Consistent OLS Estimation of AR(1) Dynamic Panel Data Models with Short Time Series

Kazuhiko Hayakawa*†
Department of Economics
Hitotsubashi University

January 19, 2006

Abstract

In this paper, we examine the usefulness of the modified OLS (MOLS) estimator by Wansbeek and Knaap (1999) from two angles: inference and testing. First, we compare the MOLS estimator with Bun and Carree's (2005) estimator and the GMM estimator in terms of accuracy of inference. Second, we propose to use the Hausman test based on these three estimators to test the null of no individual effects. Simulation results show that the MOLS estimator and the Hausman test based on the MOLS estimator perform better than the other two estimators.

*2-1, Naka, Kunitachi, Tokyo. 186-8601.

†I would like to thank Taku Yamamoto and Hiroaki Chigira for helpful comments. Financial support from the JSPS Fellowship is also gratefully acknowledge. All remaining errors are my own.

1 Introduction

Statistical analysis of dynamic panel data models is a topic of growing interest in both theoretical and applied econometrics. In the estimation of dynamic panel data models, the generalized method of moments (GMM) has been widely used since the work of Arellano and Bond (1991). However, when using instrumental variables, we must pay careful attention to the choice of instruments, because which instruments are chosen for the estimation, and how many, can lead to two serious problems, namely the weak instruments problem and the many instruments problem.[1] Hence, it is preferable to avoid using instrumental variables, if possible.

An alternative to the GMM estimator is the least squares dummy variable (LSDV) estimator. However, as Nickell (1981) has shown, the LSDV estimator is inconsistent under large N and fixed T asymptotics where N and T are the size of the cross-sectional and time series dimensions. Kiviet (1995), Hahn and Kuersteiner (2002), and Bun and Carree (2005) have tried to improve the performance of the LSDV estimator when T is small. Among the methods suggested, Bun and Carree's estimator performs very well when T is small. However, one of the drawbacks of their method is that it requires a numerical optimization procedure and it may be computationally burdensome.

In this paper, we consider the modified OLS (MOLS) estimator suggested by Wansbeek and Knaap (1999). Although the authors argued that the MOLS estimator is consistent when N is large and T is fixed, and very easy to compute, they did not provide any simulation results. The first purpose of this paper is to compare Wansbeek and Knaap's MOLS estimator with the bias-corrected LSDV estimator recently proposed by Bun and Carre and the GMM estimator via Monte Carlo simulation. The second aim is to propose a test procedure based on these three estimators and using the Hausman test.[2]

The remainder of the paper is organized as follows. In Section 2, we define the model and review the MOLS estimator. In Section 3, we propose a test for the absence of individual effects. In Section 4 we examine the performance of the estimators and the tests via Monte Carlo simulation and Section 5 concludes.

---

[1]On these problems, see Blundell and Bond (1998), Alvarez and Arellano (2003), Bun and Kiviet (2006) and Hayakawa (2005, 2006a, b).

[2]Few studies have dealt with the issue of testing for the absence of individual effects in dynamic panel data models. To the best of our knowledge, only Holtz-Eakin (1988) addresses with this problem.

2 The model and the modified OLS estimator

We consider an AR(1) panel data model given by

$$y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it} \quad i = 1, ..., N \quad and \quad t = 2, ..., T \tag{1}$$

where α is the parameter of interest with $|\alpha| < 1$. Assume that $\{v_{it}\}$ $(t = 2, ..., T; i = 1, ..., N)$ are $i.i.d$ across time and individuals and independent of η_i and y_{i1} with $E(v_{it}) = 0$, $var(v_{it}) = \sigma_v^2$. We also assume that η_i are $i.i.d$ across individuals with $E(\eta_i) = 0$, $var(\eta_i) = \sigma_\eta^2$. The initial observations satisfy

$$y_{i1} = \frac{\eta_i}{1-\alpha} + w_{i1} \qquad for \quad i = 1, ..., N \tag{2}$$

where w_{i1} is $w_{i1} = \sum_{j=0}^{\infty} \alpha^j v_{i,1-j}$ and independent of η_i.

By first-differencing model (1), we have

$$\Delta y_{it} = \alpha \Delta y_{i,t-1} + \Delta v_{i,t} \quad i = 1, ..., N \quad and \quad t = 3, ..., T \tag{3}$$

The OLS estimator of this model is given by

$$\hat{\alpha}_{dols} = \frac{\sum_{t=3}^{T} \sum_{i=1}^{N} \Delta y_{i,t-1} \Delta y_{i,t}}{\sum_{t=3}^{T} \sum_{i=1}^{N} \Delta y_{i,t-1}^2} = \alpha + \frac{\sum_{t=3}^{T} \sum_{i=1}^{N} \Delta y_{i,t-1} \Delta v_{i,t}}{\sum_{t=3}^{T} \sum_{i=1}^{N} \Delta y_{i,t-1}^2} \tag{4}$$

Using the fact that y_{it} can be expressed as

$$y_{it} = \frac{\eta_i}{1-\alpha} + w_{it} \tag{5}$$

where $w_{it} = \sum_{j=0}^{\infty} \alpha^j v_{i,t-j}$, we obtain

$$\text{plim}_{N\to\infty} \hat{\alpha}_{dols} = \alpha + \text{plim}_{N\to\infty} \frac{\sum_{t=3}^{T} \sum_{i=1}^{N} \Delta y_{i,t-1} \Delta v_{i,t}}{\sum_{t=3}^{T} \sum_{i=1}^{N} \Delta y_{i,t-1}^2} = \alpha + \frac{-\sigma_v^2}{\frac{2\sigma_v^2}{1+\alpha}} \tag{6}$$

$$= \frac{\alpha - 1}{2} \tag{7}$$

Then the modified OLS estimator will be

$$\hat{\alpha}_{mols} = 2\hat{\alpha}_{dols} + 1 \tag{8}$$

Note that in the case of $T = 3$, $\hat{\alpha}_{mols}$ is equal to the bias-corrected LSDV estimator suggested by Bun and Carree (2005). It is also straightforward to show that $\hat{\alpha}_{mols}$ is consistent under large N and fixed T asymptotics:

$$\text{plim}_{N\to\infty} \hat{\alpha}_{mols} = 2\,\text{plim}_{N\to\infty} \hat{\alpha}_{dols} + 1 = \alpha \tag{9}$$

The drawback of $\hat{\alpha}_{mols}$ is that the variance of $\hat{\alpha}_{mols}$ becomes four times that of $\hat{\alpha}_{dols}$.

3 Testing for the absence of individual effects

The difficulty in estimating dynamic panel data models results from the potential presence of individual effects in the model. If individual effects are not present in the model, OLS suffices for the estimation and inference. Therefore, it is worth considering to test for the absence of individual effects. However, so far, only Holtz-Eakin (1988) has considered such test. In this section, we propose a test statistic to test for the absence of individual effects in the model that is based on the Hausman test (Hausmam, 1978) and is easier to implement than the Holtz-Eakin test. The hypothesis to be tested is H_0 : no individual effects in the model against H_1 : individual effects are present in the model. Under H_0, the OLS estimator of (1), $\hat{\alpha}_{ols}$, is consistent and efficient, while under H_1, $\hat{\alpha}_{ols}$ becomes inconsistent. As alternative estimators which become consistent under both H_0 and H_1, we consider $\hat{\alpha}_{bc}$, $\hat{\alpha}_{gmm}$, and $\hat{\alpha}_{mols}$.[3] Let us define

$$\hat{q}_{bc} = \hat{\alpha}_{bc} - \hat{\alpha}_{ols} \tag{10}$$

$$\hat{q}_{gmm} = \hat{\alpha}_{gmm} - \hat{\alpha}_{ols} \tag{11}$$

$$\hat{q}_{mols} = \hat{\alpha}_{mols} - \hat{\alpha}_{ols} \tag{12}$$

Since $\hat{\alpha}_{ols}$ is efficient under H_0, it follows that

$$var(\hat{q}_{bc}) = var(\hat{\alpha}_{bc}) - var(\hat{\alpha}_{ols}) \tag{13}$$

$$var(\hat{q}_{gmm}) = var(\hat{\alpha}_{gmm}) - var(\hat{\alpha}_{ols}) \tag{14}$$

$$var(\hat{q}_{mols}) = var(\hat{\alpha}_{mols}) - var(\hat{\alpha}_{ols}) \tag{15}$$

Then the test statistics are

$$H_{bc} = \frac{\hat{q}_{bc}^2}{var(\hat{q}_{bc})} \sim^a \chi^2(1) \tag{16}$$

$$H_{gmm} = \frac{\hat{q}_{gmm}^2}{var(\hat{q}_{gmm})} \sim^a \chi^2(1) \tag{17}$$

$$H_{mols} = \frac{\hat{q}_{mols}^2}{var(\hat{q}_{mols})} \sim^a \chi^2(1) \tag{18}$$

In the next section, we examine the performances of these test statistics.

[3] $\hat{\alpha}_{bc}$ is Bun and Carree's estimator (Bun and Carree, 2005). $\hat{\alpha}_{gmm}$ is a generalized method of moments (GMM) estimator where the individual effects in the model are removed by the Helmert transformation and all past level instruments, i.e. $(y_{i,1}, ..., y_{i,t-2})$, are used. A more detailed explanation of $\hat{\alpha}_{gmm}$ is provided by Alvarez and Arellano (2003).

4 Monte Carlo experiments

In this section, we conduct Monte Carlo experiments to examine the performance of the MOLS estimator and the test statistic. We consider the following AR(1) model:

$$y_{i,t} = \alpha y_{i,t-1} + \eta_i + v_{it} \tag{19}$$

where $\eta_i \sim iidN(0, \sigma_\eta^2)$, $y_{i,1} \sim iidN(\eta_i/(1-\alpha), \sigma_v^2/(1-\alpha^2))$, and $v_{it} \sim iidN(0, \sigma_v^2)$. The sample sizes we consider are $N = 50, 100, 500$ and $T = 4, 7, 10$. α is set to $\alpha = 0.3, 0.6, 0.9$. The number of replications is 5000 for all cases.

First, we consider the performance of $\hat{\alpha}_{mols}$. Table 1 shows the simulation results for $\hat{\alpha}_{bc}$, $\hat{\alpha}_{gmm}$, and $\hat{\alpha}_{mols}$ in the case of $\sigma_\eta^2 = \sigma_v^2 = 1$. We computed the mean (Mean), the standard deviation (Std. Dev.), and the size (Size) of the Wald test $H_0 : \alpha = \alpha_0$ with a 5% significant level for $\hat{\alpha}_{bc}$, $\hat{\alpha}_{gmm}$, and $\hat{\alpha}_{mols}$. Looking at the results, we find that the biases of both $\hat{\alpha}_{bc}$ and $\hat{\alpha}_{mols}$ are very small for all cases. However, $\hat{\alpha}_{mols}$ is more dispersed than $\hat{\alpha}_{bc}$. With regard to the size property, the size of $\hat{\alpha}_{bc}$ is not stable. In the case of $T = 4$, $\hat{\alpha}_{bc}$ becomes undersized, and in the case of $T = 7, 10$, it becomes oversized. Because the size distortion of the test generally comes from a bias in the estimated coefficient and/or the bias of the estimated variance, we can say that there is a bias in the estimation of the standard error in $\hat{\alpha}_{bc}$ since simulation results show that $\hat{\alpha}_{bc}$ is almost unbiased.[4] In contrast, the sizes of $\hat{\alpha}_{mols}$ are close to the nominal level in all cases. This implies that there is no bias in the estimation of the standard error in $\hat{\alpha}_{mols}$. With regards to the performance of $\hat{\alpha}_{gmm}$, both the bias and the variance are larger than $\hat{\alpha}_{bc}$ and $\hat{\alpha}_{mols}$ in almost all the cases. Furthermore, the size of $\hat{\alpha}_{gmm}$ is very unstable. Therefore, $\hat{\alpha}_{gmm}$ is inferior to $\hat{\alpha}_{bc}$ and $\hat{\alpha}_{mols}$.

Next, we consider the performance of the proposed test statistic. Table 2 shows the simulation results for the test statistics H_{bc}, H_{gmm}, and H_{mols}. We computed the size and the power of the test. In computing the power of the test, we set $\sigma_\eta^2 = 0.5, 1$. With regard to the size of the test, we find that the sizes of H_{bc} and H_{gmm} are very unstable. The sizes of H_{bc} and H_{gmm} differ depending on T, N, and α. Therefore, H_{bc} and H_{gmm} are not useful in practice. In contrast, the size of H_{mols} is very close to the nominal level. These results are related to the size distortion in the Wald statistic discussed above. Although $\hat{\alpha}_{bc}$ is almost unbiased, there exists a bias in the estimation of the variance. This bias leads to the unstable Hausman test statistic H_{bc}. Finally, we consider the power of H_{mols}.[5] H_{mols} has

[4]Simulation results which show the presence of a bias in the estimation of the standard error in $\hat{\alpha}_{bc}$ are available from the author upon request.

[5]We do not consider H_{bc} and H_{gmm} since their sizes are distorted.

reasonable power in the cases of $\alpha = 0.3$ and 0.6 for any N and T. In the case of $\alpha = 0.9$, the power is not high unless both N and T are large. When both N and T are as large, for example when $T = 10$ and $N = 500$, the power of H_{mols} is greater than 0.8.

5 Conclusion

In this paper, we considered the modified OLS estimator which is very simple to implement. The modified OLS estimator is consistent when N is large, and the simulation results showed that it performs very well for a wide range of N, T and α in terms of bias and accuracy of inference. We also considered a test for the absence of individual effects in the model. The simulation results showed that the test based on the modified OLS estimator has reasonable power and size except for strongly persistent data.

References

[1] Alvarez, J. and M. Arellano (2003) " The Time Series and Cross-Section Asymptotics of Dynamic Panel Data Estimators," *Econometrica*, 71, 1121-1159.

[2] Arellano, M. and S.R. Bond (1991) " Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies* , 58, 277-297.

[3] Arellano, M. and O. Bover (1995)" Another Look at the Instrumental Variable Estimation of Error-Component Models," *Journal of Econometrics*, 68, 29-45.

[4] Blundell, R. and S. Bond (1998)" Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 87, 115-43.

[5] Bun, M.J.G. and M.A. Carree (2005)" Bias-Corrected Estimation in Dynamic Panel Data Models," *Journal of Business and Economic Statistics*, 23, 200-210.

[6] Bun, M.J.G. and J.F. Kiviet (2006) " The Effects of Dynamic Feedbacks on LS and MM Estimator Accuracy in Panel Data Models," *Journal of Econometrics*, forthcoming.

[7] Hahn, J. and G. Kuersteiner (2002) " Asymptotically Unbiased Inference for a Dynamic Panel Model with Fixed Effects When Both n and T are Large" *Econometrica*, 70, 1639-1657.

[8] Hausman, J. (1978) " Specification Tests in Econometrics" *Econometrica*, 46, 1251-1271.

[9] Hayakawa, K. (2005) " Small Sample Bias Properties of the System GMM Estimator in Dynamic Panel Data Models," Hi-Stat Discussion Paper, No.82, Hitotsubashi University.

[10] Hayakawa, K. (2006a) " The Asymptotic Properties of the System GMM Estimator in Dynamic Panel Data Models When Both N and T are Large," Hi-Stat Discussion Paper, No.129, Hitotsubashi University.

[11] Hayakawa, K. (2006b) " Efficient GMM Estimation of Dynamic Panel Data Models Where Large Heterogeneity May Be Present," Hi-Stat Discussion Paper, No.130, Hitotsubashi University.

[12] Kiviet, J.F. (1995) " On Bias, Inconsistency, and Efficiency of Various Estimators in Dynamic Panel Data Models," *Journal of Econometrics* , 68, 53-78.

[13] Nickell, S. (1981) " Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49, 1417-1426.

Table 1: Simulation Results

Case			Mean			Std. Dev			Size		
T	N	α	$\hat{\alpha}_{bc}$	$\hat{\alpha}_{gmm}$	$\hat{\alpha}_{mols}$	$\hat{\alpha}_{bc}$	$\hat{\alpha}_{gmm}$	$\hat{\alpha}_{mols}$	$\hat{\alpha}_{bc}$	$\hat{\alpha}_{gmm}$	$\hat{\alpha}_{mols}$
4	50	0.3	0.300	0.260	0.306	0.154	0.258	0.166	0.025	0.062	0.058
4	50	0.6	0.596	0.458	0.604	0.170	0.502	0.183	0.024	0.091	0.057
4	50	0.9	0.894	0.169	0.902	0.178	0.885	0.192	0.019	0.145	0.051
4	100	0.3	0.302	0.280	0.305	0.109	0.183	0.121	0.019	0.055	0.063
4	100	0.6	0.599	0.529	0.603	0.117	0.330	0.127	0.017	0.079	0.053
4	100	0.9	0.897	0.217	0.900	0.128	0.848	0.138	0.014	0.147	0.053
4	500	0.3	0.299	0.296	0.300	0.049	0.081	0.053	0.025	0.053	0.055
4	500	0.6	0.599	0.587	0.599	0.054	0.144	0.057	0.018	0.056	0.052
4	500	0.9	0.902	0.643	0.903	0.057	0.742	0.062	0.011	0.121	0.051
7	50	0.3	0.300	0.250	0.302	0.076	0.107	0.102	0.057	0.077	0.053
7	50	0.6	0.598	0.479	0.601	0.084	0.150	0.112	0.075	0.139	0.048
7	50	0.9	0.896	0.443	0.902	0.090	0.263	0.121	0.065	0.469	0.044
7	100	0.3	0.298	0.273	0.300	0.054	0.078	0.073	0.057	0.066	0.056
7	100	0.6	0.600	0.536	0.602	0.058	0.111	0.079	0.076	0.093	0.049
7	100	0.9	0.898	0.521	0.901	0.064	0.245	0.085	0.071	0.401	0.046
7	500	0.3	0.300	0.295	0.301	0.024	0.036	0.032	0.057	0.053	0.051
7	500	0.6	0.600	0.585	0.600	0.027	0.052	0.035	0.076	0.061	0.047
7	500	0.9	0.899	0.756	0.899	0.029	0.152	0.038	0.074	0.191	0.045
10	50	0.3	0.297	0.254	0.300	0.056	0.074	0.083	0.069	0.097	0.056
10	50	0.6	0.597	0.510	0.600	0.058	0.091	0.091	0.096	0.165	0.054
10	50	0.9	0.895	0.592	0.902	0.062	0.155	0.095	0.121	0.621	0.048
10	100	0.3	0.299	0.277	0.301	0.039	0.052	0.056	0.061	0.065	0.050
10	100	0.6	0.598	0.549	0.599	0.041	0.068	0.063	0.093	0.116	0.049
10	100	0.9	0.897	0.653	0.900	0.044	0.141	0.068	0.126	0.535	0.048
10	500	0.3	0.300	0.296	0.300	0.018	0.024	0.026	0.059	0.050	0.055
10	500	0.6	0.600	0.590	0.600	0.018	0.031	0.029	0.102	0.061	0.051
10	500	0.9	0.900	0.813	0.901	0.020	0.076	0.030	0.123	0.222	0.042

Table 2: Test for the Absence of Individual Effects

Case			Size (5%)			Power ($\sigma_\eta^2 = 0.5$)			Power ($\sigma_\eta^2 = 1$)		
T	N	α	H_{bc}	H_{gmm}	H_{mols}	H_{bc}	H_{gmm}	H_{mols}	H_{bc}	H_{gmm}	H_{mols}
4	50	0.3	0.026	0.070	0.075	0.513	0.467	0.588	0.745	0.567	0.787
4	50	0.6	0.026	0.077	0.067	0.253	0.286	0.349	0.357	0.308	0.440
4	50	0.9	0.019	0.117	0.060	0.044	0.185	0.095	0.051	0.202	0.093
4	100	0.3	0.025	0.060	0.073	0.801	0.651	0.836	0.964	0.743	0.970
4	100	0.6	0.023	0.067	0.066	0.458	0.320	0.559	0.619	0.346	0.708
4	100	0.9	0.017	0.084	0.058	0.053	0.189	0.112	0.057	0.191	0.120
4	500	0.3	0.023	0.057	0.076	1.000	0.997	1.000	1.000	1.000	1.000
4	500	0.6	0.024	0.055	0.070	0.995	0.663	0.997	1.000	0.652	1.000
4	500	0.9	0.011	0.057	0.053	0.186	0.173	0.329	0.210	0.181	0.330
7	50	0.3	0.078	0.087	0.061	0.997	0.973	0.912	1.000	0.994	0.991
7	50	0.6	0.111	0.102	0.057	0.939	0.785	0.654	0.982	0.835	0.793
7	50	0.9	0.090	0.213	0.053	0.258	0.592	0.121	0.268	0.632	0.131
7	100	0.3	0.070	0.073	0.059	1.000	0.999	0.995	1.000	1.000	1.000
7	100	0.6	0.114	0.077	0.060	0.997	0.897	0.910	1.000	0.925	0.975
7	100	0.9	0.093	0.147	0.052	0.379	0.545	0.189	0.410	0.578	0.204
7	500	0.3	0.078	0.050	0.064	1.000	1.000	1.000	1.000	1.000	1.000
7	500	0.6	0.113	0.056	0.055	1.000	1.000	1.000	1.000	1.000	1.000
7	500	0.9	0.093	0.065	0.055	0.917	0.388	0.647	0.929	0.412	0.680
10	50	0.3	0.107	0.121	0.060	1.000	1.000	0.980	1.000	1.000	0.999
10	50	0.6	0.183	0.132	0.059	0.999	0.990	0.846	1.000	0.997	0.931
10	50	0.9	0.151	0.297	0.054	0.492	0.824	0.153	0.523	0.861	0.176
10	100	0.3	0.104	0.086	0.059	1.000	1.000	1.000	1.000	1.000	1.000
10	100	0.6	0.170	0.100	0.056	1.000	0.999	0.984	1.000	1.000	0.999
10	100	0.9	0.165	0.191	0.047	0.718	0.785	0.273	0.767	0.814	0.289
10	500	0.3	0.105	0.054	0.060	1.000	1.000	1.000	1.000	1.000	1.000
10	500	0.6	0.178	0.056	0.059	1.000	1.000	1.000	1.000	1.000	1.000
10	500	0.9	0.166	0.083	0.050	0.999	0.676	0.831	1.000	0.706	0.865