. This assumption is also very strong if age-trended variables are driving the process (Heckman, 1981; Hyslop, 1998; Hsiao, 2003).

is correlated with α_i but it is uncorrelated with u_{it} $(t \geq 1)$.The random-effects assumption implies that α_i is uncorrelated with u_{i1}.Thus, the approximated conditional distribution of initial values is specified as follows:

$$f_{i1}\left(y_{i1}|z_{i1},\alpha_i;\pi,\lambda\right) = \left\{ \begin{array}{ll} 1-\Phi\left[(z_{i1}'\pi+\lambda\sigma_\alpha\alpha_i)/\sigma_u\right] & y_{i1}=0 \\ (1/\sigma_u)\phi\left[(y_{i1}-z_{i1}'\pi-\lambda\sigma_\alpha\alpha_i)/\sigma_u\right] & y_{i1}>0 \end{array} \right\} \tag{8}$$

with $Var[\epsilon_{i1}] = \lambda^2\sigma_\alpha^2 + \sigma_u^2$ and the correlation between ϵ_{i1} and unobserved individual-effects $(\rho_{\epsilon_{i1}\alpha_i})$ is

$$\rho_{\epsilon_{i1}\alpha_i} = Corr(\epsilon_{i1},\alpha_i) = \frac{\lambda\sigma_\alpha}{\sqrt{\lambda^2\sigma_\alpha^2+\sigma_u^2}} = \frac{\psi}{\sqrt{\psi^2+1}} \tag{9}$$

where $\psi = \lambda\sigma_\alpha/\sigma_u$. The parameters of the structural system (2) and the approximate reduced-form conditional probability (8) can be simultaneously estimated without imposing any restriction (Heckman, 1981; Hsiao, 2003).

Another solution method is suggested by Wooldridge (2005) which is a simple alternative to Heckman's reduced-form approximation. This method considers the distribution of unobserved individual-effects to be conditioned on initial values and exogenous variables. Specifying the distribution on these variables can lead to very tractable functional forms, and consistent estimators in censored dynamic random-effects models as well as in similar probit, logit and Poisson models (Honore, 2002; Wooldridge, 2005).

This method suggests specifying $f\left(\alpha_i|\left\{x_{it}\right\}_{t=1}^{T}, y_{i1}\right)$ instead of $f_{i1}(.)$ using a similar strategy to Chamberlain's (1984) correlated-effects model. It is based on the following auxiliary distribution of unobserved individual-effects.

$$\alpha_i = \xi_0 + \xi_1 y_{i1} + \xi_2\overline{x}_i + \eta_i \tag{10}$$

where $\alpha_i|y_{i1},\overline{x}_i \sim N\left[\xi_0+\xi_1 y_{i1}+\xi_2\overline{x}_i,\sigma_\eta^2\right]$ and η_i is a new unobserved individual-effects which is assumed as $\eta_i \sim iidN\left[0,\sigma_\eta^2\right]$; y_{i1} is the initial sample values; $\overline{x}_i$ is the within-means of time-variant exogenous variables defined as $\overline{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}$. Thus, we obtain a conditional likelihood which is based on the joint distribution of the observations conditional on initial values. This likelihood function will be like those in standard static random-effect censored model and the parameters can be easily estimated using a commercial random-effects software.

The likelihood function (5) of the censored dynamic random-effects model which is adopted here, involves only a single integral, which can be effectively implemented using Gaussian-Hermite Quadrature (Butler and Muffitt, 1982). This method is much less time consuming and efficient in comparison with the other alternative based on simulation with a proper simulator, such as frequency (natural) by direct Monte Carlo sampling from normal distribution and GHK. (Gourieroux and Monfort, 1993; Hajivassiliou and Ruud,1994). In this paper, we therefore prefer to use Gaussian-Hermite Quadrature in all likelihood computations.

3 Monte Carlo experiments and the results

In order to compare the finite sample performance of the solution methods several designs of MCE are considered that differ on the length of the panel, number of individuals, the relative sizes of the key parameters and on the data generating process for the explanatory variables.[5] We apply the following strategy: We first analyze the bias for the case in which the *initial values are known* constants in order to check the possible bias when the initial values problem is not exist. Second, we design cases in which the initial values problem is severe and analyze *naive exogenous initial values method* as a worst scenario. Third, we use the same data sets to analyze and to compare the performance of *Heckman's reduced-form approximation* and *simple method of Wooldridge*.

The data generating process based on the censored dynamic random-effects model is specified as follows.

$$y_{i0} = \max(0, \frac{\beta x_{i0}}{1-\gamma} + \frac{\alpha_i}{1-\gamma} + \frac{u_{i0}}{\sqrt{1-\gamma^2}}) \tag{11}$$

$$y_{it} = \max(0, \beta x_{it} + \gamma y_{i,t-1} + \alpha_i + u_{it}) \tag{12}$$

where $i = 1, ..., N$ and $t = 1, ..., T$; $\alpha_i \sim iidN\,[0, \sigma_\alpha^2]$; and $u_{it} \sim iidN\,[0, \sigma_u^2]$. The design adopted for the initial values y_{i0} aims first to include correlation between initial values and unobserved individual-effects, and second, to create mean stationarity in the stochastic process. All the results presented here are based on $L = 200$ conditioning data sets. We produced a new set of panel data for each experiment and the same data set is also used for each solution methods. The number of individuals is set to $N = 200$. The behavior of bias is also analyzed for large number of individuals by using $N = 300$, 500, 750 and 1000. The durations of the panel data sets are set to $T = 3, 5, 8, 15, 20$. Number of quadrature points (nodes and weights) used in the optimization procedure of the likelihood function is set to 30.[6]

[5] Our MCE is designed in Fortran software, and the optimization for the likelihood functions is performed using $ZXMIN$, which is very fast and robust. The routines written for the experiments can be provided by the author upon request.

[6] We used different number of quadrature nodes and weights in order to check stability of estimated parameters. It is observed that 30 quadrature points produce very stable results.

3.1 MCE_1 : Benchmark design. A normal explanatory variable

The benchmark design consists of one strictly exogenous explanatory variable which is obtained by using independent and identically distributed standard normal random variates,

$$x_{it} \sim N\left[0, 1\right] \tag{13}$$

True values of the parameters β and σ_u^2 are set to 1; two values for the true state-dependence $\gamma = -0.5$ and $\gamma = 0.5$ are used. The variance of unobserved individual-effects is first set to $\sigma_\alpha^2 = 1$ and then increased to $\sigma_\alpha^2 = 3$, in order to analyze the size of the variance of unobserved-effects on the estimated parameters. The design is produced, in average, $45 - 55\%$ censored observations.

The results of MCE_1 are summarized in Table 1a, 1b, 1c and 1d. Tables report results only for the key parameters: $\widehat{\beta}$, $\widehat{\gamma}$, $\widehat{\sigma}_\alpha^2$ and $\widehat{\sigma}_u^2$. In addition to the *mean bias* and *root mean square error* ($RMSE$), the *median bias* and *median absolute error* (MAE) are also reported since the estimators of the type considered here often do not have finite theoretical moments. The median bias and MAE are also less sensitive to outliers compared to other two measures. A negative sign on both mean and median bias shows an underestimation and a positive sign shows an overestimation.

Table 1a about here

We focus first on the case in which initial values are *known* (Table 1a) in a sense that the sample data and the process start at the same time and also initial values are nonstochastic ($y_{i0} = 0$). Thus, there is no initial values problem and the bias is very small even with panels of short durations. The mean and median bias are very close to each other meaning that the bias has a symmetric distribution. The variation around the true values is reduced as T increases. A larger true value for the variance of unobserved individual-effects ($\sigma_\alpha = \sqrt{3}$) causes a slight increase in the bias and variation. Last row of Table 1a results by number of individuals (N) for a constant number of time periods ($T = 5$). The bias seems not to be affected by the number of individuals in the panel set.

Table 1b about here

As a second step, the process is operated 25 periods before the sample data are collected in order to create a initial values problem.[7] Table 1b presents the results for the

[7] We operate the system through 25 periods before the sample data is observed. For example, when $T = 3$, the sample data contain the $(y_{i26}, y_{i27}, y_{i28})$ and we use it as (y_{i1}, y_{i2}, y_{i3}). Where y_{i1} are the initial sample values.

naive method based on exogenous initial values assumption. In this case, this assumption is wrong and, as expected, it causes large bias in γ and σ_α. γ is highly overestimated while σ_α is highly underestimated. The bias is $40-50\%$ when $T=3$ for these two parameters. The bias is also remarkable reduced by the duration of panel (especially for $T>10$). A large value of σ_α increased the bias substantially (70% for γ and more than 100% for σ_α). The other two parameters (β and σ_u^2) are found not be largely biased in almost every case.

Table 1c and Table 1d present results for Heckman's reduced-form approximation and simple method of Wooldridge, respectively. Heckman's approximation method performs very well for all durations of the panels. γ and σ_α are almost $3-5\%$ biased when $T=3$, and a large value of σ_α causes the bias to be larger ($5-10\%$). The simple method of Wooldridge also performs well but not as well as Heckman's approximation. The simple method of Wooldridge also tends to overestimate γ and underestimate σ_α for small samples as naive method. The bias produced by this method is about $15-25\%$ for $T=3$. For a duration which is greater than $T=5$, the size of the bias produced by the simple method of Wooldridge method tends to be equal to the Heckman's approximation. Additionally, all methods perform equally well for the panels which are longer than $T=10-15$.

Table 1c about here

Table 1d about here

3.2 MCE_2 : A non-normal explanatory variable

As pointed out by Honore and Kyriazidou (2000), normally distributed explanatory variables can make the bias appear smaller than it is for other distributions of the explanatory variables, which can largely affect the results in Monte Carlo studies. We, therefore, modify MCE_1 by changing the distribution of the explanatory variable to one degrees of freedom chi-square distributed random variable $\chi^2(1)$, which has a skewed distribution. We standardize this random variable to transform it to the same mean and variance with the exogenous variable given above.[8]

$$x_{it} \sim \frac{\chi^2_{(1)} - 1}{\sqrt{2}} \tag{14}$$

[8] Note that $Z = \left(\chi^2_{(k)} - k\right) / \sqrt{2k}$, where k is the degrees of freedom. Z is the standardized χ^2 random variable.

The data-generating process for dependent variable (11-12) is the same as for the benchmark case and the only difference is explanatory variable used in the estimation. True values of the parameters are set to: $\beta = 1$, $\gamma = 0.5$, $\alpha_i \sim iidN\left[0, \sigma_\alpha^2 = 1\right]$ and $u_i \sim iidN\left[0, \sigma_u^2 = 1\right]$. The process is operated through 25 periods before the samples are observed with durations of $T = 3, 5, 8, 15, 20$, and the average number of observations that are censored is almost the same as the benchmark design.

Table 2 about here

The results of MCE_2 are summarized in Table 2. A comparison between the results in Table 1a-1d and the corresponding results in Table 2 suggests that the results in benchmark design MCE_1 are very robust. The methods do not produce significantly larger bias with non-normal explanatory variables. The bias has symmetric distribution with a decreasing variance. The performance order between the methods is clear: The smallest bias is obtained by Heckman's reduce-form approximation and it is followed by the simple method of Wooldridge for short panels. The initial values problem tended to be not important source of bias when the duration of the panel is increased.

3.3 MCE_3 : An autocorrelated explanatory variable

MCE_3 is based on a relatively complicated data generating process for explanatory variable which contains higher degree of intra-group variations. In this design, there is only one strictly exogenous variable x_{it} based on following first order autoregressive process

$$x_{it} = \rho x_{it-1} + \psi_{it} \tag{15}$$

where ψ_{it} is a standard normal random variable $\psi_{it} \sim N\left[0, 1\right]$, $\rho = 0.5$ and $\psi_{i1} = x_{i1}$. True values of the parameters are set to: $\beta = 1.0$, $\gamma = 0.5$, $\alpha_i \sim iidN\left[0, \sigma_\alpha^2 = 1\right]$ and $u_i \sim iidN\left[0, \sigma_u^2 = 1\right]$. The data generating process for dependent variable is kept the same as in (11-12) and the process is operated through 25 periods before the samples are observed with the durations $T = 3, 5, 8, 15, 20$. The number of the censored observations is almost the same as those produced in first two MCE.

Table 3 about here

The results of the MCE_3 are reported in Table 3. Introducing more intra-group variation to explanatory variable does not change the results found above. The magnitude of the bias and the performance order among the solution methods are the same as those obtained in other two MCE.

Figure 1 shows the Q-Q plots based on the quantiles of normal distribution, by solution methods. We present only for the true state-dependence and variance of unobserved individual-effects. These figures show whether the asymptotic distribution of the estimators used here can be approximated by normal distribution for our MCE samples. We plot the empirical quantiles of the estimated Monte Carlo parameters in MCE_3 against those of normal distribution, where $T = 10$ and number of MCE replication is $L = 200$.

Figure 1 about here

The Q-Q plots support the normality approximation. The empirical quantiles of estimated Monte Carlo parameters in MCE_3 lie mostly in straight lines for all solution methods.

4 Discussion and conclusions

The performance of some widely used solution methods of initial values problem in censored dynamic random-effects models is analyzed using several designs of Monte Carlo experiments. We first presented results for the case in which the initial values are known constants implying that there is no initial values problem. Second, we designed cases in which the initial values problem is severe, and the naive method based on exogenous initial values is analyzed to simulate the effect of a mistreatment for the problem. Third, the performance of the Heckman's (1981) reduced-form approximation and simple method of Wooldridge (2005) are analyzed and compared using the same conditioning data.

The initial values problem can lead to misleading results on the magnitude of true and spurious state-dependence. The naive exogenous initial values method can produce substantial bias especially for the panels of short duration. It causes true state-dependence to be highly overestimated while the variance of unobserved individual-effects is highly underestimated. Considering the durations of the micro-panel data sets encountered in the practice, which generally have thousands of individuals and small number of periods, the conditional distribution of initial values must be specified. Among the solution methods based on specifying the conditional distribution, Heckman's reduced-form approximation is the best choice for the small samples, but for moderate samples there is no clear performance order between Heckman's and Wooldridge's methods with respect to bias that they produce. The message is that the simple method of Wooldridge can be used instead of Heckman's approximation for the panels of moderate duration (such as, time periods $T = 5 - 10$ time periods). Another intuitive message is that all methods which are compared here tend to perform equally well for panels of long duration (such as, time periods $T > 10 - 15$)

From an empirical point of view, Heckman's approximation constitutes a computationally challenging task especially with an *unbalanced* panel data set. As explained in Honore (2002), ad hoc treatments of the initial values problem are in particular unappealing with unbalanced panel data sets, which are the ones generally used in empirical applications. As seen in the Monte Carlo studies above, the simple method of Wooldridge is attractive especially with panels of moderate durations and also it can be easily applied using a standard random-effect software with either balanced or unbalanced panel data sets.

References

[1] An, M.Y., and M. Liu (2000), Using Indirect Inference to Solve the Initial Conditions Problem, *Review of Economics and Statistics*, **4**: 656-667

[2] Arellano, M., O. Bover, and J. Labeaga (1997), Autoregressive Models with Sample Selectivity for Panel Data, Working Paper No. 9706, CEMFI.

[3] Arellano, M., and B. Honore (2001), Panel Data Models. Some Recent Developments, *Handbook of Econometrics*, 5, Elsevier Science, Amsterdam.

[4] Arellano, M. and O. Bover (1997), Estimating Dynamic Limited Dependent Variable Models From Panel Data, *Investigaciones Economicas*, **21**: 141-65.

[5] Arellano, M. and R. Carrasco (2003), Binary Choice Panel Data Models with Predetermined Variables, *Journal of Econometrics*, **115**: 125-157

[6] Arellano, M. and J. Hahn (2006), Understanding Bias in Nonlinear Panel Models: Some Recent Developments. In: R. Blundell, W. Newey, and T. Persson (eds.): Advances in Economics and Econometrics, Ninth World Congress, Cambridge University Press, *forthcoming*.

[7] Blundell, R.W. and R.J. Smith (1991), Initial Conditions and Efficient Estimation in Dynamic Panel Data Models, *Annales d'Economie et de Statistique*, **20/21**: 109-123.

[8] Bulundell, R., and S. Bond (1998), Initial conditions and Moment Conditions in Dynamic Panel Data Models, *Journal of Econometrics*, **87**: 115-143

[9] Butler, J.S., and R. Moffitt (1982), A Computationally Efficient Quadrature Procedure for the One Factor Multinomial Probit Model, *Econometrica*, **50**: 761-764.

[10] Chamberlain, G. (1984), Panel Data, in *Handbook of Econometrics*, Vol. II, edited by Zvi Griliches and Michael Intriligator. Amsterdam: North Holland.

[11] Heckman, J. (1981), The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time - Discrete Data Stochastic Process, in Structural Analysis of Discrete Panel Data with Econometric Applications, ed. by C. Manski and D. McFadden, Cambridge: MIT Press.

[12] Hajivassiliou, V., and P. Ruud (1994), Classical Estimation Methods for LDV Models using Simulation, in R. Engle & D. McFadden, eds. *Handbook of Econometrics*, IV, 2384-2441.

[13]

[14] Honore, B. (1993), Orthogonality-Conditions for Tobit Model with Fixed Effect and Lagged Dependent Variables, *Journal of Econometrics,* **59**: 35-61

[15] Honore, B., and E. Kyriazidou (2000), Panel Data Discrete Choice Models with Lagged Dependent Variables, *Econometrica,* **68**: 839-874

[16] Honore, B., and L. Hu (2001), Estimation of Cross Sectional and Panel Data Censored Regression Models with Endogeneity, unpublished manuscript, Princeton University.

[17] Honore, B. (2002), Nonlinear Models with Panel Data, *Portuguese Economic Journal*, **1**: 163–179.

[18] Honore, B., and E. Tamer (2006), Bounds on Parameters in Panel Dynamic Discrete Choice Models, *Econometrica*, **74**: 611-629

[19] Honore, B., and L. Hu (2004), Estimation of Cross Sectional and Panel Data Censored Regression Models with Endogeneity, *Journal of Econometrics*, **122**(2): 293–316,

[20] Hsiao, C. (2003), Analysis of Panel Data, 2nd ed. Cambridge University Press, Cambridge.

[21] Hu, L. (2002), Estimation of a censored dynamic panel-data model, *Econometrica*, **70**: 2499-2517

[22] Hyslop, D. R. (1999), State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women, *Econometrica*, **67**, 1255–1294.

[23] Gourieroux, C., and A. Monfort (1993), Simulation-based Inference: A Survey with Special Reference to Panel Data Models, *Journal of Econometrics*, **59**: 5-33.

[24] Greene, W. (2004), Fixed effects and bias due to the incidental parameters problem in the Tobit model, *Econometric Reviews*, **23**:125-147

[25] Keane, M. P. (1994), A Computationally Practical Simulation Estimator for Panel Data, *Econometrica*, **62**: 95-116

[26] Lee, L.F. (1999), Estimation of dynamic and ARCH Tobit models, *Journal of Econometrics*, **92**: 355-390

[27] McFadden, D., Ruud, P. (1994). Estimation by Simulation. *Review of Economics and Statistics*, **76**: 591-608.

[28] Mundlak, Y. (1978), On the Pooling of Time Series and Cross Section Data, *Econometrica*, **46**: 69-85

[29] Nerlove, M. (1971), Further evidence on the estimation of dynamic economic relations from a time series of cross sections, *Econometrica*, **39**: 359-383

[30] Neyman, J. and E. Scott (1948), Consistent Estimates Based on Partially Consistent Observations, *Econometrica*, **16**: 1-32.

[31] Powell, J. (1984), Least Absolute Deviations Estimation for the Censored Regression Model, *Journal of Econometrics*, **25**: 303-25.

[32] Wooldridge, J.M. (2005), Simple solutions to the initial conditions problem in dynamic, nonlinear panel-data models with unobserved heterogeneity, *Journal of Applied Econometrics*, **20**: 39-54

Table 1a. Results of MCE_1: A normal explanatory variable. *Initial values are known*

		$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
T		Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE
3		-0.001	0.121	-0.012	0.079	0.010	0.151	0.023	0.112	-0.012	0.264	-0.011	0.130	0.008	0.124	-0.012	0.082
5		0.005	0.070	0.005	0.045	0.008	0.062	0.013	0.043	-0.010	0.162	-0.013	0.090	0.011	0.055	0.008	0.053
8		0.002	0.045	0.003	0.036	0.007	0.042	0.007	0.032	0.004	0.125	0.011	0.094	0.005	0.035	0.005	0.027
15		0.005	0.035	0.007	0.022	0.005	0.033	0.004	0.019	0.007	0.099	0.004	0.098	0.002	0.025	0.001	0.019
20		0.003	0.031	0.004	0.021	0.004	0.026	0.004	0.016	0.008	0.085	0.003	0.064	-0.001	0.021	-0.003	0.017
		$\beta=1$				$\gamma=-0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
3		-0.008	0.124	0.022	0.087	0.011	0.152	0.016	0.137	-0.016	0.244	-0.012	0.105	-0.005	0.118	-0.009	0.075
5		-0.004	0.075	0.008	0.046	0.010	0.079	0.011	0.065	0.014	0.133	0.019	0.096	-0.006	0.057	-0.003	0.036
8		0.005	0.062	0.007	0.044	0.005	0.057	0.008	0.042	0.013	0.127	0.014	0.082	0.003	0.042	0.001	0.029
15		0.002	0.048	0.006	0.030	0.004	0.035	0.009	0.027	0.005	0.095	0.007	0.059	0.001	0.026	0.000	0.022
20		0.006	0.033	-0.005	0.022	0.005	0.027	-0.002	0.019	0.002	0.092	0.005	0.043	0.002	0.024	0.002	0.016
		$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=\sqrt{3}$				$\sigma_u=1$			
3		-0.016	0.139	-0.014	0.089	0.021	0.162	0.017	0.139	-0.024	0.266	-0.021	0.206	0.006	0.102	0.017	0.087
5		-0.007	0.074	-0.008	0.047	0.016	0.078	0.013	0.077	-0.019	0.143	-0.014	0.148	0.007	0.062	0.009	0.036
8		0.001	0.056	-0.004	0.037	0.008	0.054	0.011	0.054	0.010	0.104	0.016	0.104	-0.001	0.039	0.005	0.026
15		0.005	0.035	0.003	0.027	0.005	0.046	0.008	0.028	0.014	0.102	0.018	0.066	0.004	0.026	0.004	0.019
20		0.003	0.032	0.001	0.022	0.005	0.031	0.004	0.020	0.009	0.099	0.015	0.045	0.003	0.025	0.003	0.015
T	N	$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
5	300	0.005	0.045	-0.003	0.028	0.011	0.060	0.012	0.042	-0.009	0.122	-0.004	0.093	0.012	0.042	0.008	0.031
5	500	0.004	0.031	0.005	0.018	0.010	0.056	0.013	0.038	-0.001	0.114	-0.009	0.087	0.011	0.033	0.010	0.027
5	750	0.003	0.030	0.004	0.019	0.011	0.056	0.011	0.038	0.002	0.103	0.003	0.086	0.013	0.027	0.009	0.023
5	1000	0.003	0.027	0.003	0.013	0.009	0.052	0.010	0.037	0.007	0.097	0.007	0.081	0.012	0.025	0.007	0.020

Note: Monte Carlo design is based on a standard normal explanatory variable $x_{it} \sim N[0,1]$; Initial values are all constant (censored $y_{i0}=0$); $\beta=1$; $\gamma=(0.5,-0.5)$; $\sigma_\alpha=(1,\sqrt{3})$; $\sigma_u=1$; All results are based on $L=200$ Monte Carlo replications; number of individuals is $N=200$, if else is not stated; number of nodes in Gaussian-Hermite Quadrature is 30; RMSE is the root mean square error; and MAE is the mean absolute error.

Table 1b. Results of MCE_1: A normal explanatory variable. *Naïve exogenous initial values method (when it is wrong)*

T	N	$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
		Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE
3		-0.041	0.125	-0.038	0.083	0.242	0.265	0.245	0.245	-0.461	0.419	-0.445	0.445	0.057	0.115	0.059	0.067
5		-0.023	0.076	-0.018	0.057	0.099	0.146	0.098	0.098	-0.174	0.177	-0.165	0.165	0.024	0.053	0.028	0.040
8		-0.012	0.061	-0.015	0.052	0.038	0.081	0.038	0.058	-0.081	0.134	-0.082	0.093	0.011	0.039	0.010	0.029
15		0.004	0.040	0.005	0.025	0.014	0.034	0.016	0.021	-0.031	0.101	-0.025	0.048	0.008	0.023	0.006	0.016
20		0.001	0.025	0.002	0.024	0.011	0.027	0.009	0.019	-0.015	0.094	-0.013	0.042	0.002	0.020	0.005	0.015
		$\beta=1$				$\gamma=-0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
3		-0.031	0.118	-0.041	0.089	0.220	0.216	0.234	0.234	-0.499	0.438	-0.471	0.471	0.046	0.128	0.042	0.095
5		-0.014	0.082	-0.019	0.056	0.108	0.101	0.108	0.108	-0.180	0.227	-0.176	0.176	0.031	0.070	0.030	0.046
8		0.008	0.055	0.011	0.033	0.036	0.057	0.039	0.040	-0.066	0.123	-0.063	0.083	0.012	0.045	0.017	0.025
15		0.005	0.039	0.004	0.021	0.015	0.043	0.018	0.022	-0.020	0.115	-0.021	0.056	0.005	0.024	-0.004	0.021
20		-0.001	0.035	-0.002	0.020	0.008	0.036	0.011	0.023	-0.012	0.104	-0.010	0.043	-0.001	0.022	0.001	0.017
		$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=\sqrt{3}$				$\sigma_u=1$			
3		-0.053	0.151	-0.066	0.117	0.351	0.388	0.360	0.360	-1.154	1.112	-1.101	1.001	0.052	0.161	0.055	0.103
5		-0.027	0.081	-0.025	0.061	0.211	0.252	0.222	0.222	-0.589	0.567	-0.556	0.556	0.035	0.098	0.039	0.079
8		-0.014	0.050	-0.013	0.035	0.072	0.096	0.081	0.083	-0.161	0.255	-0.125	0.224	0.024	0.055	0.025	0.042
15		0.008	0.045	0.007	0.024	0.032	0.047	0.034	0.032	-0.065	0.140	-0.061	0.103	0.011	0.030	0.011	0.031
20		0.003	0.034	-0.004	0.023	0.021	0.036	0.019	0.026	-0.033	0.115	-0.028	0.075	0.013	0.027	0.011	0.024
T	N	$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
5	300	-0.015	0.046	-0.017	0.023	0.092	0.136	0.093	0.093	-0.153	0.158	-0.156	0.156	0.022	0.047	0.026	0.034
5	500	-0.011	0.037	-0.010	0.024	0.091	0.130	0.092	0.092	-0.154	0.155	-0.152	0.152	0.025	0.045	0.026	0.032
5	750	-0.008	0.028	-0.007	0.022	0.091	0.129	0.093	0.093	-0.153	0.160	-0.151	0.151	0.021	0.044	0.024	0.032
5	1000	-0.005	0.022	-0.003	0.015	0.090	0.125	0.091	0.091	-0.151	0.161	-0.151	0.151	0.021	0.044	0.023	0.031

Note: Monte Carlo design is based on a standard normal explanatory variable $x_{it} \sim N[0,1]$; $y_{i0} = \max(0, x_{i0}/(1-\gamma) + \alpha_i/(1-\gamma) + u_{i0}/\sqrt{1-\gamma^2})$, $y_{it} = \max(0, \beta x_{it} + \gamma y_{i,t-1} + \alpha_i + u_{it})$ is operated through T = 25 periods before the samples are observed; $\beta = 1$; $\gamma = (0.5, -0.5)$; $\sigma_\alpha = (1, \sqrt{3})$; $\sigma_u = 1$; All results are based on $L = 200$ Monte Carlo replications; number of individuals is $N = 200$, if else is not stated; number of nodes in Gaussian-Hermite Quadrature is 30; RMSE is the root mean square error; and MAE is the mean absolute error.

Table 1c. Results of MCE_1: A normal explanatory variable. *Heckman's approximation*

T	N	$\beta=1$ Mean Bias	RMSE	Median Bias	MAE	$\gamma=0.5$ Mean Bias	RMSE	Median Bias	MAE	$\sigma_\alpha=1$ Mean Bias	RMSE	Median Bias	MAE	$\sigma_u=1$ Mean Bias	RMSE	Median Bias	MAE
3		-0.018	0.097	-0.012	0.048	0.016	0.081	0.012	0.054	-0.052	0.217	-0.070	0.128	0.035	0.093	0.041	0.072
5		-0.015	0.065	-0.014	0.041	0.011	0.059	0.011	0.041	-0.028	0.137	-0.027	0.090	0.026	0.057	0.031	0.053
8		-0.010	0.062	-0.008	0.036	0.012	0.050	0.010	0.030	-0.019	0.105	-0.017	0.070	0.015	0.041	0.029	0.032
15		-0.008	0.032	-0.011	0.023	0.011	0.035	0.010	0.026	-0.005	0.095	-0.008	0.068	0.013	0.030	0.016	0.028
20		-0.005	0.030	-0.006	0.019	0.008	0.029	0.007	0.024	-0.001	0.091	-0.003	0.060	0.012	0.025	0.014	0.019
		$\beta=1$				$\gamma=-0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
3		-0.013	0.115	-0.012	0.055	0.015	0.101	0.013	0.081	-0.055	0.205	-0.062	0.138	0.016	0.081	0.015	0.066
5		-0.004	0.081	-0.005	0.045	0.014	0.092	0.015	0.062	-0.027	0.142	-0.027	0.091	0.012	0.070	0.011	0.047
8		0.008	0.064	0.009	0.042	0.012	0.068	0.008	0.055	-0.015	0.099	-0.021	0.080	0.010	0.051	0.012	0.031
15		0.003	0.041	0.011	0.036	0.008	0.054	0.006	0.030	-0.012	0.097	-0.008	0.075	0.011	0.033	0.010	0.022
20		-0.006	0.033	-0.008	0.022	0.009	0.036	0.008	0.026	-0.005	0.089	-0.004	0.064	0.006	0.027	0.008	0.019
		$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=\sqrt{3}$				$\sigma_u=1$			
3		-0.028	0.106	-0.025	0.082	0.031	0.091	0.030	0.088	-0.082	0.291	-0.097	0.242	0.049	0.104	0.051	0.081
5		-0.017	0.078	-0.023	0.057	0.016	0.066	0.017	0.071	-0.038	0.156	-0.041	0.145	0.031	0.078	0.042	0.055
8		0.022	0.062	0.015	0.045	0.018	0.061	0.013	0.049	-0.012	0.112	-0.020	0.085	0.023	0.061	0.030	0.038
15		-0.015	0.034	-0.012	0.024	0.017	0.046	0.015	0.038	-0.012	0.107	-0.014	0.082	0.012	0.039	0.026	0.025
20		0.010	0.031	0.009	0.021	0.012	0.037	0.010	0.027	-0.011	0.098	-0.007	0.071	0.007	0.027	0.012	0.019
T	N	$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
5	300	-0.013	0.052	-0.014	0.031	0.008	0.042	0.008	0.026	-0.018	0.124	-0.029	0.081	0.024	0.052	0.030	0.043
5	500	-0.013	0.046	-0.013	0.027	-0.005	0.039	0.004	0.023	-0.015	0.112	-0.017	0.078	0.024	0.050	0.029	0.041
5	750	-0.012	0.045	-0.014	0.029	0.004	0.031	0.007	0.025	-0.016	0.101	-0.016	0.075	0.022	0.048	0.030	0.045
5	1000	-0.011	0.038	-0.012	0.022	0.004	0.026	-0.003	0.018	-0.014	0.102	-0.016	0.074	0.020	0.046	0.031	0.047

Note: Monte Carlo design is based on a standard normal explanatory variable $x_{it} \sim N[0,1]$; $y_{i0} = \max(0, x_{i0}/(1-\gamma) + \alpha_i/(1-\gamma) + u_{i0}/\sqrt{1-\gamma^2})$, $y_{it} = \max(0, \beta x_{it} + \gamma y_{i,t-1} + \alpha_i + u_{it})$ is operated through T = 25 periods before the samples are observed; $\beta = 1$; $\gamma = (0.5, -0.5)$; $\sigma_\alpha = (1, \sqrt{3})$; $\sigma_u = 1$; All results are based on $L = 200$ Monte Carlo replications; number of individuals is $N = 200$, if else is not stated; number of nodes in Gaussian-Hermite Quadrature is 30; RMSE is the root mean square error; and MAE is the mean absolute error.

Table 1d. Results of MCE_1: A normal explanatory variable. *Simple method of Wooldridge*

		$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
T		Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE
3		0.045	0.104	0.051	0.082	0.098	0.188	0.111	0.139	-0.145	0.262	-0.152	0.185	0.039	0.103	0.038	0.071
5		0.014	0.086	0.017	0.071	0.022	0.074	0.023	0.051	-0.062	0.198	-0.051	0.117	0.011	0.065	0.012	0.044
8		0.011	0.052	0.012	0.052	0.001	0.030	0.002	0.023	-0.010	0.103	-0.010	0.082	0.004	0.041	0.009	0.027
15		-0.008	0.036	-0.005	0.023	-0.006	0.027	-0.006	0.021	-0.003	0.098	-0.005	0.074	-0.005	0.026	0.000	0.019
20		-0.006	0.030	-0.004	0.021	-0.002	0.021	-0.003	0.019	0.001	0.094	0.002	0.063	0.002	0.023	0.003	0.018
		$\beta=1$				$\gamma=-0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
3		0.031	0.123	0.043	0.118	0.096	0.166	0.099	0.136	-0.156	0.245	-0.153	0.179	0.024	0.145	0.028	0.083
5		0.013	0.085	0.021	0.073	0.025	0.079	0.029	0.057	-0.055	0.183	-0.061	0.121	0.008	0.075	0.011	0.051
8		0.010	0.067	0.009	0.056	0.010	0.036	0.014	0.029	-0.017	0.112	-0.029	0.094	0.004	0.046	0.013	0.033
15		0.011	0.042	0.008	0.032	-0.004	0.032	-0.003	0.025	-0.011	0.095	-0.017	0.082	0.003	0.032	0.008	0.021
20		0.005	0.029	-0.002	0.024	-0.007	0.026	-0.005	0.021	-0.005	0.093	-0.011	0.066	0.002	0.022	0.007	0.017
		$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=\sqrt{3}$				$\sigma_u=1$			
3		0.041	0.156	0.055	0.147	0.116	0.191	0.133	0.153	-0.336	0.326	-0.314	0.319	0.021	0.150	0.033	0.092
5		0.015	0.087	0.020	0.089	0.026	0.082	-0.021	0.077	-0.138	0.245	-0.144	0.178	0.010	0.081	0.017	0.061
8		0.010	0.058	0.017	0.065	0.021	0.036	-0.024	0.057	-0.067	0.158	-0.055	0.125	-0.010	0.053	-0.011	0.037
15		-0.009	0.046	-0.011	0.040	-0.002	0.035	-0.012	0.033	-0.031	0.121	-0.028	0.084	-0.007	0.036	-0.011	0.029
20		-0.008	0.025	-0.010	0.022	-0.004	0.024	-0.012	0.024	-0.011	0.101	-0.010	0.067	-0.004	0.023	-0.009	0.023
T	N	$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
5	300	0.017	0.068	0.012	0.056	0.021	0.067	0.020	0.046	-0.051	0.175	-0.053	0.089	0.010	0.047	0.010	0.033
5	500	0.015	0.054	0.011	0.042	0.020	0.052	0.019	0.042	-0.049	0.167	-0.051	0.086	0.010	0.048	0.011	0.031
5	750	0.015	0.048	0.013	0.038	0.018	0.040	0.018	0.040	-0.050	0.152	-0.048	0.075	0.011	0.036	0.011	0.030
5	1000	0.013	0.047	0.013	0.027	0.019	0.039	0.020	0.041	-0.042	0.135	-0.044	0.072	0.009	0.029	0.008	0.027

Note: Monte Carlo design is based on a standard normal explanatory variable $x_{it} \sim N[0,1]$; $y_{i0} = \max(0, x_{i0}/(1-\gamma) + \alpha_i/(1-\gamma) + u_{i0}/\sqrt{1-\gamma^2})$, $y_{it} = \max(0, \beta x_{it} + \gamma y_{i,t-1} + \alpha_i + u_{it})$ is operated through T = 25 periods before the samples are observed; $\beta = 1$; $\gamma = (0.5, -0.5)$; $\sigma_\alpha = (1, \sqrt{3})$; $\sigma_u = 1$; All results are based on $L = 200$ Monte Carlo replications; number of individuals is $N = 200$, if else is not stated; number of nodes in Gaussian-Hermite Quadrature is 30; RMSE is the root mean square error; and MAE is the mean absolute error.

Table 2. Results of MCE_2 : A non-normal explanatory variable (*chi-squared distribution with one degrees of freedom*).

	$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
T	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE
Known initial values																
3	0.005	0.096	0.013	0.054	-0.003	0.086	-0.005	0.043	-0.010	0.120	-0.005	0.071	0.004	0.121	0.005	0.072
5	0.006	0.057	0.006	0.035	-0.001	0.046	0.001	0.030	-0.004	0.073	-0.007	0.055	0.002	0.074	-0.006	0.041
8	0.007	0.042	0.004	0.024	0.004	0.025	0.004	0.018	-0.003	0.072	-0.006	0.050	0.003	0.045	-0.001	0.029
15	0.005	0.034	0.003	0.024	0.001	0.018	0.001	0.014	-0.009	0.063	-0.003	0.036	0.002	0.029	0.003	0.017
20	0.001	0.021	0.001	0.023	0.000	0.015	-0.001	0.010	-0.007	0.060	-0.006	0.035	0.001	0.020	0.002	0.016
Naïve exogenous initial values method (when it is wrong)																
3	-0.016	0.088	-0.017	0.056	0.219	0.218	0.211	0.211	-0.542	0.612	-0.412	0.412	0.042	0.116	0.042	0.076
5	0.013	0.066	0.015	0.039	0.102	0.107	0.119	0.119	-0.158	0.178	-0.156	0.156	0.015	0.064	0.015	0.043
8	0.007	0.045	0.009	0.031	0.046	0.043	0.044	0.045	-0.066	0.093	-0.060	0.065	0.011	0.047	0.004	0.028
15	0.006	0.032	0.008	0.015	0.011	0.018	0.008	0.014	-0.023	0.062	-0.018	0.041	0.009	0.029	0.003	0.020
20	0.004	0.024	0.006	0.014	0.005	0.016	0.007	0.010	-0.011	0.055	-0.010	0.038	0.007	0.025	0.005	0.017
Heckman's approximation																
3	-0.011	0.095	-0.010	0.065	-0.001	0.081	0.002	0.048	-0.058	0.217	-0.070	0.128	0.022	0.096	0.020	0.074
5	0.010	0.065	0.008	0.042	0.005	0.053	0.003	0.030	-0.028	0.137	-0.027	0.090	0.018	0.045	0.015	0.056
8	0.008	0.042	0.005	0.026	0.001	0.034	0.001	0.021	-0.019	0.105	-0.017	0.072	0.011	0.044	0.014	0.035
15	-0.005	0.030	-0.004	0.022	0.000	0.028	-0.003	0.018	-0.005	0.071	-0.008	0.058	0.008	0.036	0.010	0.022
20	-0.002	0.022	-0.002	0.015	0.000	0.014	0.001	0.012	-0.001	0.062	-0.003	0.040	0.004	0.028	0.008	0.020
Simple method of Wooldridge																
3	0.011	0.186	0.014	0.067	0.098	0.147	0.110	0.139	-0.175	0.252	-0.172	0.163	0.020	0.108	0.038	0.075
5	0.010	0.085	0.015	0.048	0.029	0.074	0.030	0.051	-0.061	0.198	-0.057	0.107	0.010	0.065	0.012	0.046
8	0.009	0.045	0.010	0.029	0.001	0.030	0.002	0.021	-0.010	0.103	-0.010	0.062	0.008	0.037	0.010	0.025
15	0.008	0.028	0.007	0.017	-0.006	0.025	-0.006	0.016	-0.003	0.069	-0.005	0.043	0.003	0.026	0.006	0.026
20	-0.004	0.023	-0.005	0.015	-0.002	0.015	-0.003	0.012	-0.001	0.053	0.002	0.039	0.002	0.025	0.007	0.023

Note: Monte Carlo design is based on a standardized one degrees of freedom chi-squared explanatory variable $x_{it} \sim \chi^2_{(1)} - 1\big/2$; for the case initial values are known $y_{i0} = 0$; for other cases the process $y_{i0} = \max(0, x_{i0}/(1-\gamma) + \alpha_i/(1-\gamma) + u_{i0}/\sqrt{1-\gamma^2})$, $y_{it} = \max(0, \beta x_{it} + \gamma y_{i,t-1} + \alpha_i + u_{it})$ is operated through T = 25 periods before the samples are observed; $(\beta, \gamma, \sigma_\alpha, \sigma_u) = (1, 0.5, 1, 1)$. All results are based on $L = 200$ Monte Carlo replications; number of individuals is $N = 200$; number of nodes in Gaussian-Hermite Quadrature is 30; RMSE is the root mean square error; and MAE is the mean absolute error.

Table 3. Results of MCE_3: An autocorrelated explanatory variable

T	$\beta=1$				$\gamma=0.5$				$\sigma_\alpha=1$				$\sigma_u=1$			
	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE	Mean Bias	RMSE	Median Bias	MAE
Known initial values																
3	0.009	0.050	-0.002	0.046	-0.015	0.140	-0.017	0.096	-0.024	0.200	-0.018	0.114	0.012	0.149	0.019	0.097
5	0.004	0.048	0.006	0.040	-0.003	0.065	0.002	0.047	-0.015	0.125	-0.014	0.072	0.008	0.075	0.010	0.070
8	0.004	0.047	0.004	0.034	-0.005	0.041	0.001	0.030	-0.009	0.095	-0.008	0.063	0.005	0.053	0.008	0.041
15	0.006	0.035	0.003	0.029	-0.000	0.023	0.001	0.015	-0.015	0.089	-0.015	0.061	0.003	0.034	0.006	0.031
20	0.005	0.020	0.003	0.014	-0.001	0.021	0.000	0.014	-0.010	0.088	-0.016	0.060	0.004	0.029	0.002	0.018
Naïve exogenous initial values method (when it is wrong)																
3	-0.046	0.056	-0.034	0.041	0.207	0.214	0.208	0.208	-0.487	0.587	-0.482	0.482	0.053	0.228	0.047	0.199
5	-0.017	0.038	-0.016	0.026	0.088	0.104	0.093	0.093	-0.154	0.207	-0.167	0.172	0.033	0.129	0.027	0.082
8	-0.003	0.033	-0.003	0.012	0.032	0.049	0.034	0.037	-0.068	0.119	-0.062	0.079	0.012	0.075	0.016	0.041
15	0.003	0.015	0.004	0.012	0.005	0.028	0.004	0.022	-0.022	0.093	-0.024	0.066	0.004	0.036	0.005	0.029
20	0.002	0.011	0.006	0.010	0.002	0.021	0.002	0.014	-0.010	0.092	-0.012	0.058	0.006	0.034	0.007	0.020
Heckman's approximation																
3	0.004	0.043	0.005	0.024	0.001	0.088	-0.002	0.063	-0.043	0.219	-0.052	0.150	0.014	0.113	0.011	0.077
5	0.007	0.024	0.008	0.018	0.004	0.049	0.003	0.033	-0.030	0.136	-0.023	0.081	0.011	0.067	0.005	0.051
8	0.004	0.021	0.003	0.016	0.005	0.038	0.002	0.024	-0.015	0.111	-0.014	0.072	0.005	0.046	0.009	0.042
15	0.003	0.020	0.002	0.013	0.001	0.025	0.002	0.019	-0.011	0.095	-0.012	0.066	0.004	0.035	0.007	0.029
20	0.002	0.014	0.002	0.012	0.001	0.021	0.000	0.015	-0.009	0.089	-0.010	0.061	0.007	0.033	0.009	0.026
Simple method of Wooldridge																
3	-0.022	0.053	0.011	0.054	0.064	0.181	0.056	0.108	-0.083	0.290	-0.091	0.191	0.011	0.145	0.017	0.117
5	0.011	0.044	0.010	0.042	0.026	0.076	0.024	0.046	-0.027	0.148	-0.028	0.093	0.013	0.088	0.014	0.062
8	0.009	0.041	0.010	0.021	-0.001	0.042	-0.003	0.028	-0.017	0.109	-0.011	0.074	0.008	0.062	0.009	0.054
15	0.004	0.022	0.007	0.017	0.003	0.029	-0.011	0.019	-0.016	0.088	-0.010	0.062	0.007	0.054	-0.001	0.027
20	0.005	0.020	0.006	0.014	0.000	0.022	-0.001	0.016	-0.011	0.087	-0.009	0.059	-0.003	0.042	0.006	0.022

Note: Monte Carlo design is based on an autocorrelated explanatory variable $x_{it} = \rho x_{i,t-1} + \psi_{it}$, $\psi_{it} \sim N[0,1]$, $\rho = 0.5$ and $x_{i0} = \psi_{i1}$; for the case initial values are known $y_{i0} = 0$; for other cases the process $y_{i0} = \max(0, x_{i0}/(1-\gamma) + \alpha_i/(1-\gamma) + u_{i0}/\sqrt{1-\gamma^2})$, $y_{it} = \max(0, \beta x_{it} + \gamma y_{i,t-1} + \alpha_i + u_{it})$ is operated through T = 25 periods before the samples are observed; $(\beta, \gamma, \sigma_\alpha, \sigma_u) = (1, 0.5, 1, 1)$. All results are based on $L = 200$ Monte Carlo replications; number of individuals is $N = 200$; number of nodes in Gaussian-Hermite Quadrature is 30; RMSE is the root mean square error; and MAE is the mean absolute error.

Figure 1. Q-Q plots based on the quantiles of normal distribution by solution methods for true state-dependence and the variance of unobserved individual-effects. $L = 200$ estimated Monte Carlo parameters based on $T = 10$ and the design; $x_{it} = \rho x_{it-1} + \psi_{it}$, $\psi_{it} \sim N[0, 1]$, $\rho = 0.5$ and $x_{i0} = \psi_{i1}$; $y_{i0} = \max(0, \beta x_{i0}/(1-\gamma)+\alpha_i/(1-\gamma)+u_{i0}/\sqrt{1-\gamma^2})$; $y_{it} = \max(0, \beta x_{it}+\gamma y_{i,t-1}+\alpha_i+u_{it})$; $(\beta, \gamma, \sigma_\alpha, \sigma_u) = (1, 0.5, 1, 1)$; and number of individuals is $N = 200$.

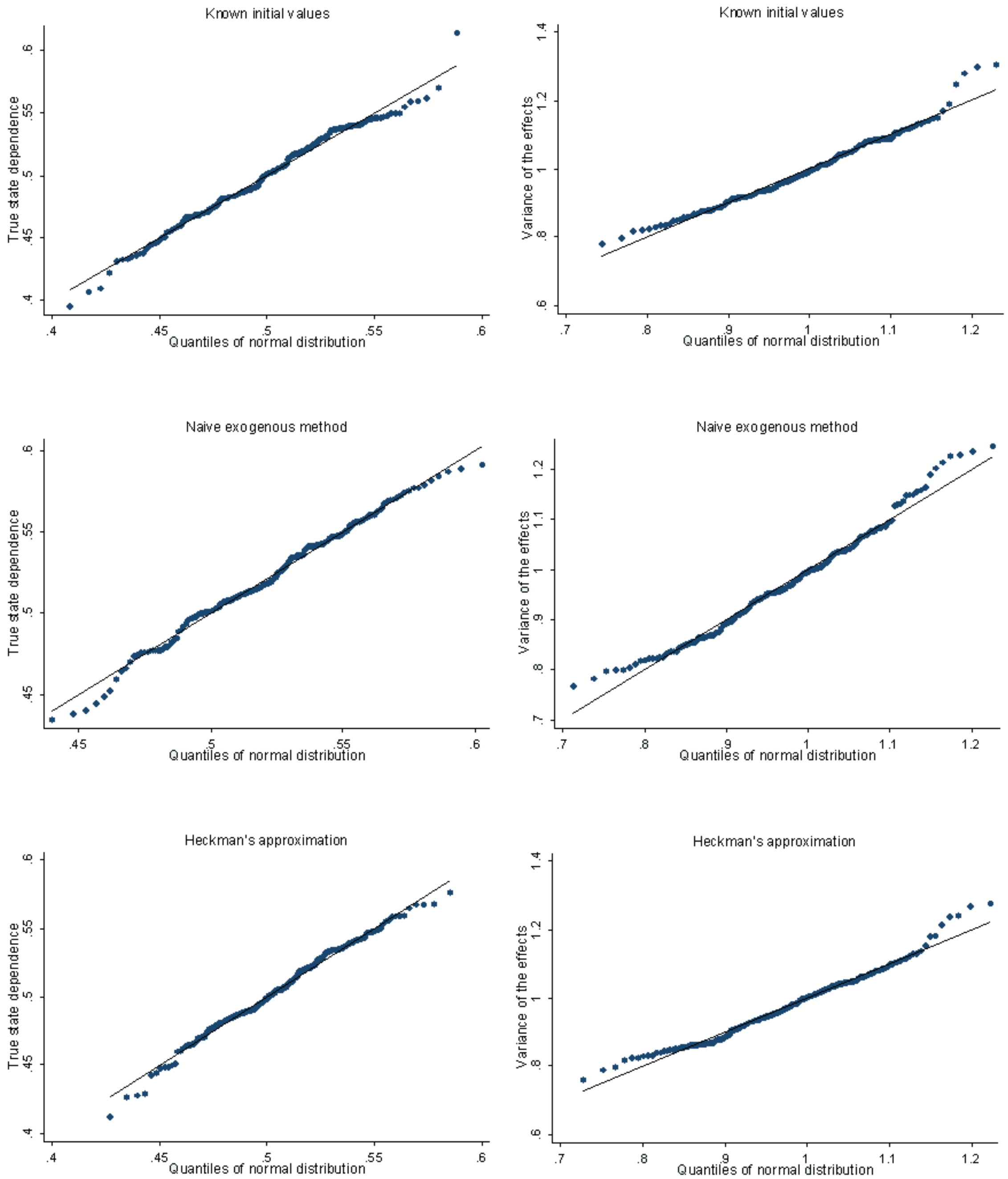

Figure 1. *Continued*