Far East J. Theo. Stat. **20**(2) (2006), 179-196

COMPARISON OF GMM WITH SECOND-ORDER LEAST SQUARES ESTIMATION IN NONLINEAR MODELS

TARANEH ABARIN and LIQUN WANG
Department of Statistics
University of Manitoba
Winnipeg, Manitoba, Canada R3T 2N2
e-mail: liqun_wang@umanitoba.ca

Abstract

Generalized method of moments (GMM) is an estimation technique which estimates unknown parameters by matching theoretical moments with sample moments. It may provide a poor approximation to the finite sample distribution of the estimator. Moreover, increasing the number of moment conditions requires substantial increase of the sample size. Second-order least squares (SLS) estimation is an extension of the ordinary least squares method by adding to the criterion function the distance of the squared response variable to its second conditional moment. It is shown in this paper that the SLS is asymptotically more efficient than the GMM when both use the same moment conditions. Moreover, Monte Carlo simulation studies show that SLS performs better than the GMM estimators using three or four moment conditions.

1. Introduction

It has been more than two decades since econometricians have

2000 Mathematics Subject Classification: Primary 62F10, 62F12; Secondary 62J02.

Keywords and phrases: asymptotic efficiency, consistency, identifiability, least squares method, method of moments, minimum distance estimator, nonlinear regression.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

Received June 14, 2006

suggested using a generalized version of Pearson's [6] method of moment approach, commonly known as the *generalized method of moment* (GMM) estimation. It is an estimation technique which estimates unknown parameters by matching theoretical moments with the sample moments. Large sample properties of GMM estimators have been developed by Hansen [2]. In recent years, there has been growing literature on finite sample behavior of GMM estimators, see, e.g., Windmeijer [11] and references therein. Although GMM estimators are consistent and asymptotically normally distributed under general regularity conditions, it has long been recognized that this asymptotic distribution may provide a poor approximation to the finite sample distribution of the estimators.

Identification is another issue in GMM method (e.g., Stock and Wright [7] and Wright [12]). In particular, the number of moment conditions needs to be equal to or greater than the number of parameters, in order for them to be identified. This restriction makes problems when parametric dimension increases. It seems that adding over-identifying restrictions (moment conditions) will increase precision, however this is not always the case (Anderson and Sorenson [1]). More recently, for a linear model with heteroscedasticity, Koenker and Machado [4] showed that an effective sample size can be given by n/q_n^3, where q_n is the number of moment conditions used. This means that very large sample size is required to justify conventional asymptotic inference. See also Huber [3, p. 164].

Recently, Wang [8, 9] proposed a second-order least squares (SLS) estimator which is based on the first two conditional moments of the response variable given the predictor variables. This estimator extends the ordinary least squares estimation by including in the criterion function the distance of the squared response variable to its second conditional moment. Wang and Leblanc [10] compared the SLS estimator with the ordinary least squares (OLS) estimator in nonlinear models. They showed that SLSE is asymptotically more efficient than the OLSE when the third moment of the random error is nonzero. An interesting question that arises naturally is how does SLS compare with GMM estimators? This question is partially addressed in the current paper.

In Section 2, we introduce the SLS and its asymptotic properties. In Section 3, we compare the asymptotic covariance matrix of the (efficient) SLS and GMM when both use the same number of moment conditions. Since a theoretical comparison of these two methods is extremely difficult when SLS and GMM use different number of moment conditions, we compare these two estimators through Monte Carlo simulation studies, which consist of Section 4. Finally, conclusions are given in Section 5.

2. Second-order Least Squares Estimation

Consider a general regression model

$$Y = g(X; \theta) + \varepsilon, \tag{1}$$

where $Y \in \mathbb{R}$ is the response variable, $X \in \mathbb{R}^k$ is the predictor variable, $\theta \in \mathbb{R}^p$ is the unknown regression parameter and ε is the random error satisfying $E(\varepsilon \mid X) = 0$ and $E(\varepsilon^2 \mid X) = \sigma^2$. Under model (1) the first two conditional moments of Y given X are respectively, $E_\gamma(Y \mid X) = g(X; \theta)$ and $E_\gamma(Y^2 \mid X) = g^2(X; \theta) + \sigma^2$, where $\gamma = (\theta', \sigma^2)'$. Suppose $(Y_i, X_i')'$, $i = 1, 2, \ldots, n$ is an i.i.d. random sample. Following Wang and Leblanc [10], the *second-order least squares estimator* $\hat{\gamma}_{SLS}$ for γ is defined as the measurable function that minimizes

$$Q_n(\gamma) = \sum_{i=1}^{n} \rho_i'(\gamma) W_i \rho_i(\gamma), \tag{2}$$

where $\rho_i(\gamma) = (Y_i - g(X_i; \theta), Y_i^2 - g^2(X_i; \theta) - \sigma^2)'$ and $W_i = W(X_i)$ is a 2×2 nonnegative definite matrix which may depend on X_i. It is easy to see that this estimator is an extension of the ordinary least squares estimator by adding the distance of the squared response variable to its second conditional moment into the criterion function.

Wang and Leblanc [10] proved that, under some regularity conditions, the SLSE is consistent and has an asymptotic normal distribution with the asymptotic covariance matrix given by $A^{-1}BA^{-1}$, where

$$A = E\left[\frac{\partial \rho_i'(\gamma)}{\partial \gamma} W_i \frac{\partial \rho_i(\gamma)}{\partial \gamma'}\right], \quad B = E\left[\frac{\partial \rho_i'(\gamma)}{\partial \gamma} W_i \rho_i(\gamma)\rho_i'(\gamma) W_i \frac{\partial \rho_i(\gamma)}{\partial \gamma'}\right]$$

and

$$\frac{\partial \rho_i'(\gamma)}{\partial \gamma} = -\begin{pmatrix} \frac{\partial g(X_i;\, \theta)}{\partial \theta} & 2g(X_i;\, \theta)\frac{\partial g(X_i;\, \theta)}{\partial \theta} \\ 0 & 1 \end{pmatrix}.$$

They pointed out that the best choice for the weighting matrix is $W_i = U_i^{-1}$, where $U_i = E(\rho_i(\gamma)\rho_i'(\gamma)|X_i)$, which gives the smallest variance-covariance matrix

$$E^{-1}\left[\frac{\partial \rho_i'(\gamma)}{\partial \gamma} U_i^{-1} \frac{\partial \rho_i(\gamma)}{\partial \gamma'}\right].$$

In the following we provide a formal proof of this fact.

Theorem 1. *Denote* $C = \frac{\partial \rho_i(\gamma)}{\partial \gamma'}$ *and* $D = \rho_i(\gamma)\rho_i'(\gamma)$. *Then the asymptotic covariance matrix of the most efficient SLSE is* $E^{-1}(C'U^{-1}C)$, *where* $U = E(D \mid X_i)$.

Proof. First, it is easy to see that $A = E(C'WC)$ and $B = E(C'WDWC)$ $= E(C'WUWC)$, because C and W do not depend on Y_i. Further, let $\alpha = E^{-1}(C'WUWC)E(C'WC)$. Then we have

$$\begin{aligned} & E(C - UWC\alpha)' U^{-1}(C - UWC\alpha) \\ &= E(C'U^{-1}C) - E(C'WC)\alpha - \alpha' E(C'WC) + \alpha' E(C'WUWC)\alpha \\ &= E(C'U^{-1}C) - E(C'WC)E^{-1}(C'WUWC)E(C'WC) \\ &= E(C'U^{-1}C) - AB^{-1}A \end{aligned}$$

which is nonnegative definite. It follows that $E^{-1}(C'U^{-1}C) \le A^{-1}BA^{-1}$, and equality holds if $W = U^{-1}$ in both A and B.

In the rest of this paper, SLSE always refers to the most efficient second-order least squares estimator using the weight $W = U^{-1}$.

3. Comparison with GMM Estimation

In this section, we compare the SLS with GMM estimator when both use the same set of moment conditions. Given the i.i.d. random sample $(Y_i, X_i')'$, $i = 1, 2, \ldots, n,$ the *GMM estimator* using the first two conditional moments is defined as the measurable function which minimizes

$$G_n(\gamma) = \left(\sum_{i=1}^{n} \rho_i(\gamma)\right)' W_n \left(\sum_{i=1}^{n} \rho_i(\gamma)\right), \tag{3}$$

where $\rho_i(\gamma)$ is defined in (2) and W_n is a nonnegative definite weighting matrix. It can be shown (e.g., Mátyás [5]) that under some regularity conditions, the GMM estimator has an asymptotic normal distribution and the asymptotic covariance matrix of the efficient GMM is given by

$$\left[E\left(\frac{\partial \rho_i'(\gamma)}{\partial \gamma}\right) V^{-1} E\left(\frac{\partial \rho_i(\gamma)}{\partial \gamma'}\right)\right]^{-1},$$

where $V = E[\rho_i(\gamma)\rho_i'(\gamma)] = E(U)$ is the optimum weighting matrix. The next theorem compares the asymptotic covariance matrices of the SLS and GMM estimators.

Theorem 2. *The SLSE is asymptotically more efficient than the GMM estimator using the first two moment conditions, i.e.,* $E^{-1}(C'U^{-1}C) \leq [E(C')V^{-1}E(C)]^{-1}$.

Proof. The proof is similar to that of Theorem 1. Let $\alpha = V^{-1}E(C)$. Then the result follows from

$$\begin{aligned} & E(C - U\alpha)' U^{-1}(C - U\alpha) \\ &= E(C'U^{-1}C) - E(C')\alpha - \alpha' E(C) + \alpha' E(U)\alpha \\ &= E(C'U^{-1}C) - E(C')V^{-1}E(C) \geq 0. \end{aligned}$$

If $\hat{\theta}$ and $\hat{\sigma}^2$ denote the estimators of the regression and variance parameters, respectively, then the above theorem implies that $V(\hat{\theta}_{GMM}) \geq V(\hat{\theta}_{SLS})$ and $V(\hat{\sigma}^2_{GMM}) \geq V(\hat{\sigma}^2_{SLS})$ asymptotically. Given these theoretical comparison results, an interesting question is that does GMM perform better than SLS if using more than two moment conditions? This question is examined in the next section.

4. Monte Carlo Simulation Studies

In order to verify the finite sample behavior of the second-order least squares (SLS) estimation approach and the generalized method of moment (GMM) estimation, several simulation scenarios are considered.

4.1. Design of the studies

We consider the following exponential, logistic and linear exponential models, each with two or three parameters.

1. Exponential model with two parameters: $Y = 10e^{\theta X} + \varepsilon$, where $X \sim U(0.1, 10)$, $\varepsilon \sim N(0, \sigma^2)$, and true parameter values are $\theta = -0.5$ and $\sigma^2 = 1$.

2. Exponential model with three parameters: $Y = \theta_1 e^{\theta_2 X} + \varepsilon$, where $X \sim U(0.1, 10)$, $\varepsilon \sim N(0, \sigma^2)$, and true parameters are $\theta_1 = 10$, $\theta_2 = -0.5$ and $\sigma^2 = 1$.

3. Logistic model with two parameters: $Y = \dfrac{20}{1 + \exp[-(X - \theta)/34]} + \varepsilon$, where $X \sim U(20, 100)$, $\varepsilon \sim N(0, \sigma^2)$, and true parameters are $\theta = 50$ and $\sigma^2 = 1$.

4. Logistic model with three parameters: $Y = \dfrac{20}{1 + \exp[-(X - \theta_1)/\theta_2]} + \varepsilon$, where $X \sim U(20, 80)$, $\varepsilon \sim N(0, \sigma^2)$, and true parameters are $\theta_1 = 50$, $\theta_2 = 34$ and $\sigma^2 = 1$.

5. Linear-exponential model with two parameters: $Y = 5e^{\theta_1 X} + 10e^{\theta_2 X} + \varepsilon$, where $X \sim U(0.1, 10)$, $\varepsilon \sim N(0, 1)$, and true parameters are $\theta_1 = -3$ and $\theta_2 = -1$.

6. Linear-exponential model with three parameters: $Y = \theta_3 e^{\theta_1 X} + 10e^{\theta_2 X} + \varepsilon$, where $X \sim U(0.1, 10)$, $\varepsilon \sim N(0, 1)$, and true parameters are $\theta_1 = -3$, $\theta_2 = -1$ and $\theta_3 = 5$.

In each model we compare SLS with two versions of GMM estimator, one (GMM3) using the first three and another (GMM4) using the first four moment conditions. Both SLS and GMM estimators are computed in two steps. In the first step, identity weighting matrix is used to obtain initial parameter estimates. In the second step, first the optimal weight for the SLS is calculated according formula (7) in Wang and Leblanc [10], and optimal weight for GMM is calculated as $W_n = n^{-1}\sum_{i=1}^{n} \rho_i(\hat{\gamma})\rho_i(\hat{\gamma})'$. Then the final estimates are computed using the estimated weights.

As is frequently the case in nonlinear numerical optimization, convergence, numerical complaints and other problems will be encountered. To avoid potential optimization problems involved in the iterative procedures, a direct grid search method is applied. In particular, $n_0 = 5000$ grid points per parameter are generated in each iteration. For each model, 1000 Monte Carlo repetitions are carried out for each of the sample sizes $n = 20$, $n = 30$, $n = 50$, $n = 100$, $n = 200$, $n = 500$. The Monte Carlo means (SLS, GMM3, GMM4) and their root mean squared errors (RMSE) are computed. The numerical computation is done using the statistical computer language R for Windows on a PC with standard configuration.

4.2. Summary of simulation results

Tables 1-6 report the results, where GMM3 denotes the estimator based on the first three moment conditions, and GMM4 denotes the estimator based on the first four moment conditions. We report the Monte Carlo means and their root mean squared errors.

The results show that the SLS performs reasonably well, especially for sample sizes above 100. Moreover, all estimates seem to be unbiased within the range $\pm 3SSE$ (Simulation Standard Error), whereas the GMM converges slower than SLS and, in some cases, it is still biased even for relatively large sample sizes. For example Figure 1 suggests a downward bias in the GMM3 and GMM4 for σ^2 but not in SLS.

In the most cases, SLS performs better than GMM3 and GMM4 in the sense that it has smaller RMSE and it decreases with the increase of a number of observations. Figure 2 compares RMSE of estimators for exponential model with two parameters. As we can see in this figure, SLS has smaller RMSE for all sample sizes than GMM3 and GMM4, and it decreases as sample size increases. Another fact in this figure is that GMM3 has smaller RMSE than GMM4. Since GMM4 uses more information than GMM3, we expect to have more precise estimators for GMM4 than GMM3, however this is not always true. In Koenker and Machado [4], they imply that GMM with higher number of moment equations needs more sample size to justify conventional asymptotic inference.

We have also simulated all models with normally distributed X and obtained similar results, except in logistic model with three parameters where SLS has larger RMSE for small sample sizes than GMMs. However, as sample size gets larger, SLS starts to dominate both GMM3 and GMM4. Figure 3 shows this result.

Table 6 shows the case $(\theta_1 = -3)$ that SLS has larger RMSE than GMMs even for sample size equal to 500. This fact can be seen in Figure 4 with the same model but normal X. In this case, we examined larger sample sizes $(n = 1000, 2000)$ and observed that SLS needs more sample size to perform better than GMMs.

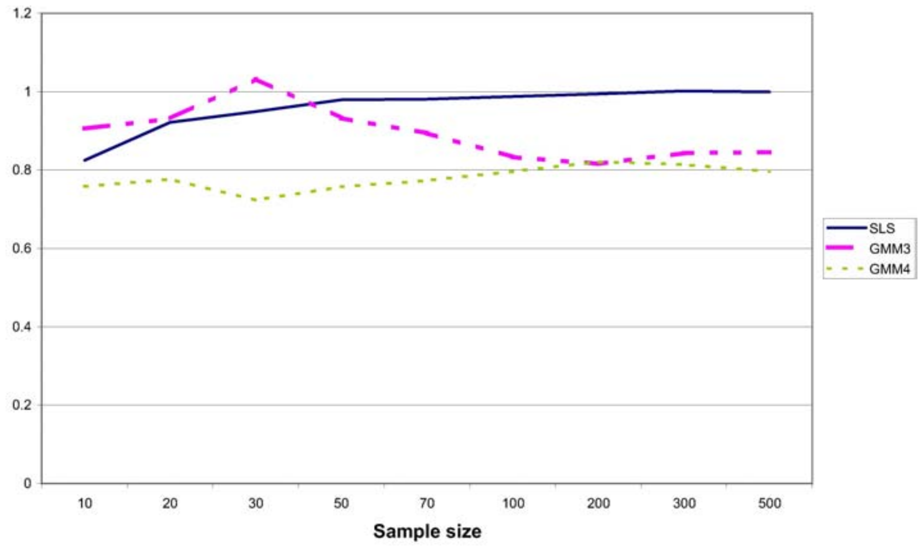

Figure 1. Estimation of Sigma 2 in exponential model with two parameters.

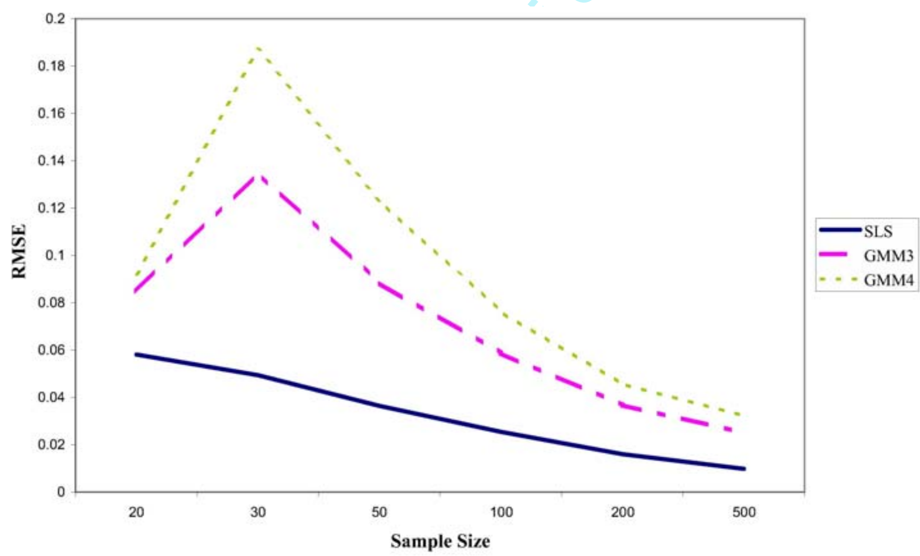

Figure 2. RMSE of Theata of estimators in exponential model.

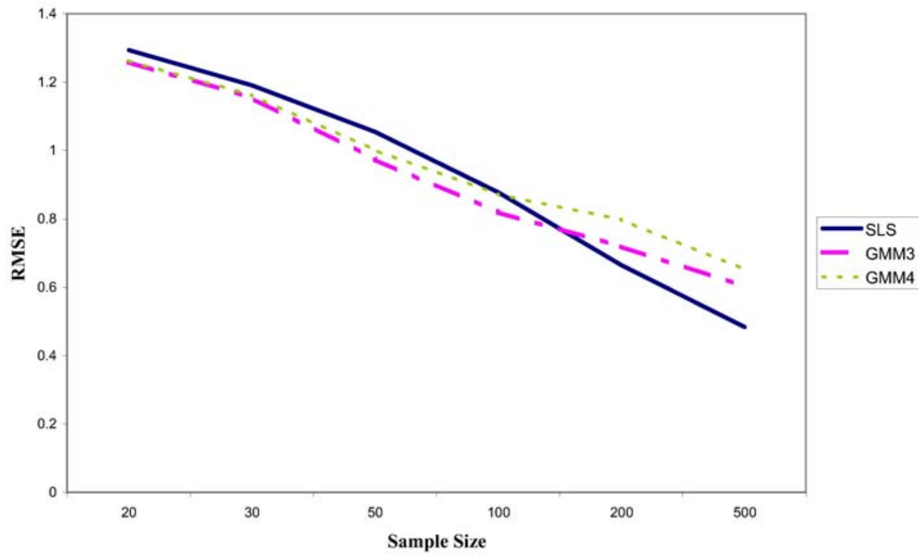

Figure 3. RMSE of Theata 0 of estimators in logistic model.

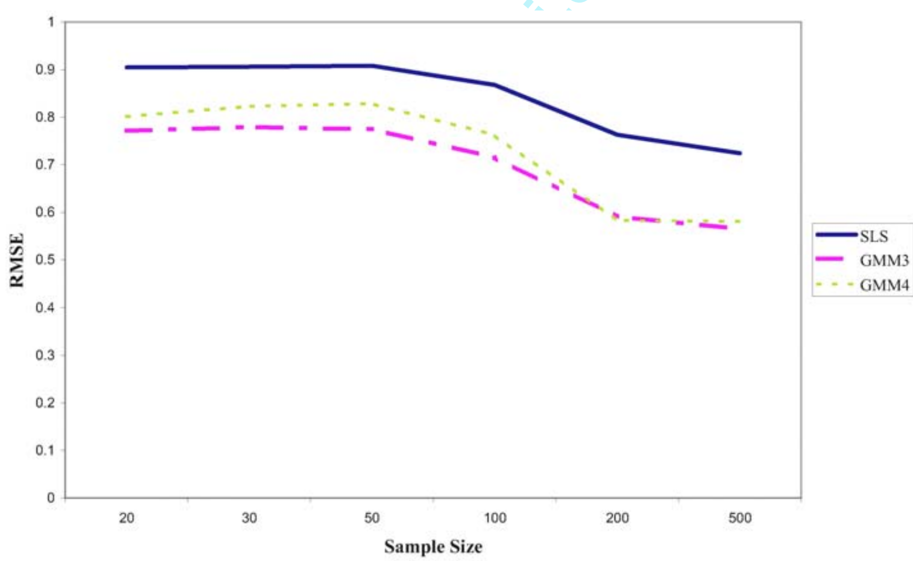

Figure 4. RMSE of Theata 0 of estimators in linear exponential model.

Table 1. Exponential model with two parameters

	$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
			$\theta = -0.5$			
SLS	–0.5083	–0.5092	–0.5052	–0.5031	–0.5001	–0.5004
	(0.0581)	(0.0494)	(0.0365)	(0.0255)	(0.0160)	(0.0098)
GMM3	–0.5113	–0.5378	–0.5126	–0.4988	–0.4933	–0.4949
	(0.0850)	(0.1351)	(0.0884)	(0.0588)	(0.0368)	(0.0251)
GMM4	–0.5096	–0.5535	–0.5173	–0.4986	–0.4925	–0.4914
	(0.0922)	(0.1871)	(0.1234)	(0.0764)	(0.0456)	(0.0323)
			$\sigma^2 = 1$			
SLS	0.9218	0.9494	0.9798	0.9882	0.9949	1.000
	(0.3109)	(0.2591)	(0.2005)	(0.1427)	(0.1003)	(0.0622)
GMM3	0.9317	1.0333	0.9331	0.8335	0.8159	0.8451
	(0.4005)	(0.3968)	(0.4086)	(0.4309)	(0.4399)	(0.4371)
GMM4	0.7772	0.7231	0.7579	0.7966	0.8211	0.7972
	(0.4698)	(0.5000)	(0.4729)	(0.4567)	(0.4398)	(0.4493)

Table 2. Logistic model with two parameters

	n = 20	n = 30	n = 50	n = 100	n = 200	n = 500
			$\theta = 50$			
SLS	50.0883	50.0104	50.0572	50.0112	50.0315	50.0101
	(1.3735)	(1.2320)	(1.0403)	(0.7832)	(0.5426)	(0.3419)
GMM3	50.0204	49.9089	49.9203	49.8564	49.9216	49.8778
	(1.4145)	(1.2863)	(1.1391)	(0.8998)	(0.6909)	(0.4849)
GMM4	49.9939	49.8795	49.8722	49.7912	49.8827	49.8232
	(1.4381)	(1.3484)	(1.1965)	(0.9950)	(0.7768)	(0.5993)
			$\sigma^2 = 1$			
SLS	0.9287	0.9661	0.9676	0.9927	0.9974	0.9956
	(0.2819)	(0.2477)	(0.1952)	(0.1357)	(0.0940)	(0.0628)
GMM3	0.8056	0.8121	0.8775	0.8370	0.8097	0.7939
	(0.5286)	(0.4923)	(0.4460)	(0.4290)	(0.4482)	(0.4452)
GMM4	0.7806	0.8037	0.8230	0.8150	0.8193	0.7978
	(0.5396)	(0.5025)	(0.4725)	(0.4463)	(0.4464)	(0.4543)

Table 3. Linear exponential model with two parameters

	$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
			$\theta_1 = -3$			
SLS	–2.9661	–3.0189	–3.0154	–3.0766	–3.0451	–3.0234
	(0.9426)	(0.7893)	(0.7482)	(0.7898)	(0.5711)	(0.3840)
GMM3	–3.1525	–3.0284	–2.9683	–3.0474	–2.9454	–2.9497
	(0.7459)	(0.5987)	(0.5675)	(0.5816)	(0.5727)	(0.5462)
GMM4	–3.2763	–3.0239	–2.9858	–3.0469	–2.9861	–2.9404
	(0.7951)	(0.6010)	(0.5753)	(0.5833)	(0.5677)	(0.5693)
			$\theta_2 = -1$			
SLS	–1.0993	–1.0299	–1.0235	–1.0182	–1.0104	–1.0026
	(0.2806)	(0.1612)	(0.1100)	(0.0798)	(0.0722)	(0.0381)
GMM3	–1.1712	–1.0620	–1.0435	–1.0180	–1.0351	–1.0274
	(0.4111)	(0.3026)	(0.1840)	(0.0949)	(0.1417)	(0.1136)
GMM4	–1.2492	–1.1064	–1.0639	–1.0268	–1.0354	–1.0392
	(0.4967)	(0.4240)	(0.2735)	(0.1204)	(0.1882)	(0.1667)

Table 4. Exponential model with two parameters

	n = 20	n = 30	n = 50	n = 100	n = 200	n = 500
			$\theta_1 = 10$			
SLS	10.0744	10.0608	10.0208	10.0227	10.0054	6.9832
	(1.2046)	(0.9414)	(0.6252)	(0.6030)	(0.3614)	(0.2789)
GMM3	9.9750	10.1089	10.2047	10.1560	10.2193	10.1981
	(1.1147)	(1.0884)	(0.8478)	(0.9315)	(0.8113)	(0.9006)
GMM4	10.0109	10.2007	10.3319	10.2152	10.4046	10.2912
	(1.1200)	(1.1262)	(0.9112)	(1.0557)	(0.9776)	(1.0189)
			$\theta_2 = -0.5$			
SLS	–0.5145	–0.5117	–0.5065	–0.5041	–0.5016	–0.4996
	(0.0911)	(0.0807)	(0.0508)	(0.0381)	(0.0256)	(0.0191)
GMM3	–0.5035	–0.5102	–0.5420	–0.5149	–0.5364	–0.5195
	(0.1048)	(0.1157)	(0.1424)	(0.0939)	(0.1236)	(0.0992)
GMM4	–0.5114	–0.5243	–0.5944	–0.5269	–0.5915	–0.5391
	(0.1349)	(0.1405)	(0.2136)	(0.1317)	(0.2019)	(0.1411)
			$\sigma^2 = 1$			
SLS	0.8667	0.9172	0.9540	0.9743	0.9843	0.9966
	(0.3241)	(0.2663)	(0.2119)	(0.1491)	(0.1082)	(0.0773)
GMM3	0.8935	0.8308	0.8481	0.8518	0.8719	0.8427
	(0.4064)	(0.4354)	(0.4318)	(0.4360)	(0.4245)	(0.4319)
GMM4	0.8215	0.8405	0.7808	0.8216	0.7825	0.8082
	(0.4412)	(0.4314)	(0.4678)	(0.4424)	(0.4558)	(0.4489)

Table 5. Logistic model with three parameters

	$n = 20$	$n = 30$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
			$\theta_1 = 50$			
SLS	49.9784	50.0449	50.1037	50.0983	50.0688	50.0295
	(1.3575)	(1.2445)	(1.0435)	(0.8200)	(0.5724)	(0.3400)
GMM3	49.8844	49.9033	49.9606	49.9089	49.8856	49.8674
	(1.3567)	(1.2304)	(1.0257)	(0.8463)	(0.6755)	(0.5429)
GMM4	49.8889	49.8748	49.9328	49.8543	49.8439	49.8212
	(1.3801)	(1.2498)	(1.0853)	(0.9465)	(0.8295)	(0.6843)
			$\theta_2 = 34$			
SLS	34.0087	34.0222	34.0016	33.9482	33.9556	33.9424
	(1.6347)	(1.5990)	(1.5192)	(1.3130)	(1.0907)	(0.7140)
GMM3	33.9469	33.9619	33.9106	33.9385	34.0086	34.0284
	(1.2782)	(1.2179)	(1.1494)	(1.1587)	(1.1623)	(1.1621)
GMM4	33.8767	33.9728	33.9098	33.9530	33.9826	34.0038
	(1.3000)	(1.2078)	(1.1578)	(1.1738)	(1.1249)	(1.1623)
			$\sigma^2 = 1$			
SLS	0.8941	0.9314	0.9641	0.9695	0.9848	0.9943
	(0.3011)	(0.2514)	(0.2020)	(0.1452)	(0.1002)	(0.0637)
GMM3	0.8330	0.7942	0.7966	0.7940	0.8108	0.8098
	(0.4976)	(0.4754)	(0.4596)	(0.4501)	(0.4361)	(0.4385)
GMM4	0.8189	0.8038	0.8084	0.7883	0.8137	0.8125
	(0.4929)	(0.4692)	(0.4503)	(0.4580)	(0.4311)	(0.4495)

Table 6. Linear exponential model with three parameters

	n = 20	n = 30	n = 50	n = 100	n = 200	n = 500
			θ_3 = 5			
SLS	4.9695	5.0889	5.1074	4.9739	5.0194	5.0629
	(1.6352)	(1.5366)	(1.2622)	(1.4251)	(0.9281)	(0.8469)
GMM3	4.9011	4.9196	5.0838	4.9434	5.1693	5.1909
	(1.1845)	(1.1895)	(1.1312)	(1.1403)	(1.0668)	1.0453
GMM4	4.9146	5.0329	5.0846	5.0251	5.2054	5.2024
	(1.1500)	(1.1421)	(1.1084)	(1.1481)	(1.1331)	(1.0956)
			θ_1 = –3			
SLS	–2.9963	–2.9494	–2.9362	–3.0091	–2.9845	–3.0040
	(0.8174)	(0.7654)	(0.7793)	(0.7339)	(0.6662)	(0.5772)
GMM3	–3.0569	–2.9786	–2.9660	–3.0118	–2.9210	–2.9919
	(0.6075)	(0.5907)	(0.5848)	(0.5904)	(0.5760)	(0.5631)
GMM4	–3.0787	–2.9685	–2.9638	–3.0021	–2.9519	–2.9644
	(0.5908)	(0.5729)	(0.5688)	(0.5685)	(0.5684)	(0.5921)
			θ_2 = –1			
SLS	–1.0466	–1.0569	–1.0296	–1.0175	–1.0133	–1.0057
	(0.1479)	(0.1947)	(0.1172)	(0.0916)	(0.0647)	(0.0396)
GMM3	–1.0362	–1.0432	–1.0558	–1.0242	–1.0826	–1.0096
	(0.1451)	(0.2103)	(0.2045)	(0.1327)	(0.2371)	(0.1792)
GMM4	–1.0503	–1.0729	–1.1147	–1.0499	–1.1092	–1.0841
	(0.1717)	(0.2650)	(0.3400)	(0.2035)	(0.3153)	(0.2681)

5. Conclusion

The generalized method of moments is widely used in econometrics, statistics, and many other fields. Although GMM estimators are consistent and asymptotically normally distributed under general regularity conditions, it has long been recognized that this asymptotic distribution may provide a poor approximation to the finite sample distribution. Moreover, GMM needs more moment equations than unknown parameters, which cause problems when parameter dimension increases. The second-order least squares estimation uses only the first two conditional moments of the response variable given the predictor variables. In this paper, it has been shown that when the number of moment equations and parameters are equal, the SLS is asymptotically more efficient than that of the GMM in the case of two parameters. Simulation studies show that SLS performs better than GMM even when it uses more moment equations.

References

[1] T. G. Anderson and B. E. Sorenson, GMM estimation of a stochastic volatility model: A Monte Carlo study, J. Bus. Econom. Statist. 14 (1996), 328-352.

[2] L. Hansen, Large sample properties of generalized method of moments estimators, Econometrica 50 (1982), 1029-1054.

[3] P. J. Huber, Robust Statistics, Wiley, New York, 1981.

[4] R. Koenker and J. A. F. Machado, GMM inference when the number of moments conditions is large, J. Econometrics 93 (1999), 327-344.

[5] L. Mátyás, Generalized Method of Moments Estimation, Cambridge University Press, 1999.

[6] K. Pearson, On the criterion that a given system of deviations from the probable in the case of a correlated system of variables in such that it can be reasonably supposed to have arisen from random sampling, Philosophical Magazine Series 50 (1900), 157-175.

[7] J. H. Stock and J. H. Wright, GMM with weak identification, Econometrica 68 (2000), 1055-1096.

[8] L. Wang, Estimation of nonlinear Berkson-type measurement error models, Statist. Sinica 13 (2003), 1201-1210.

[9] L. Wang, Estimation of nonlinear models with Berkson measurement errors, Ann. Statist. 32 (2004), 2559-2579.

[10] L. Wang and A. Leblanc, Second-order nonlinear least squares estimation, Working Paper, Department of Statistics, University of Manitoba, Canada, 2005.

[11] F. Windmeijer, A finite sample correction for the variance of linear efficient two-step GMM estimators, J. Econometrics 126 (2005), 25-51.

[12] J. H. Wright, Detecting lack of identification in GMM, Econometric Theory 19 (2003), 322-330.

■