Empirical Likelihood, Exponential Tilting, and GMM Estimators with a number of Moment Conditions

K. Hitomi
Kyoto Institute of Technology, Kyoto, Japan

Abstract: In an effort to improve the small sample properties of GMM, a number of alternative estimators have been suggested. These include Empirical Likelihood (EL) and Exponential Tilting (ET) estimators. This paper compares conventional GMM estimator to Empirical Likelihood (EL) and Exponential Tilting estimators when the number of moment conditions increases with the number of observations. The estimators are subject to a Monte Carlo investigation using the following specification. A linear equation includes endogenous explanatory variables and there exist a number of instrumental variables. The number of the instrumental variables is increased as the number of observations is increased. The main findings of the experiments show the following. Small sample biases of EL and ET are considerably smaller than GMM if the number of moment conditions is less than 12–15% of the number of observations. When the number of the moment conditions exceed 12–15% of the number of observations, small sample biases of EL and ET increase almost linearly as the number of moment conditions increase and the growth rates of the biases are greater than the one of GMM. The small sample bias of bias corrected GMM (Newey and Smith 2001) is always the same or smaller than the biases of other estimators. The standard deviations of GMM and bias corrected GMM estimators are decreased as the number of moment conditions is increased. The standard deviations of EL and ET estimators, however, are increased when the number of moment conditions exceeds 12–15% of the number of observations.

Keywords: *Empirical likelihood, GMM, Exponential Tilting, small sample bias, number of moment conditions*

1 INTRODUCTION

In the econometric literatures, the generalized method of moments (GMM) estimation method has been quite popular in the past decade. This approach has an attractive feature that it has rather broad applicability and it is easily implemented in statistical analyses. However, it has been known that there is a serious bias problem in the GMM estimation when there are many instruments in econometric models.

In recently, Empirical Likelihood (EL) and Exponential Tilting (ET) estimators have been proposed as alternative to the GMM method and have been gotten some attention in the statistical and econometric literatures. These methods give asymptotically efficient estimator in the semi-parametric sense and improve the serious bias problem known in the GMM method when the number of instruments is large in econometric models.

Newey and Smith (2001) studied higher order properties of generalized empirical likelihood (GEL) estimator by the stochastic expansion method and find that EL and ET have two theoretical advantages. First, asymptotic bias of EL and ET do not grew with the number of moment conditions, while the bias of the GMM does. Second, the bias corrected EL is higher efficient relative to the bias corrected GMM.

In econometric modeling, relatively large numbers of moment conditions are often used and infinite moment conditions are available in same cases, i.e. Carrasco and Florens (2000) and Press (1972).

This paper examines the small sample properties of Empirical Likelihood (EL) estimator with relatively large numbers of moment conditions and compares conventional GMM estimator to Empirical Likelihood (EL) and Exponential Tilting estimators when the number of moment conditions increases with the number of observations.

The rest of the paper is organized as followings. Section 2 provides brief explanations of the model and the estimators. Section 3 presents the design of Monte Carlo experiments. The results are presented in Section 4. Some conclusions are given in Section 5.

2 THE MODEL AND ESTIMATORS

Let z_i, $(i = 1, \ldots, n)$, be i.i.d. observations on a data vector z. Also, let θ be a $q \times 1$ parameter vector and $g(z, \theta)$ be an $m \times 1$ vector of functions of the data observation z and the parameter, where $m \geq q$. The model has a true parameter θ_0 satisfying the moment condition

$$E[g(z, \theta_0)] = 0, \tag{1}$$

where $E[.]$ denotes expectation taken with respect to distribution of z_i.

The standard two-step GMM solution to this estimation problem (1) is to estimate θ_0 as the solution

$$\min_{\theta} \left(\frac{1}{n} \sum_{i=1}^{n} g(z_i, \theta) \right)' \hat{W}^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} g(z_i, \theta) \right) \tag{2}$$

where $\hat{W}$ is a consistent estimator of $E[g(z_i, \theta_0) g(z_i, \theta_0)']$.

A new alternative to GMM is Empirical Likelihood. The idea is due to Art Owen (1988, 2001) and has been extended to the GMM context by Qin and Lawless (1994). It provided a non-parametric analog to likelihood estimation suitable for the GMM context. The idea is the follows. The distribution of z may be well approximated non-parametrically by multinomial distribution which places probability p_i at each observations z_i, with the constraints that $\sum_{i=1}^n p_i = 1$ and $\sum_{i=1}^n p_i g(z_i, \theta) = 0$. The latter constraint imposes the moment condition (1).

The Empirical Likelihood (EL) estimator for the vector of unknown parameter θ in (1) is defined as the solution to

$$\max_{\theta, p_1, \ldots, p_n} \sum_{i=1}^{n} \log p_i \tag{3}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} p_i g(z_i, \theta) = 0, \sum_{i=1}^{n} p_i = 1, \forall i \; p_i > 0.$$

The log likelihood statistic $-\sum_{i=1}^n \log(np_i)$ could be viewed as a measure of distance of $(p_1, p_2, \ldots, p_n)$ from the empirical measure $(1/n, 1/n, \ldots, 1/n)$. Other distances could be used for constructing another estimator. The Exponential Tilting (ET) estimator introduced by Kitamura and Stutzer (1997) and Imbens, Spady, and Johnson (1998) uses the Kullback-Leibler Information Criterion (KLIC) distance instead of the log likelihood. The ET estimator is defined as the solution to

$$\min_{\theta, p_1, \ldots, p_n} \sum_{i=1}^{n} p_i \log p_i \tag{4}$$

$$\text{s.t.} \quad \sum_{i=1}^{n} p_i g(z_i, \theta) = 0, \sum_{i=1}^{n} p_i = 1, \forall i \; p_i > 0.$$

All of these three estimators are consistent, asymptotically normal, and have the same asymptotic covariance matrix $(G'\Omega^{-1}G)^{-1}$ where $G = E[\partial g(z_i, \theta_0)/\partial\theta]$ and $\Omega = (E[g(z_i, \theta_0) g'(z_i, \theta_0)])^{-1}$. Small sample properties of these estimators, however, might be different each other. Newey and Smith (2001) have argued that ET and EL estimators have smaller bias than GMM estimator when the number of moment conditions, m, is large and the covariance matrix of the bias corrected EL estimator coincides the one of the maximum likelihood estimator with in the second order expansion.

3 Monte Carlo Design

In order to investigate the finite sample properties of GMM, EL, and ET estimators with a number of moment conditions, we have done a set of numerical simulations.

For this purpose, we set a simple linear model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i. \tag{5}$$

u_i is n.i.d (0,1) and $x_i = (x_{i1}, x_{i2})$ is correlated with u_i in the following way

$$\begin{aligned} x_{i1} &= w_{i1} + 0.5 u_i \qquad (6) \\ x_{i2} &= w_{i2} + 0.5 u_i \qquad (7) \end{aligned}$$

where $w_i = (w_{i1}, w_{i2})$ is n.i.d with mean (0,0) and covariance matrix

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

For setting up the m moment conditions, the following instrumental variables are used in the experiments

$$v_{ik} = \begin{cases} w_{i1} + e_{ik} & \text{if } k \text{ is odd} \\ w_{i2} + e_{ik} & \text{if } k \text{ is even} \end{cases} \tag{8}$$

where e_{ik} is uncorrelated n.i.d (0,1).

The moment conditions for estimating $\theta_0 = (\beta_0, \beta_1, \beta_2)$ are

$$g(z_i, \theta) = \begin{pmatrix} y_i - b_0 - b_1 x_{i1} - b_2 x_{i2} \\ (y_i - b_0 - b_1 x_{i1} - b_2 x_{i2}) v_{i1} \\ (y_i - b_0 - b_1 x_{i1} - b_2 x_{i2}) v_{i2} \\ \cdots \\ (y_i - b_0 - b_1 x_{i1} - b_2 x_{i2}) v_{im} \end{pmatrix}$$

where z_i is $z_i = (y_i, x_{i1}, x_{i2}, v_{i1}, v_{i2}, \ldots, v_{im})$. True parameter $\theta_0 = (\beta_0, \beta_1, \beta_2) = (1, 1, 1)$ is chosen for the true parameter values.

The following combinations of the number of observations, n, and the number of moment conditions, m, were examined.

- $n = 50$ and $m = 5, 7, 9, 11, 13, 15, 25$.
- $n = 100$ and $m = 5, 7, 9, 11, 13, 15, 25, 45$.

- $n = 200$ and $m = 5, 7, 9, 11, 13, 15, 25, 45, 65$.

The number of replication was 5000 for all cases.

On the above settings, we compared small sample biases and standard deviations of GMM, EL, and ET estimators. In addition, we examined finite sample properties of bias corrected GMM, EL, and ET estimators that were described in Newey and Smith (2001).

4 SIMULATION RESULTS

The results were summarized in Figure 1 to 9. In all figures, horizontal axes were the number of moment conditions and vertical axes indicated biases or standard deviations. Each line in Figures had a label something like el50-bc, the el indicates EL estimator (et is Exponential Tilting and gmm is GMM), 50 implies the number of observation is 50, and bc denotes bias corrected version of the estimator.

Figure 1 reports biases of constant term, b_0, when the number of observations is 50. As shown in the figure, biases were negligible for all estimator and all the numbers of moment conditions. It was also the same in the case of $n = 100$ and $n = 200$, so these two cases were not reported.

Figure 2 – 4 displayed the biases of the slope coefficient, b_1. As shown in the figures, the bias of GMM estimator increased as the number of moment conditions increased in all cases. Small sample biases of EL and ET are almost zero if the number of moment conditions is less than 12–15% of the number of observations. When the number of the moment conditions exceed about 12–15% of the number of observations, small sample biases of EL and ET increase almost linearly as the number of moment conditions increase and the growth rates of biases were much faster than GMM. The small sample bias of bias corrected GMM (Newey and Smith 2001) is always the same or smaller than the biases of other estimators, and considerably smaller than other estimators when the number of moment conditions exceed about 12–15% of the number of observations. It might be noteworthy that ET estimators had smaller biases than EL estimators did.

Figure 5 and 6 represent the standard deviations of each estimator of the constant term b_0 when the numbers of observations are 50 and 200, respectively. The standard deviations of EL and ET estimators were increased when the number of moment conditions rose above 12–15% of the number of observations. The standard deviations of GMM estimator were almost independent of the number of observations.

Figure 7 – 9 represent the standard deviations of each estimator of the slope coefficient b_1 when the numbers of observations are 50, 100, and 200, respectively. Since more a lot of moment conditions imply more a lot of information, it is expected that the standard deviations of estimators will be decreased as the number of moment conditions increased. The standard deviations of GMM and bias corrected GMM estimators are decreased as the number of moment conditions is increased as expected. The standard deviations of EL and ET estimators, however, are increased when the number of moment conditions exceeds 12–15% of the number of observations.

Figure 10 to 12 are the distribution of EL estimator of the slope parameter b_1 when the number of observations is 100 and the numbers of moment conditions are 5, 25 and 45, respectively. It could be observed that the distribution of b_1 was shifting to right and became flatter as the number of moment conditions increased.

To sum up the results of the simulations,

- The biases of EL and ET are very small until the number of moment conditions is less than 12–15% of the number of observations. On the other hand, the bias of GMM increases almost linearly as the number of moment conditions increases.

- when the number of moment conditions exceeds 12–15% of the number of moment conditions, the bias of EL and ET increases rapidly than the bias of GMM. The bias of the bias corrected GMM is the same or smaller bias than the ones of other estimators in all the numbers of moment conditions.

- the standard deviation of GMM decreases as the number of moment conditions increases. The standard deviations of EL and ET increase as the number of moment conditions increases when the number of moment conditions exceeds 12–15% of the number of observations.

5 CONCLUDING REMARKS

It has been pointed out a number of times that the GMM estimators have serious small bias problem when the number of moment conditions is large in econometric models (Hansen, Heaton and Yaron 1996). Information theoretic estimators, like empirical likelihood (EL) and exponential tilting (ET) estimator, have been expected to improve the small sample bias problem. Our result of the Monte Carlo simulations, however, draws a different picture. The

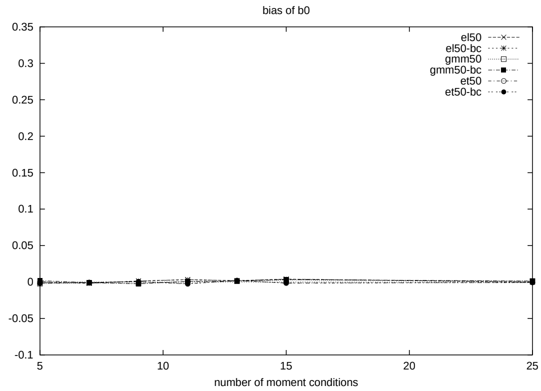

Figure 1: Bias of b_0 n=50

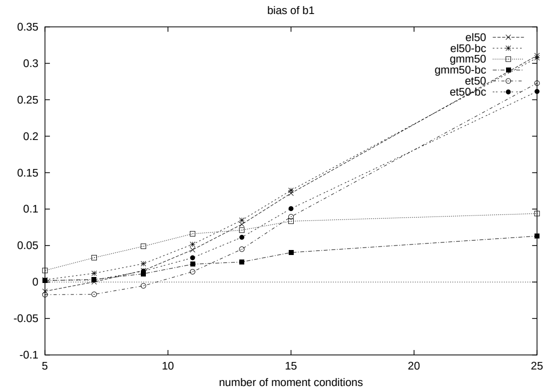

Figure 2: Bias of b_1 n=50

biases and the standard deviations of EL and ET estimators are lager than ones of GMM estimators at least the number of moment conditions is large.

This paper examines small sample properties of GMM, EL and ET estimators when the number of moment conditions was relatively large. Our Monte Carlo experiments shows that EL and ET estimators have small biases when the number of moment conditions are about less than 12–15% of the number of observations. However, the number of moment conditions are increased, the biases and the standard deviations of ET and EL estimators also increased. The bias corrected GMM has best small sample properties with respect to bias and standard deviations in our linear model settings.

The asymptotic theory of EL and ET estimators implicitly assumes that the origin was included in interior of the convex hull of $g(z_i, \theta_0)$ $i = 1, \ldots n$. With a finite sample, this convex hull conditions might be violated with a positive probability. This violation might explain the bad small sample properties of EL and ET estimators when the number of moment conditions was large.

References

M. Carrasco and J.-P. Florens. Generalization of gmm to a continuum of moment conditions. *Econometric Theory*, 16(6):797–834, 2000.

R. H. S. Guido W. Imbens and P. Johnson. Information theoretic approaches to inference in moment condition models. *Econometrica*, 66(2):333–357, 1998.

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982.

L. P. Hansen, J. Heaton, and A. Yaron. Finite-sample properties of some alternative gmm estimators. *Journal of Business and Economic Statistics*, 14:262–280, 1996.

Y. Kitamura and M. Stutzer. An information-theoretic alternative to generalized method of moments estimation. *Econometrica*, 65(4):861–874, 1997.

W. K. Newey and R. J. Smith. Higher order properties of gmm and generalized empirical likelihood estimators. 2001.

A. B. Owen. Empirical likelihood ratio confidence regions. *The Annals of Statistics*, 22:300–325, 1990.

A. B. Owen. Empirical likelihood for linear models. *The Annals of Statistics*, 19:1725–1747, 1991.

A. B. Owen. *Empirical likelihood*. Number 92 in Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, 2001.

V. K. S. P. Rilstone and A. Ullah. The second-order bias and mean-squared error of nonlinear estimators. *Journal of Econometrics*, 75:369–395, 1996.

J. Press. Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association*, 67(340):842–846, 1972.

J. Qin and J. Lawless. Empirical likelihood and general estimating equations. *The Annals of Statistics*, 20(1):300–325, 1994.

R. J. Smith. Alternative semi-parametric likelihood approaches to generalized method of moments estimation. *Economic Journal*, 107:503–519, 1997.

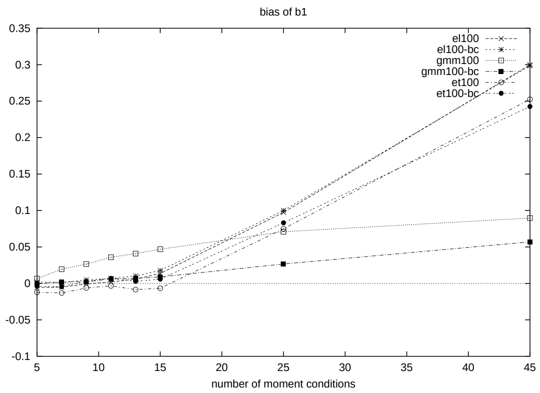

Figure 3: Bias of b_1 n=100

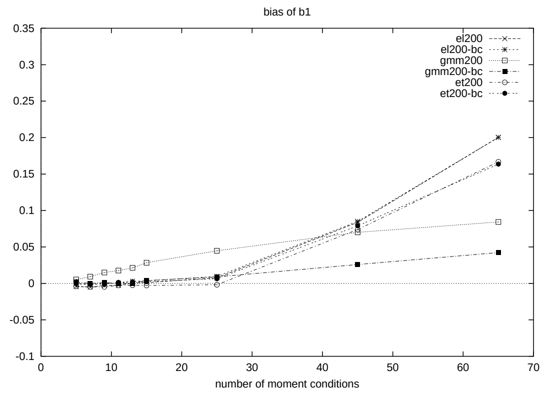

Figure 4: Bias of b_1 n=200

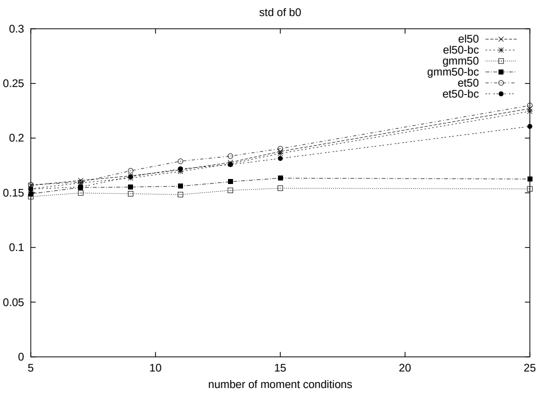

Figure 5: Standard Deviations of b_0 n=50

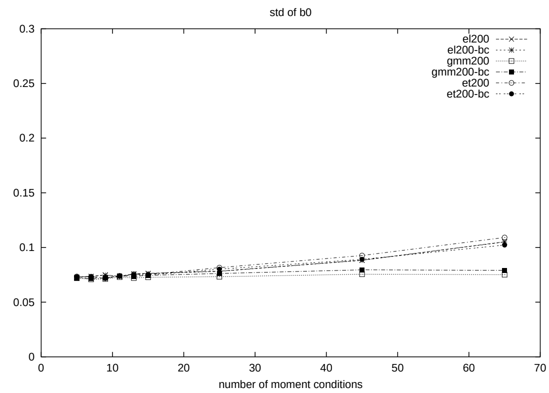

Figure 6: Standard Deviations of b_0 n=200

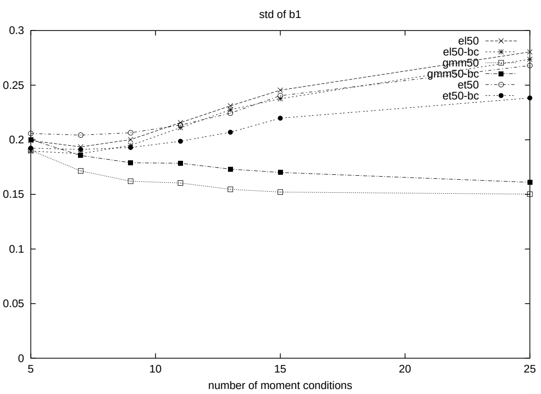

Figure 7: Standard Deviations of b_1 n=50

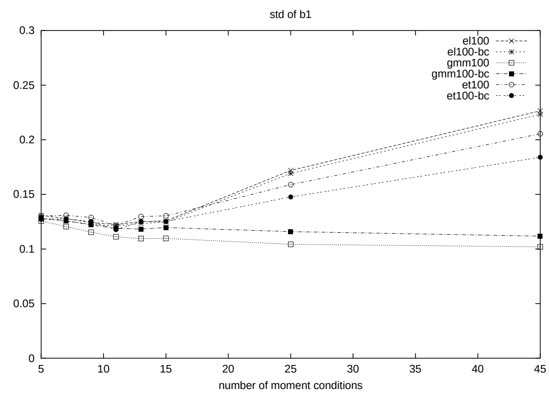

Figure 8: Standard Deviations of b_1 n=100

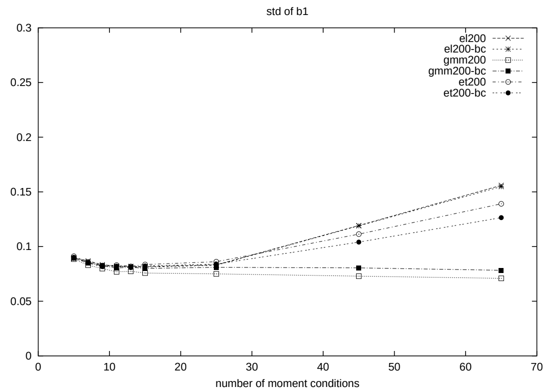

Figure 9: Standard Deviations of b_1 n=200

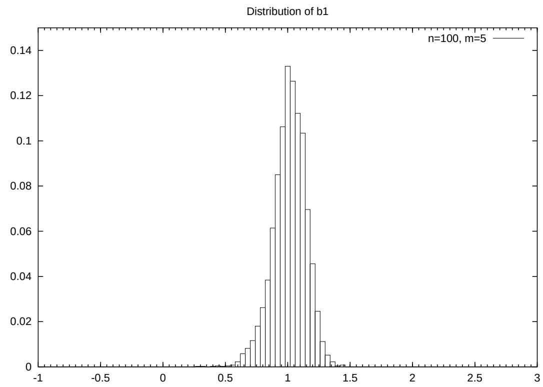

Figure 10: Distribution of EL estimator of b_1 n=100 and m=5

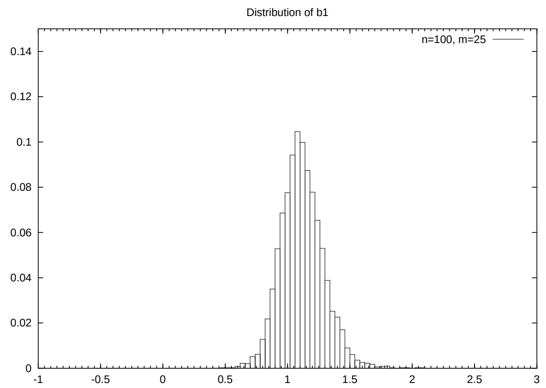

Figure 11: Distribution of EL estimator of b_1 n=100 and m=25

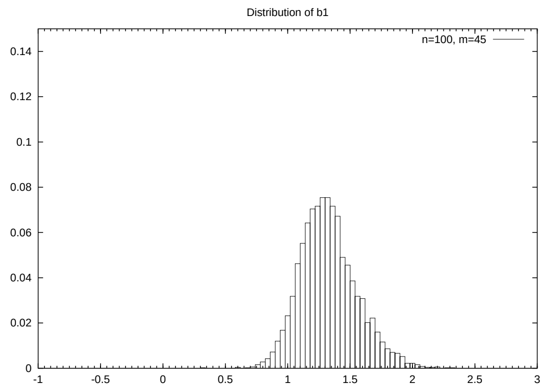

Figure 12: Distribution of EL estimator of b_1 n=100 and m=45