

Entropy Type Estimator to Simple Linear Measurement Error Models

Amjad D. Al-Nasser
Yarmouk University, Jordan

Abstract: The classical maximum likelihood estimation fails to estimate the simple linear measurement error model, with or without equation error, unless additional assumptions are made about the structural parameters. In the literature there are six different assumptions that could be added in order to solve the measurement error models. In this paper, we proposed an entropy-type estimator based on the generalized maximum entropy estimation approach, which allows one to abstract away from the additional assumptions that are made in the classical method. Monte Carlo experiments were carried out in order to investigate the performance of the proposed estimators. The simulation results showed that the entropy-type estimator of unknown parameters has outperformed the classical estimators in terms of mean square error criterion.

Zusammenfassung: Ohne zusätzliche Annahmen über die Strukturparameter zu treffen versagt die klassische Maximum Likelihood Methode bei der Schätzung im einfachen linearen Measurement Error Modell mit oder ohne Gleichungsfehler. In der Literatur gibt es sechs verschiedene Annahmen, die herangezogen werden können um das Messfehlermodell zu lösen. In diesem Aufsatz schlagen wir einen Schätzer basierend auf den Ansatz der generalisierten Maximum Entropie Schätzung vor, der es uns erlaubt ohne zusätzliche Annahmen auszukommen, welche bei der klassischen Methode gemacht werden. Um die Güte des vorgeschlagenen Schätzers zu untersuchen wurden Monte Carlo Experimente durchgeführt. Die Ergebnisse der Simulation zeigten, dass der auf die Entropie basierende Schätzer für die Parameter die klassischen Schätzer bezüglich des mittleren quadratischen Fehlers übertrifft.

Keywords: Generalized Maximum Entropy, Maximum Likelihood Estimation, Equation-Error Model and No-Equation-Error Model.

1 Introduction

Consider a bivariate random variable (ξ, η) satisfying a linear relationship $\eta = \alpha + \beta\xi$, where α and β are unknown parameters, ξ and η are unobservable latent variables that can only be observed with additive errors. Thus, instead of observing ξ and η directly, one observes the variables $x = \xi + \delta$ and $y = \eta + \varepsilon$. Such relationship is a generalization of the regression model and well known as measurement error model (MEM) without equation error. For a sample of size n , $\{(x_i, y_i) : i = 1, \dots, n\}$, the no-equation-error-model can be formulated as

$$\begin{aligned}x_i &= \xi_i + \delta_i \\y_i &= \eta_i + \varepsilon_i\end{aligned}\tag{1}$$

where

$$\eta_i = \alpha + \beta\xi_i, \quad i = 1, \dots, n \quad (2)$$

and (ξ_i, η_i) are unobservable true variables and $(\delta_i, \varepsilon_i)$ are additive errors.

We make the customary assumptions that $\xi \sim N(\mu, \sigma^2)$, i.e., we are dealing with structural relationships, where ε and δ are uncorrelated $N(0, \sigma_1^2)$ and $N(0, \sigma_2^2)$, respectively, and distributed independently of ξ . Thus, the joint distribution of (x, y) is bivariate normal $N(\mu, \alpha + \beta\mu, \sigma^2 + \sigma_2^2, \beta^2\sigma^2 + \sigma_1^2, \rho)$, where

$$\rho = \frac{\beta\sigma^2}{\sqrt{(\beta^2\sigma^2 + \sigma_1^2)(\sigma^2 + \sigma_2^2)}}. \quad (3)$$

Consider the situation in which (ξ, η) are not perfectly linearly related and there is an error in the equation, say ζ . Hence, a no-equation-error model can be modified such that (2) is replaced by

$$\eta_i = \alpha + \beta\xi_i + \zeta_i, \quad i = 1, \dots, n,$$

where $\zeta \sim N(0, \sigma_3^2)$ is independent of ξ . Furthermore, it is assumed that ξ , ζ and ε , δ are independent. Thus, the joint distribution of (x, y) for the equation-error model is also bivariate normal $N(\mu, \alpha + \beta\mu, \sigma^2 + \sigma_2^2, \beta^2\sigma^2 + \sigma_1^2 + \sigma_3^2, \rho_1)$, where

$$\rho_1 = \frac{\beta\sigma^2 + \sigma_{12}}{\sqrt{(\beta^2\sigma^2 + \sigma_1^2 + \sigma_3^2)(\sigma^2 + \sigma_2^2)}} \quad (4)$$

and σ_{12} is the covariance between ε and δ .

Lindley (1947) first demonstrated the inability to obtain maximum likelihood estimators (MLE) due to the fact that these parameters are unidentified, and remarked on the need to make some additional assumption(s) about the parameter values to alleviate this difficulty. Extensive bibliographies that clarified this problem were given by Madansky (1959), Kendall and Stuart (1979), Cheng and Ness (1999), and Fuller (1987). The later state that there are six side assumptions found in the literature, any one of which will make the normal structural no-equation-error model is identifiable. The most common assumption is that the ratio of the residual variances $\lambda = \sigma_1^2/\sigma_2^2$ is known. Other assumptions are: the reliability ratio is known, either σ_1^2 , σ_2^2 or both are known and the last prior assumption that could be added is that the intercept is known. However, for the equation-error model, the new formulation of the main equation by adding the error term might mean that at least some of the side assumptions no longer imply the identifiability of the slope. In fact, among the six side assumptions that can be used in the no-equation-error model, $\lambda = \sigma_1^2/\sigma_2^2$ and σ_1^2 are not enough to be known in order to identify the slope. The typical assumption used is that σ_3^2 is unknown but the measurement error covariance matrix

$$\Omega = \begin{bmatrix} \sigma_2^2 & \sigma_{12} \\ \sigma_{12} & \sigma_1^2 \end{bmatrix}$$

is known. More details can be found in Chen and Van Ness (1999). The MLE solutions for the normal structural MEM under different assumptions are given in Table 1. Therein we use the quantities

$$\bar{x} = n^{-1} \sum_{i=1}^n x_i, \quad \bar{y} = n^{-1} \sum_{i=1}^n y_i,$$

Table 1: MLE of the slope for normal structural MEM models

	No-Equation-Error Model	Equation-Error Model
Known Parameter	$\hat{\beta}$	$\hat{\beta}$
λ	$\left[\frac{s_{yy} - \lambda s_{xx} + \sqrt{(s_{yy} - \lambda s_{xx})^2 + 4\lambda s_{xy}^2}}{2s_{xy}}, \right. \\ \left. s_{xy} \neq 0 \right]$	Unidentified
Reliability Ratio	$\kappa_{\xi}^{-1} s_{xy} / s_{xx} \\ s_{xx} \neq 0$	$\kappa_{\xi}^{-1} s_{xy} / s_{xx} \\ s_{xx} \neq 0$
σ_1^2	$(s_{yy} - \sigma_1^2) / s_{xy}^*$	Unidentified
σ_2^2	$s_{xy} / (s_{xx} - \sigma_2^2)^{**}$	$s_{xy} / (s_{xx} - \sigma_2^2)^{**}$
σ_1^2 and σ_2^2	$\left[\frac{s_{yy} - \lambda s_{xx} + \sqrt{(s_{yy} - \lambda s_{xx})^2 + 4\lambda s_{xy}^2}}{2s_{xy}}, \right. \\ \left. s_{xy} \neq 0 \right]$	$s_{xy} / (s_{xx} - \sigma_2^2)^{**}$
α (Intercept)	$(\bar{y} - \alpha) / \bar{x}, \\ \bar{x} \neq 0$	$(\bar{y} - \alpha) / \bar{x}, \\ \bar{x} \neq 0$

* assumes that $s_{yy} > \sigma_1^2$ and $s_{xx} \geq s_{xy}^2 / (s_{yy} - \sigma_1^2)$,

** assumes that $s_{xx} > \sigma_2^2$ and $s_{yy} \geq s_{xy}^2 / (s_{xx} - \sigma_2^2)$.

$$s_{xx} = n^{-1} \sum_{i=1}^n (x_i - \bar{x})^2, \quad s_{yy} = n^{-1} \sum_{i=1}^n (y_i - \bar{y})^2, \quad s_{xy} = n^{-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$\kappa_{\xi} = \sigma^2 / (\sigma^2 + \sigma_2^2), \quad \hat{\alpha} = \bar{y} - \hat{\beta} \bar{x}.$$

However, when the data exist in terms of noisy observations, the generalized maximum entropy (GME) estimator proposed by Golan et al. (1996) allows one to abstract away from the distribution and additional assumptions that are made in the traditional MLE method.

The remainder of this paper is divided into three sections. Section 2 presents the Generalized maximum entropy estimation approach to MEM with and without equation-error. Section 3 presents Monte Carlo evidence on the numerical performance of GME and MLE. Section 4 presents concluding remarks and suggestions for future works.

2 Methodology: GME Estimation Procedure

Entropy is a concept in probability theory and the maximum entropy is applicable when we are determining a function that can be regarded as probability distribution. The En-

tropy of a distribution has a rich history dating back to Shannon (1948), Jaynes (1957b), Jaynes (1957a), Kullback (1959), Levine (1980), Skilling (1989), Csiszar (1991), Donho et al. (1992), Golan et al. (1996), Golan et al. (1997), Al-Nasser (2003a), Al-Nasser (2003b), and Al-Nasser (2004).

The idea underling the GME approach is to view each unknown parameters, and error terms as the expected value of some proper probability distribution defined over some supports. The researcher supplies these supports. Then by maximizing the joint entropies (Shannon's entropy) subject to the data, represented by each unobserved values, and the requirement for proper probability distributions, we can achieve better estimates than the traditional one with less assumptions. Note that the no-equation-error model given in (1) and (2) can be rewritten as

$$y_i = \alpha + \beta x_i - \beta \delta_i + \varepsilon_i, \quad i = 1, \dots, n.$$

Then by using the GME the problem can be solved after some reformulation of the unknowns α , β , δ_i , ε_i , $i = 1, \dots, n$, by reparameterizing their possible outcome values probabilistically as a convex combination of random variables. This combination is presented as an expected value of some proper probability distribution.

Consistent with this specification, let α be represented by a discrete random variable, a_r , $r = 1, \dots, R$ with $R \geq 2$ possible realizations and corresponding probabilities q_1, \dots, q_R

$$\alpha = \sum_{r=1}^R a_r q_r, \quad \text{where } q_r \in (0, 1), \quad \text{and } \sum_{r=1}^R q_r = 1.$$

Similarly we can rewrite β as

$$\beta = \sum_{k=1}^K z_k p_k, \quad \text{where } p_k \in (0, 1), \quad \text{and } \sum_{k=1}^K p_k = 1.$$

The restriction imposed on the parameter space through (a, z) for (α, β) , reflects a prior knowledge about these parameters. However, if we know the possible values of the parameters from the theory then we specify (a, z) accordingly. If we don't, then we specify them to be uniformly symmetric around 0 with high lower upper bounds. For example, $z = (-c, 0, c)$, with c being a large value. Moreover, assuming one specifies (a, z) to span the true values of (α, β) , then the GME is a consistent estimator. Which is an advantage of this method. Furthermore, the empirical GME literature indicates that, in general, the values of R and K are 5.

The disturbance δ_i can be treated in a similar fashion. For each observation the associated disturbance, δ_i , is assumed to be bounded between two finite values, v_t^* and v_T^* , with corresponding probability weights w_{1t}^* and w_{nT}^* . That is, each disturbance may be modelled as

$$\delta_i = \sum_{t=1}^T v_t^* w_{it}^*, \quad \text{where } w_{it}^* \in (0, 1) \quad \text{and } \sum_{t=1}^T w_{it}^* = 1, \quad i = 1, \dots, n.$$

In a similar fashion, we can reparameterize the other disturbance part

$$\varepsilon_i = \sum_{j=1}^J v_j w_{ij}, \quad \text{where } w_{ij} \in (0, 1) \quad \text{and } \sum_{j=1}^J w_{ij} = 1, \quad i = 1, \dots, n.$$

The actual bounds for v_t^* and v_j depend on the observed sample as well as any conceptual or empirical information about the underlying error. However, if such conceptual or empirical information does not exist, then v_t^* , v_j may be specified to be uniformly and symmetrically distributed around zero. Chebychev's inequality may be used as a conservative means of specifying sets of error bounds. For any random variable X , such that $E(X) = 0$ and $\text{var}(X) = \theta^2$, the inequality provides

$$P_r(|X| < d\theta) \geq 1 - 1/d^2, \quad d \geq 1.$$

Then the Chebchyev's error bounds are $v_1^* = -d\theta$ and $v_T^* = d\theta$. One can use a 3σ rule (see Pukelsheim, 1994) to specify the error bounds, which state that the probability for X falling away from the mean by more than 3 standard deviations is at most 5%. Hence, the bound of v_t^* can be chosen such that $v_t^* \in [-3S_x, 3S_x]$, where S_x is the sample standard deviation associated with x_i . Further, v_j can be observed from y_i by using the same rules. The empirical GME literature indicates that, in general, the values of J and T are 3. Moreover, the unobservable ξ_i can be obtained from the differences between the observed data values x_i and its correspondent disturbance term δ_i .

2.1 GME Formulation and Solution for the No-Equation-Error Model

Using the maximum entropy formalism, the generalized stochastic nonlinear maximum entropy principle may be stated in scalar summations with four nonnegative probability components, as

$$\begin{aligned} \max_{q,p,w,w^*} H(q,p,w,w^*) = & - \sum_{r=1}^R q_r \log(q_r) - \sum_{k=1}^K p_k \log(p_k) \\ & - \sum_{i=1}^n \sum_{j=1}^J w_{ij} \log(w_{ij}) - \sum_{i=1}^n \sum_{t=1}^T w_{it}^* \log(w_{it}^*), \end{aligned}$$

subject to the following constraints

$$\begin{aligned} y_i = & \sum_{r=1}^R a_r q_r + \sum_{k=1}^K z_k p_k \left(x_i - \sum_{t=1}^T v_t^* w_{it}^* \right) + \sum_{j=1}^J v_j w_{ij}, \quad i = 1, \dots, n \\ \sum_{r=1}^R q_r = & 1, \quad \sum_{k=1}^K p_k = 1, \quad \sum_{j=1}^J w_{ij} = 1, \quad \sum_{t=1}^T w_{it}^* = 1, \quad i = 1, \dots, n. \end{aligned}$$

Here, we have $3n + 2$ constraints and $R + K + n(T + J)$ unknowns. To solve this system, we have to find the first order conditions, which can be obtained from the Lagrangian function given by

$$L = H(q,p,w,w^*) + \sum_{i=1}^n \gamma_i \left[y_i - \sum_{r=1}^R a_r q_r - \sum_{k=1}^K z_k p_k \left(x_i - \sum_{t=1}^T v_t^* w_{it}^* \right) - \sum_{j=1}^J v_j w_{ij} \right]$$

$$\begin{aligned}
& +\mu_1 \left(1 - \sum_{r=1}^R q_r \right) + \mu_2 \left(1 - \sum_{k=1}^K p_k \right) \\
& + \sum_{i=1}^n \psi_i \left(1 - \sum_{j=1}^J w_{ij} \right) + \sum_{i=1}^n \tau_i \left(1 - \sum_{t=1}^T w_{it}^* \right)
\end{aligned}$$

where $\gamma_i, \mu_1, \mu_2, \tau_i, \psi_i, i = 1, \dots, n$, are Lagrangian multipliers. Then using Lagrange's method of undetermined multipliers, this leads to the solution

$$\hat{q}_r = \frac{\exp \left(-a_r \sum_{i=1}^n \hat{\gamma}_i \right)}{\sum_{h=1}^R \exp \left(-a_h \sum_{i=1}^n \hat{\gamma}_i \right)}, \quad r = 1, \dots, R \quad (5)$$

$$\hat{p}_k = \frac{\exp \left(-z_k \sum_{i=1}^n \hat{\gamma}_i \left(x_i - \sum_{t=1}^T v_t \hat{w}_{it}^* \right) \right)}{\sum_{j=1}^K \exp \left(-z_j \sum_{i=1}^n \hat{\gamma}_i \left(x_i - \sum_{t=1}^T v_t \hat{w}_{it}^* \right) \right)}, \quad k = 1, \dots, K \quad (6)$$

$$\hat{w}_{ij} = \frac{\exp(-\hat{\gamma}_i v_j)}{\sum_{h=1}^J \exp(-\hat{\gamma}_i v_h)}, \quad i = 1, \dots, n, \quad j = 1, \dots, J \quad (7)$$

$$\hat{w}_{it}^* = \frac{\exp \left(-\hat{\gamma}_i v_t^* \sum_{k=1}^K z_k \hat{p}_k \right)}{\sum_{j=1}^T \exp \left(-\hat{\gamma}_i v_j^* \sum_{k=1}^K z_k \hat{p}_k \right)}, \quad i = 1, \dots, n, \quad t = 1, \dots, T \quad (8)$$

Hence, the estimated parameters can be obtained from

$$\begin{aligned}
\hat{\alpha} &= \sum_{r=1}^R a_r \hat{q}_r, & \hat{\beta} &= \sum_{k=1}^K z_k \hat{p}_k, & \hat{\delta}_i &= \sum_{t=1}^T v_t^* \hat{w}_{it}^*, \\
\hat{\delta}_i &= \sum_{t=1}^T v_t^* \hat{w}_{it}^*, & \hat{\varepsilon}_i &= \sum_{j=1}^J v_j \hat{w}_{ij}, & \hat{\xi}_i &= x_i - \hat{\delta}_i, \quad i = 1, \dots, n, \quad (9)
\end{aligned}$$

where $\hat{q}_r, \hat{p}_k, \hat{w}_{it}^*$, and \hat{w}_{ij} are given in (5)–(8), respectively.

Consequently, the structural parameters in (1)–(2) can be estimated based on the estimators in (9) as

$$\begin{aligned}
\text{var}(\hat{\varepsilon}) &= \hat{\sigma}_1^2 = \text{var} \left(\sum_{i=1}^n \sum_{j=1}^J v_j \hat{w}_{ij} \right) = \sum_{i=1}^n \sum_{j=1}^J \hat{w}_{ij}^2 \text{var}(v_j) \\
&= \sum_{i=1}^n \sum_{j=1}^J \hat{w}_{ij}^2 \left[\sum_{r=1}^J v_r^2 \hat{w}_{ik} - \left(\sum_{l=1}^J v_l \hat{w}_{il} \right)^2 \right] \quad (10)
\end{aligned}$$

Similarly,

$$\text{var}(\hat{\delta}) = \hat{\sigma}_2^2 = \sum_{i=1}^n \sum_{t=1}^T \hat{w}_{it}^{*2} \left[\sum_{r=1}^T v_r^{*2} \hat{w}_{ir}^* - \left(\sum_{l=1}^T v_l^* \hat{w}_{il}^* \right)^2 \right]. \quad (11)$$

2.2 GME Formulation and Solution for the Equation-Error Model

The estimation procedure described above can be used for the equation-error-model given in (1) and (3), where the error term ζ can be reparameterized as

$$\zeta_i = \sum_{m=1}^M v_m^{**} w_{im}^{**}, \quad \text{where } w_{im}^{**} \in (0, 1) \quad \text{and} \quad \sum_{m=1}^M w_{im}^{**} = 1, \quad i = 1, \dots, n.$$

Then the generalized stochastic nonlinear maximum entropy principle for equation-error model can be stated as

$$\begin{aligned} \max_{q,p,w,w^*,w^{**}} H(q,p,w,w^*,w^{**}) = & - \sum_{r=1}^R q_r \log(q_r) - \sum_{k=1}^K p_k \log(p_k) - \sum_{i=1}^n \sum_{j=1}^J w_{ij} \log(w_{ij}) \\ & - \sum_{i=1}^n \sum_{t=1}^T w_{it}^* \log(w_{it}^*) - \sum_{i=1}^n \sum_{m=1}^M w_{im}^{**} \log(w_{im}^{**}) \end{aligned}$$

subject to the following constraints

$$\begin{aligned} y_i = & \sum_{r=1}^R a_r q_r + \sum_{k=1}^K z_k p_k \left(x_i - \sum_{t=1}^T v_t^* w_{it}^* \right) + \sum_{j=1}^J v_j w_{ij} + \sum_{i=1}^M v_m^{**} w_{im}^{**}, \\ \sum_{r=1}^R q_r = & 1, \quad \sum_{k=1}^K p_k = 1, \quad \sum_{j=1}^J w_{ij} = 1, \quad \sum_{t=1}^T w_{it}^* = 1, \quad \sum_{m=1}^M w_{im}^{**} = 1, \quad i = 1, \dots, n. \end{aligned}$$

Here, we have $4n + 2$ constraints and $R + K + n(T + J + M)$ unknowns. The optimal solution leads to the same form that obtained in no-equation-error model as given in (5)–(8), the new estimator of the additional error term is

$$\hat{\zeta}_i = \sum_{m=1}^M v_m^{**} \hat{w}_{im}^{**}, \quad i = 1, \dots, n,$$

where

$$\hat{w}_{im}^{**} = \frac{\exp(-\hat{\gamma}_i v_m^{**})}{\sum_{j=1}^M \exp(-\hat{\gamma}_i v_j^{**})}, \quad i = 1, \dots, n, \quad m = 1, \dots, M$$

and

$$\begin{aligned} \hat{\sigma}_3^2 = & \text{var} \left(\sum_{i=1}^n \sum_{m=1}^M v_m^{**} \hat{w}_{im}^{**} \right) = \sum_{i=1}^n \sum_{m=1}^M \hat{w}_{im}^{**2} \text{var}(v_m^{**}) \\ = & \sum_{i=1}^n \sum_{m=1}^M \hat{w}_{im}^{**2} \left[\sum_{r=1}^M v_r^{**2} \hat{w}_{ir}^{**} - \left(\sum_{l=1}^M v_l^{**} \hat{w}_{il}^{**} \right)^2 \right]. \end{aligned} \quad (12)$$

3 Monte Carlo Evidence On Numerical Performance

To illustrate the performance of GME and MLE procedures, a simulation study was carried out by generating 1000 samples according to the structural relationship $y_i = 1 + x_i + \varepsilon_i$ and $x_i = \xi_i + \delta_i$, $i = 1, \dots, n$, where $\xi \sim N(0, \sigma^2)$, $\varepsilon \sim N(0, \sigma_1^2)$, and $\delta \sim N(0, \sigma_2^2)$, with $\sigma_1^2 = \sigma_2^2 = 0.1$, $\sigma^2 = 0.5$, and sample sizes $n = 15, 25, 40$, and 50 . For the equation-error model we assume that $\zeta \sim N(0, 0.5)$ and $\sigma_{12} = 0$. Moreover, GME estimators were implemented with five support values for the parameters α and β support interval $[-100, -50, 0, 50, 100]$. Based on the empirical three-standard-deviations rule, the error terms were implemented with 3 support values in the interval $[-3S_x, 0, 3S_x]$ and $[-3S_y, 0, 3S_y]$. However, the equation error term in the case of equation-error model were implemented with 3 support values in the interval $[-3(S_y - \sigma_1), 0, 3(S_y - \sigma_1)]$. Two experiments were conducted in order to study the comparison between the two methods.

Experiment 1 is carried out to study the performance of GME and MLE in estimating the intercept and the slope of the simple MEM with and without equation error in terms of mean square error (MSE). The results of this experiment are presented in Table 2. It can be noted that GME estimators have a lower MSE for all simulations results. Hence, the GME approach out performs the MLE with respect to the MSE criterion.

Experiment 2 is carried out to study the performance of GME in estimating the correlation coefficient in (3) and (4), based on the structural parameters β , σ^2 , σ_1^2 , σ_2^2 , and σ_3^2 . Note that the correlation coefficient of no-equation-error model can be estimate by

$$\hat{\rho} = \frac{\hat{\beta}\hat{\sigma}^2}{\sqrt{(\hat{\beta}^2\hat{\sigma}^2 + \hat{\sigma}_1^2)(\hat{\sigma}^2 + \hat{\sigma}_2^2)}}.$$

Unfortunately, the MLE (by considering the most common assumption when the ratio of the residual variances $\lambda = \sigma_1^2/\sigma_2^2$ is known) can be obtained by substituting the slope estimate as given in Table 1. The other structural parameter estimates are $\hat{\sigma}^2 = s_{xy}/\hat{\beta}$, $\hat{\sigma}_2^2 = (s_{yy} - 2\hat{\beta}s_{xy} + \hat{\beta}^2)/(\lambda + \hat{\beta}^2)$, and $\hat{\sigma}_1^2 = \lambda\hat{\sigma}_2^2$ as given in Chen and Van Ness (1999). Moreover, the GME estimate of the correlation coefficient can be obtained by substitute the parameter estimates from equations (9)–(11).

However, the MLE for the correlation coefficient of equation-error model, by assuming that both variances are known, can be estimate by

$$\hat{\rho}_1 = \frac{\hat{\beta}\hat{\sigma}^2 + \sigma_{12}}{\sqrt{(\hat{\beta}^2\hat{\sigma}^2 + \sigma_1^2 + \hat{\sigma}_2^2)(\hat{\sigma}^2 + \sigma_2^2)}},$$

where $\hat{\sigma}_3^2 = S_{yy} - \hat{\beta}S_{xy} - \sigma_1^2$. The GME estimates are obtained from (9)–(11) and (12).

According to the sampling situation described before, and for the correlation coefficients $|\rho| = 0.1, 0.2, \dots, 0.9$, we investigate the performance of MLE and GME by measuring the accuracy of the estimation procedure using the sample mean absolute errors, $\text{MAE}(\hat{\rho}) = (1000)^{-1} \sum_{i=1}^{1000} |\hat{\rho}_i - \rho|$, in Table 3. Then comparisons between the two methods in terms of mean square error $\text{MSE}(\hat{\rho}) = (1000)^{-1} \sum_{i=1}^{1000} (\hat{\rho}_i - \rho)^2$ are given in Table 4 for no-equation-error model and in Table 5 for the equation-error model.

Table 2: Comparison between GME and MLE with different samples sizes

n	Method	Known Parameter(s)	No-equation-Error Model		Equation-Error Model	
			MSE($\hat{\alpha}$)	MSE($\hat{\beta}$)	MSE($\hat{\alpha}$)	MSE($\hat{\beta}$)
15	GME		9.0978E-03	6.2656E-04	1.130E-03	4.1646E-03
	MLE	λ	4.0108E-02	3.4488E-03	**	**
		σ_1^2	9.9188E-03	8.4778E-04	**	**
		σ_2^2	1.0158E-02	9.3788E-04	1.414E-03	2.3783E-02
		σ_1^2, σ_2^2	4.0108E-02	3.4488E-03	1.414E-03	2.3783E-02
		κ_ξ	9.6745E-03	7.3721E-04	1.2145E-03	6.1510E-03
α	*	3.1536E+02	*	15.639E+01		
25	GME		1.2708E-04	1.1057E-06	5.1957E-04	1.1968E-03
	MLE	λ	1.4447E-02	2.1468E-04	**	**
		σ_1^2	3.6615E-03	2.3921E-04	**	**
		σ_2^2	3.7451E-03	2.3948E-04	6.1666E-04	5.6262E-03
		σ_1^2, σ_2^2	1.4447E-02	2.1468E-04	6.1666E-04	5.6262E-03
		κ_ξ	3.6766E-03	2.1191E-04	5.4819E-04	2.5533E-03
α	*	4.8400E+02	*	1.0384E+02		
40	GME		1.2215E-04	1.7003E-07	5.2240E-05	6.9453E-04
	MLE	λ	6.1793E-03	1.3624E-04	**	**
		σ_1^2	1.2286E-03	9.4598E-05	**	**
		σ_2^2	1.2341E-03	9.6983E-05	6.3685E-05	2.0932E-03
		σ_1^2, σ_2^2	6.1793E-03	1.3624E-04	6.3685E-05	2.0932E-03
		κ_ξ	1.2220E-03	8.7264E-05	6.1095E-05	2.0930E-03
α	*	7.5165E+02	*	2.8462E+01		
50	GME		2.4136E-05	7.7712E-08	3.6815E-05	1.2354E-04
	MLE	λ	4.3291E-03	1.0937E-05	**	**
		σ_1^2	8.4385E-04	6.4090E-05	**	**
		σ_2^2	8.4438E-04	6.2260E-05	5.3674E-05	9.2148E-04
		σ_1^2, σ_2^2	4.3291E-03	1.0937E-05	5.3674E-05	9.2148E-04
		κ_ξ	8.3576E-04	5.6538E-05	4.9606E-05	8.4136E-04
α	*	1.4433E+02	*	2.2571E+01		

*Intercept is known. **Unidentified Case

Table 3 represents the sample MAE for both methods. It should be noted that the sample MAEs are not the estimates of the true MAEs, they are just measures of the accuracy of the estimates with 1000 replications. The results suggested that the estimates of ρ are quite reasonable; for example for sample size 50 the MAE for GME approach when the true correlation is 0.9 is 0.8942 and for the MLE is 0.8837.

Table 3: Sample Mean Absolute Errors (MAE) of correlation coefficient

n	15		25		40		50	
$ \rho $	GME	MLE	GME	MLE	GME	MLE	GME	MLE
0.1	0.0831	0.0457	0.0847	0.0475	0.0776	0.0480	0.0846	0.0488
0.2	0.1573	0.1506	0.1729	0.1526	0.1656	0.1532	0.1747	0.1536
0.3	0.2765	0.2583	0.2844	0.2591	0.2620	0.2597	0.2637	0.2608
0.4	0.3727	0.3625	0.3750	0.3653	0.3522	0.3665	0.3286	0.3653
0.5	0.4973	0.4687	0.4849	0.4685	0.4338	0.4701	0.4162	0.4699
0.6	0.5656	0.5718	0.5766	0.5735	0.5097	0.5732	0.5089	0.5729
0.7	0.6811	0.6743	0.6792	0.6765	0.6388	0.6765	0.6655	0.6765
0.8	0.7155	0.7790	0.7841	0.7797	0.7552	0.7800	0.7534	0.7802
0.9	0.8539	0.8834	0.8900	0.8836	0.8930	0.8835	0.8942	0.8837

Table 4: Mean Square Error of correlation coefficient: No-Equation-Error Model

n	15		25		40		50	
$ \rho $	GME	MLE	GME	MLE	GME	MLE	GME	MLE
0.1	0.3766	0.7730	0.3119	0.7809	0.4172	0.7830	0.3155	0.7819
0.2	0.4634	0.6276	0.4357	0.6279	0.4943	0.6284	0.4396	0.6280
0.3	0.4117	0.4830	0.3993	0.4831	0.4268	0.4834	0.4015	0.4832
0.4	0.3209	0.3557	0.3161	0.3558	0.3298	0.3559	0.3175	0.3558
0.5	0.2292	0.2472	0.2278	0.2473	0.2349	0.2474	0.2289	0.2473
0.6	0.1488	0.1583	0.1488	0.1583	0.1526	0.1584	0.1497	0.1583
0.7	0.0841	0.0890	0.0846	0.0890	0.0866	0.0892	0.0853	0.0890
0.8	0.0371	0.0395	0.0377	0.0395	0.0387	0.0396	0.0382	0.0396
0.9	0.0089	0.0098	0.0093	0.0098	0.0095	0.0099	0.0096	0.0099

Table 5: Mean Square Error of correlation coefficient: Equation-Error Model

n	15		25		40		50	
$ \rho $	GME	MLE	GME	MLE	GME	MLE	GME	MLE
0.1	0.0119	0.0188	0.0075	0.0115	0.0049	0.0072	0.0039	0.0056
0.2	0.0070	0.0124	0.0044	0.0076	0.0029	0.0048	0.0024	0.0037
0.3	0.0034	0.0073	0.0022	0.0045	0.0015	0.0028	0.0012	0.0022
0.4	0.0012	0.0036	0.0008	0.0022	0.0005	0.0014	0.0004	0.0011
0.5	0.0003	0.0012	0.0001	0.0007	0.0001	0.0005	0.0000	0.0003
0.6	0.0008	0.0001	0.0003	0.0001	0.0001	0.0000	0.0001	0.0000
0.7	0.0026	0.0004	0.0012	0.0002	0.0007	0.0001	0.0005	0.0001
0.8	0.0057	0.0020	0.0030	0.0011	0.0017	0.0007	0.0013	0.0006
0.9	0.0102	0.0049	0.0055	0.0028	0.0033	0.0017	0.0025	0.0014

Table 4 and Table 5 represent the MSE of correlation coefficients using GME and MLE methods and for different sample sizes. It can be noted that, for the true correlation $|\rho| \leq 0.5$ the GME estimates outperform the MLE; while increasing the degree of

associations $|\rho| > 0.5$ the MLE estimates begin slowly to out perform the GME for all sample sizes. These results support the GME as a good alternative to the traditional MLE in estimating the correlation coefficient when both variables are subject to error.

4 Concluding Remarks and Future Works

In this paper, we have suggested the using of GME method to estimate the simple measurement error model with and without equation error.

The theory of GME allow us to abstract away the additional assumptions that could be added by using the traditional MLE in recovering the unknowns from the MEM, and in estimating the correlation coefficient when both variables are subject to error. The main advantage of GME over MLE, it does not require any additional distributional assumption, and it does not require any assumption about the disturbance variances ratio or other structural parameters. Thus, all what the GME methods needs to be applicable can be obtained from the sample. In fact, the GME estimators appear to be succeeding where the MLE failed.

Simulation results (see Table 2 – Table 5) showed that the GME dominated the MLE estimators according to the mean square error criterion. Clearly, the potential of the GME method has not been fully explored here. To examine the robustness of the GME estimators the simulation experiments could be repeated using non-normal alternatives. Additional investigation should be made about the finite sample and asymptotic properties of the GME estimators. Moreover, its needed to explore the useful parameterizations of higher dimensional MEM such as semiparametric measurement error models or non-linear measurement error models, see (Carroll et al., 1995) and Roeder et al. (1996); and this will consider as a future work.

Although the GME method is used to estimate the costumer satisfaction index (see, Al-Nasser (2003a)), it may be also regarded as a contribution to the MEM literature.

Acknowledgement

The author would like to thank Professor Herwig Friedl, Editor of the Austrian Journal of Statistics, and a referee for their valuable suggestions that improved this paper.

References

- Al-Nasser, A. D. (2003a). Customer satisfaction measurement models: Generalized maximum entropy approach. *Pakistan Journal of Statistics*, 19, 213-226.
- Al-Nasser, A. D. (2003b). Fitting nonlinear regression models using generalized maximum entropy. *Abhath Al-Yarmouk: "Basic Sci. & Eng."*, 11, 145-155.
- Al-Nasser, A. D. (2004). Estimation of multiple linear functional relationships. *Journal Of Modern Applied Statistical Methods*, 3, 181-186.
- Carroll, R. J., Ruppert, D., and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models*. London: Chapman & Hall.

- Cheng, C.-L., and Ness, J. W. V. (1999). *Statistical Regression with Measurement Error*. New York: Arlond.
- Csiszar, I. (1991). Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. *The Annals of Statistics*, 19, 2032-2066.
- Donho, D. L., Johnstone, I. M., Hoch, J. C., and Stern, A. S. (1992). Maximum entropy and nearly black object. *Journal of the Royal Statistical Society, Series B*, 54, 41-81.
- Fuller, W. A. (1987). *Measurement Error Models*. New York: John Wiley.
- Golan, A., Judge, G., and Miller, D. (1996). *A Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York: John Wiley.
- Golan, A., Judge, G., and Perloff, J. (1997). Estimation and inference with censored and ordered multinomial response data. *Journal of Econometrics*, 79, 23-51.
- Jaynes, E. T. (1957a). Information and statistical mechanics (ii). *Physics Review*, 108, 171-190.
- Jaynes, E. T. (1957b). Information and statistical mechanics i. *Physics Review*, 106, 620-630.
- Kendall, M. G., and Stuart, A. (1979). *The Advanced Theory of Statistics*. London: Griffin.
- Kullback, J. (1959). *Information Theory and Statistics*. New York: John Wiley.
- Levine, R. D. (1980). An information theoretical approach to inversion problems. *Journal of Physics, A*, 13, 91-108.
- Lindley, D. V. (1947). Regression lines and linear functional relationship. *Journal of the Royal Statistical Society, Supplement*, 9, 218-244.
- Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. *Journal of the American Statistical Society*, 55, 173-205.
- Pukelsheim, F. (1994). The three sigma rule. *The American Statistician*, 48, 88-91.
- Roeder, K., Carroll, R. J., and Lindsay, B. G. (1996). A semiparametric mixture approach to case-control studies with errors in covariables. *Journal of the American Statistical Society*, 91, 722-732.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423.
- Skilling, J. (1989). The axioms of maximum entropy. In J. Skilling (Ed.), *Maximum entropy and bayesian methods in science and engineering* (p. 173-187). Dordrecht: Kluwer Academic.

Author's address:

Amjad D. Al-Nasser
Department of Statistics
Yarmouk University
Irbid, 21163
Jordan

E-mail: amjadn@yu.edu.jo