Small Sample Bias Properties of the System GMM Estimator in Dynamic Panel Data Models

Kazuhiko Hayakawa*

Graduate School of Economics, Hitotsubashi University

and

JSPS Research Fellow

June 29, 2006

Abstract

By deriving the finite sample biases, this paper shows analytically why the system GMM estimator in dynamic panel data models is less biased than the first differencing or the level estimators even though the former uses more instruments.

Keywords: Dynamic panel data model; GMM estimator; Finite sample bias.

JEL Classification: C23

*Corresponding author: Kazuhiko Hayakawa. 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan. TEL:+81-42-580-8192, FAX:+81-42-580-8195.

1 Introduction

In the estimation of dynamic panel data models, the system GMM estimator by Blundell and Bond (1998) is widely used in empirical analyses. Since the system GMM estimator uses more instruments than the first differencing and the level estimators, one might suspect that the system GMM estimator is more biased than the first differencing and the level estimators. However some simulation results do not pose such a problem; rather the system GMM estimator is less biased than the first differencing and the level GMM estimators. The purpose of this paper is to show theoretically why the system GMM estimator, despite using more instruments, is less biased than the first differencing and level GMM estimators by deriving the second order bias of the system GMM estimator.[1]

The remainder of the paper is organized as follows. In Section 2, we define the model and the estimators used in this paper. Section 3 provides the main results concerning the small sample bias of the GMM estimators. Based on the results of Section 3, we compare the estimators numerically in Section 4. Finally, Section 5 concludes.

2 The model and estimators

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

$$y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it} \quad i = 1, ..., N; \quad t = 2, 3, 4 \tag{1}$$

where α is the parameter of interest with $|\alpha| < 1$ and v_{it} has mean zero given $\eta_i, y_{i1}, ..., y_{i,t-1}$. Let $u_{it} = \eta_i + v_{it}$. We impose the following assumptions:

Assumption 1. $\{v_{it}\}$ $(t = 2, 3, 4; i = 1, ..., N)$ *are i.i.d across time and individuals and independent of* η_i *and* y_{i1} *with* $E(v_{it}) = 0$, $var(v_{it}) = \sigma_v^2$ *and finite moments up to third order.*

Assumption 2. η_i *are i.i.d across individuals with* $E(\eta_i) = 0$, $var(\eta_i) = \sigma_\eta^2$, *and finite third order moments*

Assumption 3. *The initial observations satisfy*

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

[1] Although Bun and Kiviet (2006) derived the second order bias of the system GMM estiamtor, their purpose is different from this paper. Their purpose is to examine how the order of magnitude of bias changes when we use different set of instruments, while this paper aims to examine the structure of the bias of the system GMM estimator.

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

Under these assumptions, we consider three types of GMM estimators. These include the first differencing GMM estimator (Arellano and Bond, 1991), the level and system GMM estimator (Arellano and Bover, 1995), abbreviated as GMM(DIF), GMM(LEV) and GMM(SYS) estimators, respectively. These GMM estimators are derived from the following moment conditions;

$$E[Z'_{d,i}\Delta u_i] = 0 \qquad E(Z'_{l,i}u_i) = 0 \qquad E(Z'_{s,i}p_i) = 0 \tag{3}$$

where

$$Z_{d,i} = \begin{bmatrix} y_{i1} & 0 & 0 \\ 0 & y_{i1} & y_{i2} \end{bmatrix} \quad Z_{l,i} = \begin{bmatrix} \Delta y_{i2} & 0 \\ 0 & \Delta y_{i3} \end{bmatrix} \quad Z_{s,i} = \begin{bmatrix} Z_{d,i} & 0 \\ 0 & Z_{l,i} \end{bmatrix} \tag{4}$$

and

$$\Delta u_i = \begin{bmatrix} \Delta u_{i3} \\ \Delta u_{i4} \end{bmatrix} \quad u_i = \begin{bmatrix} u_{i3} \\ u_{i4} \end{bmatrix} \quad p_i = \begin{bmatrix} \Delta u_i \\ u_i \end{bmatrix} \tag{5}$$

Let $\Delta y'_i = (\Delta y_{i3}, \Delta y_{i4})'$, $\Delta y'_{i,-1} = (\Delta y_{i2}, \Delta y_{i3})'$, $y'_i = (y_{i3}, y_{i4})'$, $y'_{i,-1} = (y_{i2}, y_{i3})'$, $q'_i = (\Delta y'_i \; y'_i)'$, $q'_{i,-1} = (\Delta y'_{i,-1} \; y'_{i,-1})'$, and Δy, Δy_{-1} y, y_{-1}, q, and q_{-1} are stacked across individuals. Then the one-step GMM estimators for α are

$$\widehat{\alpha}_{dif} = \frac{\Delta y'_{-1} Z_d (Z'_d Z_d)^{-1} Z'_d \Delta y}{\Delta y'_{-1} Z_d (Z'_d Z_d)^{-1} Z'_d \Delta y_{-1}} \tag{6}$$

$$\widehat{\alpha}_{lev} = \frac{y'_{-1} Z_l (Z'_l Z_l)^{-1} Z'_l y}{y'_{-1} Z_l (Z'_l Z_l)^{-1} Z'_l y_{-1}} \tag{7}$$

$$\widehat{\alpha}_{sys} = \frac{q'_{-1} Z_s (Z'_s Z_s)^{-1} Z'_s q}{q'_{-1} Z_s (Z'_s Z_s)^{-1} Z'_s q_{-1}} = \hat{\gamma}\hat{\alpha}_{dif} + (1-\hat{\gamma})\hat{\alpha}_{lev} \tag{8}$$

where $Z'_d = (Z'_{d,1}, \cdots, Z'_{d,N})$, $Z'_l = (Z'_{l,1}, \cdots, Z'_{l,N})$, $Z'_s = (Z'_{s,1}, \cdots, Z'_{s,N})$, and

$$\hat{\gamma} = \frac{\Delta y'_{-1} Z_d (Z'_d Z_d)^{-1} Z'_d \Delta y_{-1}}{\Delta y'_{-1} Z_d (Z'_d Z_d)^{-1} Z'_d \Delta y_{-1} + y'_{-1} Z_l (Z'_l Z_l)^{-1} Z'_l y_{-1}} \tag{9}$$

3 Small sample biases of GMM estimators

In this section, we provide analytical forms of the finite sample biases of order N^{-1} for the GMM(DIF), GMM(LEV) and GMM(SYS) estimators.

First we define π_d and π_l as

$$\pi_d = E(Z'_{d,i} Z_{d,i})^{-1} E(Z'_{d,i} \Delta y_{i,-1})$$

$$
\begin{aligned}
&= \left[\begin{array}{ccc} \frac{-\sigma_v^2}{(1+\alpha)(C+D)} & \frac{\sigma_v^2(1-\alpha)C}{(1+\alpha)F} & \frac{\sigma_v^2(\alpha-1)}{(\alpha+1)F}[C+(\alpha+1)D] \end{array}\right]' = \left[\begin{array}{ccc} \pi_{d,1} & \pi_{d,2} & \pi_{d,3} \end{array}\right]' \\
\pi_l &= E(Z_{l,i}'Z_{l,i})^{-1}E(Z_{l,i}'y_{i,-1}) = \left[\begin{array}{cc} \frac{1}{2} & \frac{1}{2} \end{array}\right]' = \left[\begin{array}{cc} \pi_{l,1} & \pi_{l,2} \end{array}\right]'
\end{aligned}
$$

where

$$
\begin{aligned}
C &= \left(\frac{1}{1-\alpha}\right)^2 \sigma_\eta^2 \qquad D = \frac{1}{1-\alpha^2}\sigma_v^2 \\
F &= \left[\frac{1}{1+\alpha}\sigma_v^2\right]\left[2\left(\frac{1}{1-\alpha}\right)^2\sigma_\eta^2 + \left(\frac{1}{1-\alpha}\right)\sigma_v^2\right]
\end{aligned}
$$

Also, we define ϕ_d, ϕ_l and γ as

$$
\begin{aligned}
\phi_d &= E(\Delta y_{i,-1}'Z_{d,i})E(Z_{d,i}'Z_{d,i})^{-1}E(Z_{d,i}'\Delta y_{i,-1}) = \pi_d' E(Z_{d,i}'Z_{d,i})\pi_d \\
&= (\pi_{d,1}^2 + \pi_{d,2}^2 + \pi_{d,3}^2)(C+D) + 2\pi_{d,2}\pi_{d,3}(C+\alpha D) \\
\phi_l &= E(y_{i,-1}'Z_{l,i})E(Z_{l,i}'Z_{l,i})^{-1}E(Z_{l,i}'y_{i,-1}) = \pi_l' E(Z_{l,i}'Z_{l,i})\pi_l = \frac{\sigma_v^2}{1+\alpha} \\
\gamma &= \frac{\phi_d}{\phi_d + \phi_l}
\end{aligned}
$$

Next, we provide the formulas for the small sample biases of $\hat{\alpha}_{dif}$, $\hat{\alpha}_{lev}$ and $\hat{\alpha}_{sys}$.

Theorem 1. *The second order biases of $\hat{\alpha}_{dif}$, $\hat{\alpha}_{lev}$, and $\hat{\alpha}_{sys}$ are given by*

$$
\begin{aligned}
N \cdot Bias(\hat{\alpha}_{dif}) &= B_1^{dif} + B_2^{dif} + B_3^{dif} + B_4^{dif} \qquad (10) \\
N \cdot Bias(\hat{\alpha}_{lev}) &= B_1^{lev} + B_2^{lev} + B_3^{lev} + B_4^{lev} \qquad (11) \\
N \cdot Bias(\hat{\alpha}_{sys}) &= \gamma(B_1^{dif} + B_2^{dif}) + (1-\gamma)(B_1^{lev} + B_2^{lev}) \\
&\quad + \gamma^2(B_3^{dif} + B_4^{dif}) + (1-\gamma)^2(B_3^{lev} + B_4^{lev}) \\
&\quad - \frac{2}{(\phi_d+\phi_l)^2}\Psi_3 + \frac{1}{(\phi_d+\phi_l)^2}\Psi_4 \qquad (12)
\end{aligned}
$$

where

$$
\begin{aligned}
B_1^{dif} &= 0 \qquad B_2^{dif} = \frac{-\sigma_v^2}{\phi_d}\left[1 + \frac{2(C+D)^2}{F} - \frac{2(C+\alpha D)^2}{F}\right] \\
B_3^{dif} &= \frac{2\sigma_v^2}{\phi_d^2}\left[\left\{\pi_{d1}^2 + \pi_{d2}^2 + \pi_{d3}^2 + (\alpha-2)\pi_{d1}\pi_{d2}\right\}(C+D)\right. \\
&\quad \left. -\pi_{d1}\pi_{d3}\left\{(2-\alpha)C - \alpha(2\alpha-3)D\right\} + 2\pi_{d2}\pi_{d3}(C+\alpha D)\right] \\
B_4^{dif} &= -\frac{-2\sigma_v^2}{\phi_d^2}\left[\pi_{d1}\pi_{d3}^2(C+\alpha D) + \pi_{d1}\pi_{d2}\pi_{d3}(C+D)\right] \\
B_1^{lev} &= 0 \qquad B_2^{lev} = \frac{2\sigma_\eta^2}{\phi_l(1-\alpha)} \\
B_3^{lev} &= \frac{-\sigma_v^2}{\phi_l^2}\left[\frac{1}{1-\alpha}\sigma_\eta^2 + \frac{2\alpha-1}{2(1+\alpha)}\sigma_v^2\right] \qquad B_4^{lev} = \frac{\alpha-1}{4\phi_l^2(\alpha+1)}\sigma_v^4 \\
\Phi_3 &= \pi_d' \begin{bmatrix} \frac{-3\alpha+1}{(1-\alpha)^2(1+\alpha)}\sigma_\eta^2\sigma_v^2 - \frac{2\alpha-1}{1-\alpha^2}\sigma_v^4 & \frac{-2\alpha^2+3\alpha+1}{(1-\alpha)^2(1+\alpha)}\sigma_v^2\sigma_\eta^2 - \frac{\alpha(2\alpha^2-4\alpha+1)}{1-\alpha^2}\sigma_v^4 \\ \frac{-1}{1+\alpha}\sigma_\eta^2\sigma_v^2 - \frac{1}{1+\alpha}\sigma_v^4 & \frac{-3\alpha+1}{(1-\alpha)^2(1+\alpha)}\sigma_v^2\sigma_\eta^2 - \frac{\alpha(2\alpha-1)}{1-\alpha^2}\sigma_v^4 \\ \frac{-1}{1+\alpha}\sigma_\eta^2\sigma_v^2 + \frac{1}{1+\alpha}\sigma_v^4 & \frac{-3\alpha+1}{(1-\alpha)^2(1+\alpha)}\sigma_v^2\sigma_\eta^2 - \frac{2\alpha-1}{1-\alpha^2}\sigma_v^4 \end{bmatrix} \pi_l
\end{aligned}
$$

$$\Psi_4 = \pi_d' \begin{bmatrix} \pi_{l,1}\frac{2}{\alpha+1}\sigma_v^4 - \pi_{d,1}\frac{2}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 & \pi_{l,2}\frac{2\alpha(\alpha-2)}{1+\alpha}\sigma_v^4 - \pi_{d,1}\frac{2\alpha}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 \\ -\pi_{d,2}\frac{2}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 & \pi_{l,2}\frac{2\alpha}{1+\alpha}\sigma_v^4 - \pi_{d,2}\frac{2\alpha}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 - \pi_{d,3}\frac{1}{1-\alpha}\sigma_v^2\sigma_\eta^2 \\ \pi_{d,3}\frac{2\alpha}{1-\alpha^2}\sigma_\eta^2\sigma_v^2 & \pi_{l,2}\frac{2}{1+\alpha}\sigma_v^4 - \pi_{d,2}\frac{1}{1-\alpha}\sigma_v^2\sigma_\eta^2 - \pi_{d,3}\frac{2}{1-\alpha^2}\sigma_v^2\sigma_\eta^2 \end{bmatrix} \pi_l$$

Proof. Using Theorem 1 in Hahn, Hausman and Kuersteiner (2001), this result is readily obtained after very lengthy algebra. Detailed proof is available from the author upon request.

Remark We find that the bias of the system GMM estimator is composed of two elements. The first is the weighted sum of the biases of the first differencing and the level estimators (the first two rows of (12)). The second element of the bias results from using the first differencing and the level estimators jointly (the last row in (12)). Therefore, if the biases of the first differencing and the level estimators are in opposite directions, they will cancel each other out and the system estimator will have small bias. In the next section we will show that this is indeed the case.

4 Numerical analysis

Since the biases of all estimators are characterized by N, α and σ_η^2/σ_v^2, we calculate the theoretical values of the biases for the cases $\sigma_\eta^2/\sigma_v^2 = 0.25, 1, 4$ with $\alpha = 0.1, ..., 0.9$ and $N = 50$. Before we begin to analyze the direction and magnitude of the biases, we confirm how well the second order biases explain the actual biases by comparing theoretical values with simulation values. From tables 1, and 2, which describe the theoretical and simulation values of the biases of GMM(DIF), GMM(LEV) and GMM(SYS) estimators.[2] we find that the theoretical and simulation values are close when $\alpha \leq 0.5$. Hence, in the following, we focus on the case with $\alpha \leq 0.5$.

The first concern is the direction of the biases of the GMM(DIF) and GMM(LEV) estimators. Looking at the tables, we find that the GMM(DIF) estimator has a downward bias, while the GMM(LEV) estimator has an upward bias, and both biases cancel each other out in the GMM(SYS) estimator. The second concern is the value of γ and the magnitude of the biases. Even though they work in opposite directions and thus at least partly cancel each other out, the bias of the GMM(SYS) estimator will not be small in absolute value if γ takes values near zero or one, or if

[2] v_{it}, η_i and w_{i1} are independently generated by $v_{it} \sim iidN(0,1)$, $\eta_i \sim iidN(0,\sigma_\eta^2)$ and $w_{i1} \sim iidN(0, 1/(1-\alpha^2))$ where $\sigma_\eta^2 = 1, 4$. The magnitudes of relative biases are computed by $Bias(\hat{\alpha})/\alpha \times 100(\%)$

the bias of either estimator is much larger than the other in absolute value. Therefore, we need to think about γ and the magnitude of the biases. Table 4 shows the values of γ for $\alpha = 0.1, ..., 0.9$. We find that γ moves between about 0.25 and 0.5 in the region of $\alpha \leq 0.5$. We also find that although the bias of the GMM(DIF) estimator is about -20 $-$ -30% in the region of $\alpha \leq 0.5$, the bias of the GMM(LEV) estimator decreases from about 26% to 5% as α increases. Hence, the difference of the magnitude of the biases of the two estimators in absolute value gets larger as α increases. However, Table 4 shows that $1 - \gamma$, the weight on the GMM(LEV) estimator, increases as α increases, that is, the weight γ is adjusting the difference of the magnitude of the biases to be almost the same in absolute value. The column labeled "weighted sum" in Table 3 reports the theoretical values of the first element of the bias of the system estimator. In the region of $\alpha \leq 0.5$, it takes negative stable values around -0.4. The theoretical values of the second element in the bias of the GMM(SYS) estimator are given in the column labeled "correlation" in Table 3. We find that the second element of the bias of the system estimator takes positive values and gets smaller as α grows. Hence, we find that in this case, too, the two elements of the bias of the GMM(SYS) estimator partly cancel each other out, and this is the reason why the small sample bias is almost zero around $\alpha = 0.3$ or 0.4. Thus, we find that there are two reasons why the GMM(SYS) estimator is less biased. The first is that the biases of the GMM(DIF) and the GMM(LEV) estimators are in opposite directions, and the second is that the weight γ is adjusting the difference of the magnitude of both biases.

The above results pertain to the case when $\sigma_\eta^2/\sigma_v^2 = 1$. Now we turn to the case when $\sigma_\eta^2/\sigma_v^2 = 4$. The theoretical and simulation values of the biases are given in Table 2. The approximation of the small sample bias is a little worse than that of $\sigma_\eta^2/\sigma_v^2 = 1$. However, we find that there are sizable biases for any value of α for all estimators. Therefore, using the GMM(SYS) estimator in this case is problematic.

5 Conclusion

In this paper, we considered the small sample bias properties of GMM estimators in dynamic panel data models. We provided theoretical evidence why the system GMM estimator has smaller bias. We found that the bias of the system GMM estimator is a weighted sum of the biases in opposite directions of the first differencing and the level GMM estimators. In addition, we found that the role of the weight is also important since it adjusts the difference of the magnitudes of the biases. The numerical analysis showed that the fact that σ_η^2 and σ_v^2 are of almost the same value is an important reason why the system estimator has small bias. In the case when $\sigma_\eta^2/\sigma_v^2 = 4$, the biases of all the GMM estimator are sizable.

Acknowledgment

I am deeply grateful to Taku Yamamoto for his helpful and constructive comments. I also would like to thank Katsuto Tanaka, Satoru Kanoh and Yukinobu Kitamura for helpful suggestions. Financial support from the JSPS Fellowship is also gratefully acknowledge. Of course, all remaining errors are mine.

References

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

[2] Arellano, M., Bover, O., 1995. Another look at the instrumental variable estimator of error-components models. Journal of Econometrics 68, 29-52.

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

[4] Bun, M.J.G., Kiviet, J.F., 2006. The effects of dynamic feedbacks on LS and MM estimator accuracy in panel data models. Journal of Econometrics, forthcoming.

[5] Hahn, J., Hausman, J., Kuersteiner, G. 2001. Bias corrected instrumental variables estimation for dynamic panel models with fixed effects. MIT Department of Economics Working Paper No. 01-24.

Table 1: Theoretical and simulation values of relative biases (%) when $\sigma_\eta^2/\sigma_v^2 = 1$

	GMM(DIF)		GMM(LEV)		GMM(SYS)	
α	theory	simulation	theory	simulation	theory	simulation
0.1	-26.48	-27.99	25.85	26.01	5.67	5.44
0.2	-20.02	-19.79	13.20	13.42	2.35	2.26
0.3	-20.11	-20.12	8.88	8.49	0.87	0.71
0.4	-23.04	-23.05	6.65	6.32	-0.37	0.07
0.5	-29.11	-28.00	5.25	5.19	-1.77	-1.25
0.6	-40.75	-38.20	4.27	6.81	-3.53	-0.53
0.7	-65.59	-58.64	3.52	9.29	-5.76	1.47
0.8	-134.23	-76.93	2.92	9.85	-8.43	2.56
0.9	-488.87	-93.41	2.43	10.05	-11.32	5.69

Table 2: Theoretical and simulation values of relative biases (%) when $\sigma_\eta^2/\sigma_v^2 = 4$

	GMM(DIF)		GMM(LEV)		GMM(SYS)	
α	theory	simulation	theory	simulation	theory	simulation
0.1	-39.65	-40.28	91.85	96.55	64.52	62.48
0.2	-29.43	-31.15	49.20	51.77	37.53	35.50
0.3	-29.29	-28.01	34.88	40.58	28.19	28.37
0.4	-33.46	-34.14	27.65	39.65	22.87	21.20
0.5	-42.40	-44.04	23.25	36.82	18.77	20.72
0.6	-59.85	-59.04	20.27	27.81	14.92	19.88
0.7	-97.65	-71.66	18.09	28.05	10.94	20.34
0.8	-203.64	-86.89	16.42	21.01	6.71	16.87
0.9	-759.58	-101.12	15.09	10.71	2.32	9.57

Table 3: Composition of the relative bias of the GMM(SYS) estimator when $\sigma_\eta^2/\sigma_v^2 = 1$

α	Weighted sum	Correlation	$Bias(\hat{\alpha}^{sys})$
0.1	-4.571	10.243	5.672
0.2	-3.665	6.017	2.352
0.3	-3.673	4.543	0.869
0.4	-4.004	3.635	-0.369
0.5	-4.571	2.802	-1.769
0.6	-5.366	1.834	-3.531
0.7	-6.384	0.623	-5.761
0.8	-7.586	-0.844	-8.430
0.9	-8.856	-2.462	-11.318

Table 4: Theoretical values of γ

	$\sigma_\eta^2/\sigma_v^2 = 1$		$\sigma_\eta^2/\sigma_v^2 = 4$		$\sigma_\eta^2/\sigma_v^2 = 0.25$	
α	γ	$1-\gamma$	γ	$1-\gamma$	γ	$1-\gamma$
0.1	0.491	0.509	0.378	0.622	0.586	0.414
0.2	0.437	0.563	0.323	0.677	0.545	0.455
0.3	0.378	0.622	0.267	0.733	0.497	0.503
0.4	0.314	0.686	0.211	0.789	0.441	0.559
0.5	0.245	0.755	0.156	0.844	0.374	0.626
0.6	0.176	0.824	0.106	0.894	0.297	0.703
0.7	0.110	0.890	0.063	0.937	0.209	0.791
0.8	0.053	0.947	0.029	0.971	0.117	0.883
0.9	0.014	0.986	0.007	0.993	0.037	0.963