Boris Lokshin

Monte-Carlo comparison of alternative estimators for dynamic panel data models

RM/06/014

JEL code: C23

Maastricht research school of **E**conomics
of **TE**chnology and **OR**ganizations

Universiteit Maastricht
Faculty of Economics and Business Administration
P.O. Box 616
NL - 6200 MD Maastricht

phone : ++31 43 388 3830
fax : ++31 43 388 4873

Monte-Carlo comparison of alternative estimators for dynamic panel data models

Boris Lokshin[a]

[a] *Faculty of Economics and Business Administration, University of Maastricht, and UNU-MERIT*
PO Box 616, 6200 MD Maastricht, The Netherlands
E-mail: B.Lokshin@os.unimaas.nl

Abstract

This paper compares the performance of three recently proposed estimators for dynamic panel data models (LSDV bias-corrected, MLE and MDE) along with GMM. Using Monte-Carlo, we find that MLE and bias-corrected estimators have the smallest bias and are good alternatives for the GMM. System-GMM outperforms the rest in 'difficult' designs. Unfortunately, bias-corrected estimator is not reliable in these designs which may limit its applicability.

Keywords: Bias correction, Dynamic panel data, GMM, MLE

JEL classification: C23

1. Introduction.

The purpose of this note is to evaluate several recently proposed techniques for estimating dynamic panel data models when T is small. Such panels are characteristic of many microeconomic panel data sets used in applied research. Three of the recently proposed estimators for such samples, LSDV-bias-corrected (Bun and Carree, 2005), transformed fixed-effects MLE and MDE (Hsiao et al., 2002) are compared to each other and to the two, difference and system, GMM estimators (Blundell and Bond, 1998). When it comes to the estimation of a dynamic panel data model, GMM is often used, not in the least because the routine is now available in such popular packages as Stata, EViews, Gauss, PcGive and Limdep[1]. One drawback of the GMM is that it requires additional decisions by the researcher on the instruments to be used, and the estimator's growing finite-sample imprecision with an increase in instruments used.

We conduct Monte-Carlo simulations in which our goal is to provide a recommendation for an applied researcher on an appropriate technique. In particular, we focus on two issues: the accuracy of the previous studies' assessments that the new estimators are more precise than an established GMM; and second, the practicality of the new estimation techniques. The first issue is relevant because the previous comparisons were incomplete (only difference GMM was considered in previous simulations) or presented selectively. For example, the present only those results where γ MLE converged to a value less than unity and compare MLE to difference GMM only.

The second issue, practicality of the techniques, is especially relevant for applied researchers. We show for example, that the bias-corrected estimator may not have a solution altogether, while the MLE has been reported to have convergence problems in certain designs[2]. We investigate how severe these problems are, which in fact may affect the attractiveness of these estimators for an applied researcher.

The rest of the paper is organized as follows. Section two presents the empirical model and describes the Monte-Carlo simulation designs. Section three discusses results and section four presents an empirical example. Section five concludes.

2. The problem and methodology

We consider the dynamic fixed effects model

$$y_{it} = \gamma y_{i,t-1} + x_{it}\beta + \eta_i + \varepsilon_{it}, \qquad i = 1,....,N; t = 1,....,T \tag{2.1}$$

The dependent variable, y_{it}, is determined by the one-period own lag, an exogenous regressor, x_{it}, an unobserved individual effect η_i, and a random disturbance $\varepsilon_{it} \sim N(0,\sigma_\varepsilon^2), \sigma_\varepsilon^2 > 0$. We assume that x_{it} is not correlated with the general disturbance term.

In our Monte-Carlo experiments we follow Kiviet (1995). Data for y_{it} are generated by model (2.1) and the data for x_{it} are generated by

[1] Stata routine -xtabond2- written by David Roodman, Center for Global Development, Washington, DC is capable of estimating both system and difference GMM. E-views 5.0 release has a difference GMM available. Gauss and PcGive have an original DPD. Limdep 8.0 can estimate a form of difference GMM.

[2] In particular, Bun and Carree (2005) state that the BC estimator has convergence problems when the signal-to-noise ratio is low for relatively high values of γ (p. 207). Hsiao et al. (2002) state that "MLE sometimes breaks down completely… [when T and N are small] " (p. 132). See the footnote 3 below for our approach to tracking non-convergence cases.

$$x_{it} = \rho x_{i,t-1} + \xi_{it}, \qquad \xi_{it} \sim N(0, \sigma_\xi^2), \quad i = 1,...,N \text{ and } t = 1,...,T \qquad (3.1)$$

The individual effects η_i are generated by assuming $\eta_i \sim N(0, \sigma_\eta^2)$ while σ_ε is normalized to unity. In addition to β and σ_ξ^2, ρ also determines the correlation between y_{it} and x_{it} and is set at values 0.8 and 0.2. In Kiviet (1995) it is argued that the relative bias of the estimators is significantly influenced by σ_s^2, the signal-to-noise ratio of the regression. In our experiments we use a combination of relatively high σ_s^2=8 with high and low correlation and relatively low σ_s^2=2 with high and low ρ. Two different sample sizes are considered (*N, T*) = (100, 6) and (*N, T*) = (100, 3). Table 1 summarizes the parameters for each of the 16 panel designs. In generating y_{it} and x_{it} we set $y_{i,-50}$, $xi_{,-50}$ =0 and discarded the first 50 observations, using the observations t through T for estimation. We performed 1000 replications with a fixed seed for the random number generation.

3. Monte-Carlo results

The results for the 16 designs are summarized in Tables 2, and 3. Unfortunately, while the bias-corrected estimator may produce superior results in terms of bias, it is not always practical. When the signal-to-noise ratio σ_s^2 is low (designs 3, 4, 11, 12) there is a large percentage of cases with no solution[3]. The problem is exacerbated when T is relatively low, and can lead to almost 45% of such cases. We note that even with high σ_s^2, the combination of relatively high γ and ρ can lead to high share of no-solution scenario when T is relatively small (design 15). MDE and MLE estimation involves estimation of *KT* auxiliary parameters (formula B.1 in Hsiao et al., 2002), which makes matrix $\Delta\widetilde{W}_i$ relatively large even for problems of moderate size. When N is not much larger than *KT*, some authors suggest projecting Δy_{i1} on $\Delta\bar{x}_i$ rather than on Δx_i. We computed the version of MLE and MDE where $\Delta\bar{x}_i$ is used in place of Δx_i. However, this change leads to larger bias for β in MLE and MDE, while bias for γ is actually somewhat smaller. These additional results are not included in the tables and are available upon request.

[3] To avoid over- or under-reporting of non-convergence of the BC estimator, in our version of the program we rely on the exact solution, rather than on the iterative procedure proposed in the original paper. The system of equations (19) and (20) (see original paper) can be transformed into a polynomial of power T. We solve this polynomial with respect to γ. Then we use (20) to solve for true β. Note, that when *T* is odd the polynomial always has at least one real root, when *T* is even, it may have zero real roots and *T* complex roots. In our Monte-Carlo experiments, for the designs when *T*=6 we count the number of cases when the polynomial has no real roots, this is reported in Table 2; in designs when *T*=3, when there is at least one real root, we count as non-convergence those solutions that are smaller than γ-LSDV. The unique solution to the polynomial when *T*=3 and the root is less than γ-LSDV is a negative number between -3 and -4. Convergence for MLE has been always achieved. The starting values for MLE were the MDE estimates. We used NLPDD Double Dogleg optimization routine in SAS for MLE and used the outer product of gradient to compute variance. Estimates for MDE are obtained from B.5 based on initial σ_u^2 and γ, β from Anderson-Hsiao IV estimator, using $y_{i,t-2}$ and $(x_{i,t} - x_{i,t-1})$ as instruments.

4. Application to US unemployment dynamics

We applied the various methods for estimating model (2.1) to the data set used in Bun and Carree (2005) for an analysis of unemployment dynamics at the US state level observed over 1991-2000. In the model the current unemployment rate is a function of the previous period's unemployment and economic growth rate. The equation also includes the time and individual fixed effects.

Table 4 presents the recomputed results (original paper) for the BC estimator along with the MLE, MDE and GMM, estimators that we test in a Monte Carlo. For the GMM results the lagged growth rate is treated as exogenous and an IV-style instrument for it is used. MLE and BC estimates are virtually identical, implying an adjustment rate of about 40% a year. The coefficient on the lagged dependent variable is higher for both GMM estimators, while the effect of economic growth is smaller compared to the bias-corrected estimates. The former result would suggest a lower adjustment rate between 23% and 30% a year.

5. Conclusions

We have three main conclusions. MLE has the lowest average bias for γ and has the highest average ranking among all the estimators in all designs considered[4]. System-, but not difference-GMM performs better than the other estimators in "difficult" designs, with low signal-to-noise ratio coupled with high values of γ ; it has also the lowest RMSE for γ and β among considered estimators. Bias-corrected estimator can match MLE in terms of bias, and System-GMM in terms of RMSE, but unfortunately has a very high percentage of non-convergence in "difficult" designs, which makes its use in such cases problematic for applied researchers.

Acknowledgement: I thank Martin Carree for useful comments

[4] The ranking is computed by giving 5 points for the first place, 4 for the second and so on in each of the designs, and is based on average cumulative score for biases on γ and β .

References

Bun, M. and Carree M. (2005) Bias-corrected estimation in dynamic panel data models. *Journal of Business and Economic Statistics,* 23(2), 200-10.

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

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

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

Table 1 Summary of Monte-Carlo simulation parameters, N=100, σ_{ε}=1.

Design number	T	γ	β	σ_s^2	ρ	σ_{ξ}	σ_{η}	% break-downs BC
1	6	0.2	0.8	2	0.8	0.88	0.8	0
2				2	0.2	1.61	0.8	0
3	6	0.8	0.2	2	0.8	0.40	0.2	26.8
4				2	0.2	1.18	0.2	23.2
5	6	0.2	0.8	8	0.8	1.76	0.8	0
6				8	0.2	3.25	0.8	0
7	6	0.8	0.2	8	0.8	2.10	0.2	0.6
8				8	0.2	6.24	0.2	0
9	3	0.2	0.8	2	0.8	0.88	0.8	0
10				2	0.2	1.61	0.8	0
11	3	0.8	0.2	2	0.8	0.40	0.2	43.0
12				2	0.2	1.18	0.2	41.0
13	3	0.2	0.8	8	0.8	1.76	0.8	0
14				8	0.2	3.25	0.8	0
15	3	0.8	0.2	8	0.8	2.10	0.2	17.3
16				8	0.2	6.24	0.2	0.2

Table 2 Bias estimates and RMSE for γ using various estimators*

Des.	γ bias					γ RMSE				
	GMM1	GMM2	MLE	MDE	BC*	GMM1	GMM2	MLE	MDE	BC
1	-0.0199	0.0228	-0.0015	-0.0089	-0.0015	0.0044	0.0054	0.0016	0.0020	0.0016
2	-0.0073	0.0108	-0.0004	-0.0049	-0.0004	0.0016	0.0016	0.0008	0.0009	0.0008
3	-0.0611	-0.0028	0.0250	-0.0379	-0.0357	0.0143	0.0029	0.0151	0.0106	0.0060
4	-0.0551	-0.0014	0.0239	-0.0344	-0.0275	0.0124	0.0027	0.0136	0.0095	0.0052
5	-0.0127	0.0151	-0.0008	-0.0045	-0.0008	0.0026	0.0040	0.0007	0.0008	0.0007
6	-0.0025	0.0036	-0.0001	-0.0015	-0.0001	0.0005	0.0005	0.0003	0.0003	0.0003
7	-0.0590	-0.0075	0.0030	-0.0154	0.0017	0.0133	0.0035	0.0034	0.0073	0.0030
8	-0.0189	-0.0011	0.0004	-0.0076	0.0005	0.0032	0.0010	0.0010	0.0017	0.0010
9	-0.0226	-0.0021	0.0009	-0.0010	0.0003	0.0190	0.0141	0.0073	0.0099	0.0069
10	-0.0090	0.0008	0.0007	-0.0024	0.0005	0.0061	0.0049	0.0032	0.0035	0.0033
11	-0.0687	-0.0281	0.0223	-0.0250	-0.1451	0.0863	0.0143	0.0434	0.0496	0.0337
12	-0.0616	-0.0212	0.0323	-0.0211	-0.1246	0.0764	0.0127	0.0448	0.0457	0.0287
13	-0.0158	-0.0051	-0.0005	0.0007	-0.0005	0.0140	0.0148	0.0027	0.0039	0.0027
14	-0.0029	0.0004	0.0004	-0.0006	0.0003	0.0019	0.0017	0.0010	0.0010	0.0010
15	-0.0787	-0.0431	0.0330	0.0085	-0.0253	0.1100	0.0299	0.0310	0.0498	0.0137
16	-0.0222	-0.0087	0.0079	0.0069	0.0062	0.0223	0.0048	0.0076	0.0120	0.0065
AB	*0.0324*	*0.0109*	*0.0096*	*0.0113*	*0.0232*	*0.0243*	*0.0074*	*0.0111*	*0.0130*	*0.0072*

* The numbers for the BC estimator are based only on those cases where the exact solution was found (designs 3, 4, 10, 11, 15).

Table 3 Bias estimates and RMSE for β using various estimators*

Des.	β bias					β RMSE				
	GMM1	GMM2	MLE	MDE	BC	GMM1	GMM2	MLE	MDE	BC
1	0.0050	-0.0152	0.0017	0.0040	0.0017	0.0042	0.0041	0.0025	0.0025	0.0025
2	-0.0004	-0.0010	0.0005	0.0003	0.0005	0.0011	0.0009	0.0008	0.0008	0.0008
3	0.0084	0.0002	0.0037	0.0032	0.0076	0.0209	0.0056	0.0120	0.0112	0.0110
4	-0.0032	0.0004	0.0034	-0.0014	0.0006	0.0020	0.0014	0.0016	0.0015	0.0014
5	0.0024	-0.0097	0.0008	0.0021	0.0008	0.0010	0.0019	0.0007	0.0007	0.0007
6	0.0000	-0.0003	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
7	0.0014	0.0024	0.0007	0.0009	0.0007	0.0007	0.0006	0.0004	0.0004	0.0004
8	-0.0012	0.0001	0.0002	-0.0003	0.0002	0.0001	0.0000	0.0001	0.0001	0.0001
9	0.0015	0.0002	0.0038	0.0033	0.0038	0.0128	0.0085	0.0098	0.0099	0.0097
10	-0.0011	0.0000	0.0021	0.0014	0.0020	0.0033	0.0019	0.0024	0.0025	0.0025
11	0.0063	0.0094	0.0081	0.0061	0.0117	0.0624	0.0134	0.0506	0.0480	0.0406
12	-0.0029	0.0012	0.0061	0.0004	-0.0056	0.0063	0.0032	0.0050	0.0047	0.0039
13	0.0005	0.0025	0.0020	0.0017	0.0020	0.0033	0.0063	0.0024	0.0025	0.0024
14	-0.0001	0.0002	0.0009	0.0007	0.0009	0.0008	0.0005	0.0006	0.0006	0.0006
15	-0.0025	0.0144	0.0028	0.0014	0.0017	0.0026	0.0042	0.0018	0.0018	0.0017
16	-0.0016	0.0002	0.0012	0.0011	0.0010	0.0004	0.0001	0.0002	0.0003	0.0002
AB	*0.0024*	*0.0036*	*0.0024*	*0.0018*	*0.0026*	*0.0076*	*0.0033*	*0.0057*	*0.0055*	*0.0049*

*The numbers for the BC estimator are based only on those cases where the exact solution was found (designs 3, 4, 10, 11, 15).

Table 4 Empirical results for the unemployment-growth model

	BC	MLE Fixed effects[a]	Minimum distance	GMM1[b]	GMM2[c]
γ	0.6152***	0.6167***	0.5926***	0.6919***	0.7746***
SE(γ)	0.0463	0.0336	0.0335	0.0728	0.0631
β	-0.0567***	-0.0566***	-0.0579***	-0.0433***	-0.0430***
SE(β)	0.0120	0.0119	0.0118	0.0164	0.0130
LL		1747.88			

a. Standard errors are square roots of diagonals of the negative inverse of the Hessian
b. GMM1 is the difference, one-step estimator (Arellano and Bond)
c. GMM2 is the system, one-step estimator (Blundell and Bond)
*** Indicates significance at 99% level