Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
`http://siba-ese.unisalento.it/index.php/ejasa/index`
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v16n2p487

New Wald-Type Estimation Procedures for Fitting Structural Measurement Error Model

By Al Dibi'i, Abdul Rahman, Al-Nasser

14 October 2023

Electronic Journal of Applied Statistical Analysis
Vol. 16, Issue 2, 2023, 487-506
DOI: 10.1285/i20705948v16n2p487

New Wald-Type Estimation Procedures for Fitting Structural Measurement Error Model

Ro'ya Al Dibi'i[a], Rosmanjawati Abdul Rahman*[a], and Amjad D. Al-Nasser[b]

[a] *Universiti Sains Malaysia,, School of Mathematical Sciences*
[b] *Yarmouk Universi1ty, Faculty of Science*

14 October 2023

This article proposes a new estimation method to fit the structural regression model when the variables are subject to errors. The new estimation method is the extension of the Wald estimation method and involves iterative process. Several Monte Carlo simulation experiments were used to study the performance of the proposed estimators. The results were compared with the classical Wald estimation method in terms of its root mean square error (RMSE). In addition, an application for examining the relationships between Jordan's national gross domestic product (GDP) and its human development index (HDI) was presented. Numerical results showed that the GDP and HDI have a strong positive and significant correlation. Moreover, the proposed procedures with different subgroup sizes (r =3 and r =4) gave more accurate estimators than the classical estimation methods in fitting the relationships between GDP and HDI.

keywords: Measurement Error Models, Wald Estimator, Repetitive Estimator, Human Development Index, National Gross Domestic Product, Monte Carlo Simulation.

1 Introduction

Measurement error models (MEMs) are one of the most interesting models used to study the relationships between two or more variables. These models were used when both

*Corresponding author: rosmanjawati@usm.my.

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

response and predictor variables are measured with error. In modeling, when both variables are contaminated by measurement errors, they are known by many names: MEM, errors-in-variables (EIV), error in covariance when the predictor is a continuous variable, misclassification when the predictor variable is discrete, or regression model when both variables are subject to error (Stefanski (2000); Gustafson (2021)). The MEM is an extension of simple linear regression and occurs in many practical studies; for example, but not limited to, this model often occurs in healthcare studies, econometrics, ecology, and many other research areas. One critical impact of modeling the data using the MEM for estimating the unknown parameters is that the classical estimation methods such as least squares or maximum likelihood estimation (MLE) do not consistently estimate the unknown parameters (Carroll et al. (1995)). The problem of fitting the MEM has been considered by many authors (see Madansky (1959); Fuller (1987); Cheng and Van Ness (1999); Gillard (2006, 2010)). In spite of developed estimation methods for fitting the MEM, this is still an unresolved area. The difficulty with the estimation approaches in the literature, whether parametric or non-parametric, is that they rely on additional assumptions or a very complex optimization procedure, which makes these methods fail to give reliable results, or they need prior assumptions to produce consistent estimators. On the other hand, the smoothing methods cannot clearly remove the effect of the error in the predictor. Moreover, some difficulty also appears in using data-driven methods, which mainly depend on the numerical or physical properties of the data, which makes them work perfectly with some data and fail with other selected data. The problem of estimating the MEM model has received increasing attention from many authors, with the growing realization that the predictors are either weakly defined or poorly measured, (Griliches (1974)). One of the most interesting estimation procedures was proposed by (Wald (1940)), which suggests dividing the data into two groups based on the median value and then fitting a line between the mean of both groups. Wald's estimators are consistent only when the error terms are normally distributed (Gupta and Amanull (1970)). Since then, there have been some developments in the Wald estimation procedure (see, for example, (Al-Nasser et al. (2016); Cheng and Van Ness (1999)). The purpose of this paper is to shed further light on the Wald-type estimation procedure for fitting the structural MEM. This paper is divided into eight sections. Section 2 describes the structural MEM. Section 3 reviews the classical estimation procedures: Wald-type, Maximum Likelihood and Method of Moment. Section 4 presents the proposed procedures (Repetitive procedure and Iterative procedure). Then in Section 5 several Monte Carlo experiments are presented to assess the performance of the proposed estimator in fitting the structural MEM. Then, in Section 6 a real data analysis is given for illustrating the relationships between the Human Development Index (HDI) and the National Gross Domestic Product (GDP); finally, some concluding remarks are given in Section 7.

2 Structural Measurement Error Model

The structural MEM is the one from three models of MEM that depending on the assumptions about (Cheng and Van Ness (1999)), whereas, if the ξ_i are independent identically distributed random variables and independent of the errors, then the model is known as a structural model, when $E(\xi_i) = \ \mu \ and \ Var \ \ (\xi_i) = \ \sigma^2$.

Let's consider a bivariate sample of size n: (x_1, y_1), (x_2, y_2), ... (x_n, y_n) subject to error observations, such that:

$E(Xi) = \xi_i$ and $E(y_i) = \eta_i$

where ξ_i and η_i, $i = 1, 2, \ ..., \ n$ are the true values of x_i and y_i respectively. Then the observed values have a linear relationship with the exact values given as:

$$x_i = xi_i + \delta_i, \ and \ y_i = \eta_i + \epsilon_i \tag{1}$$

where ξ_i, and ϵ_i are independent error δ terms ϵ of the true values and of each other's with zero mean and constant variances σ^2_δ and σ^2_ϵ, respectively. Also, the true values are linearly related in the form of:

$$\eta_i \ = \ \alpha \ + \ \beta \ \xi_i \quad i = 1, 2, ..., n \tag{2}$$

where α and β are unknown parameters to be estimated, under the following assumptions:

- The elements of the sequences δ_i and $\epsilon_i, i = 1, 2, ..., n$ are independently and identically distributed with zero mean and finite variances σ^2_δ and σ^2_ϵ respectively, that is: $E(\delta_i) = E(\epsilon_i) = 0, foralli, \ \ E(\delta_i\delta_j) = E(\epsilon_i\epsilon_j) = 0, i \neq \ j$ and $Var(\delta_i) = \sigma^2_\delta, \ Var(\epsilon_i) = \sigma^2_\epsilon$ for all i
- The elements of the sequences δ_i and $\epsilon_i, \ i = 1, 2, ..., n$ are mutually independent, where $E(\delta_i\epsilon_j) = 0$ for all i and j.
- δ_i and ϵ_i are also independent of the true values of ξ_i and η_i, in which: $E(\xi_i\delta i) = 0$, and $E(\eta_i\epsilon_i) = 0$.
- The true values of x_i and y_i are linearly related as: $y_i = \alpha + \beta x_i, \ i = 1, 2, ..., n$

  where α and β are the parameters to be determined.
- The limit inferior is considered as $|\frac{(x_1+x_2+.....+x_m)-(x_(m+1)+x_2+...+x_n)}{n}|$, n= 2,3,...,m, where $m = \frac{n}{2}$.
- All the even and odd-order moments of the error terms exist and are finite.

Now, combining the model given in (1) and (2) we will have the following relationship:

$$y_i \ = \ \eta_i + \epsilon_i \tag{3}$$

$$y_i = \alpha + \beta\xi_i + \epsilon_i \tag{4}$$

$$y_i \;= \alpha + \beta(x_i - \delta_i) + \epsilon_i \tag{5}$$

Measurement error, which occurs when a variable of interest is not accurately observed, is a problem that frequently calls into question the reliability of an analysis. Although there are numerous methods for correcting the effects of measurement error, when their underlying assumptions are violated, they become unreliable. This issue is exacerbated when presumptions, like those pertaining to the distribution of error terms, are hard or impossible to test using the available data. In the literature on measurement error, an additive model with normally distributed errors is frequently assumed. Therefore, the predictor is not independent from the error term, which violates the basic assumption of the least squares theory. Accordingly, other estimation methods are needed to solve this problem; one of the most interesting is the free distribution estimation method known as the Wald-type grouping procedure.

3 The Classical Estimations of MEM

Measurement error, which occurs when a variable of interest is not accurately observed, is a problem that frequently calls into question the reliability of an analysis. Although there are numerous methods for correcting the effects of measurement error, when their underlying assumptions are violated, they become unreliable. This issue is exacerbated when presumptions, like those pertaining to the distribution of error terms, are hard or impossible to test using the available data. In the literature on measurement error, an additive model with normally distributed errors is frequently assumed. However, when the predictor is not independent from the error term, it violates the basic assumption of the least squares theory. Accordingly, other estimation methods are needed to solve this problem; one of the most interesting is the free distribution estimation method known as the Wald-type grouping procedure. A brief discussion on the common methods used for estimating MEM are as follows:

3.1 Wald-Type Estimation Method

A popular and simple estimation method in the context of structural MEM is the Wald-type estimation method, also known as grouping methods (Wald (1940); Gillard (2010)). The main idea of the Wald- type estimation methods is to split the data into groups (two or three groups), then estimate the slope of the MEM based on the group's centers. To illustrate this procedure, consider a random sample of size n, say $(x_i \;, \; y_i)$ i=1, 2, ..., n. Then, based on the order statistics of X, the data is divided into two groups of the same size, such that:

- $Group1 = G1 = (x_{(i)}, y_{[i]}); x_{(i)} \leq Median(x)$
- $Group2 = G2 = (x_{(i)}, y_{[i]}); x_{(i)} > Median(x)$

where $x_{(i)}$ is the ith order statistic and y[i] is the associated i-th judgmental order statistic. In the case of odd sample size, the median is eliminated. Then, the unknown slope could be estimated by joining the mean of the low values of $X(G_1)$ to the mean of the high values of $X(G_2)$. Accordingly, we have the following two group estimators:

$$\hat{\beta} = \frac{\bar{y}_{G_2} - \bar{y}_{G_1}}{\bar{x}_{G_2} - \bar{x}_{G_1}} \tag{6}$$

and

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \tag{7}$$

where:

$$\bar{y}_{G_1} = \frac{y_{(1)} + \dots\dots + y_{(m)}}{m} \qquad \bar{y}_{G_2} = \frac{y_{(m+1)} + \dots\dots + y_{(n)}}{m}$$
$$\bar{x}_{G_1} = \frac{x_{(1)} + \dots + x_{(m)}}{m} \qquad \bar{x}_{G_2} = \frac{x_{(m+1)} + \dots + x_{(n)}}{m}$$
$$\bar{y} = \frac{\bar{y}_{G_1} + \bar{y}_{G_2}}{n} \quad \text{and} \quad \bar{x} = \frac{\bar{x}_{G_1} + \bar{x}_{G_2}}{n}$$

It could be noted that,

$$E\left(\hat{\beta}\right) = E\left(\frac{\bar{y}_{G_2} - \bar{y}_{G_1}}{\bar{x}_{G_2} - \bar{x}_{G_1}}\right) = E\left(\frac{(\alpha + \beta\bar{x}_2) - (\alpha + \beta\,\bar{x}_1)}{\bar{x}_2 - \bar{x}_1}\right) = \beta \tag{8}$$

with the associated variance given as:

$$Var\left(\hat{\beta}\right) = \frac{1}{(\bar{x}_2 - \bar{x}_1)^2}\left\{Var(\bar{y}_2, \bar{y}_1)\right\} = \frac{var(\bar{y}_2) + var(\bar{y}_1) - 2cov(\bar{y}_2, \bar{y}_1)}{(\bar{x}_2 - \bar{x}_1)^2} \tag{9}$$

Moreover,

$$E(\hat{\alpha}) = E(\bar{y} - \beta\,\bar{x}) = (\alpha + \beta\,\bar{x} - \beta\,\bar{x}) = \alpha \tag{10}$$

with

$$var(\hat{\alpha}) = var\left(\bar{y} - \hat{\beta}\,\bar{x}\right) = var(\bar{y}) + var\left(\hat{\beta}\right) - 2\,cov(\bar{y}_2, \bar{y}_1) \tag{11}$$

An extension of the two groups procedure was proposed by (Bartlett (1949); Nair and Shrivastava (1942)) where they suggested the splitting of the data into three equally sized sub-groups, G_1, G_2 and G_3 and after eliminating the middle group from the analysis, the unknown parameters were estimated as:

$$\hat{\beta} = \frac{\bar{y}_{G3} - \bar{y}_{G1}}{\bar{x}_{G3} - \bar{x}_{G1}} and \hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}. \tag{12}$$

where:

$$\bar{y}_{G1} = \frac{y_{(1)} + \dots\dots + y_{(m)}}{m} \qquad \bar{y}_{G3} = \frac{y_{(m+k+1)} + \dots\dots + y_{(n)}}{m}$$
$$\bar{x}_{G1} = \frac{x_{(1)} + \dots + x_{(m)}}{m} \qquad \bar{x}_{G3} = \frac{x_{(m+k+1)} + \dots + x_{(n)}}{m}$$

$$\bar{y} = \frac{\bar{y}_{G1} + \bar{y}_{G3}}{n}; \bar{x} = \frac{\bar{x}_{G1} + \bar{x}_{G2}}{n}. \tag{13}$$

In similar way that given in 5 and 7, it could be show that,

$$E\ \left(\hat{\beta}\right) = \beta \tag{14}$$

$$E\ \left(\hat{\alpha}\right) = \alpha \tag{15}$$

3.2 Maximum Likelihood Estimation Method

Several estimation methods for fitting the structural measurement error model (MEM) had been discussed in the literature. Among these, the least squares and maximum likelihood estimation (MLE) methods are the most commonly used. Based on a certain prior assumption(s), the MLE method is preferred (Lindley (1947)). earlier, (Madansky (1959)) provided detailed summary of using the MLE in the context of MEM. An alternative estimation approach is the grouping method that proposed by (Nair and Shrivastava (1942)). Moreover, (Gillard (2010)) proposed a Bayesian technique using the expectation-maximization (EM) algorithms to calculate the MLE for MEMs with or without equation error. Similarly,(Griliches (1974)) developed an iterative MLE procedure for estimating a heteroscedastic MEM.

The classical MLE of model (3) can be obtained by solving the log likelihood function, which is given by:

$$logL(\alpha, \beta, \sigma_\delta^2, \sigma_\epsilon^2, \xi_1,\ ..., \xi_n, \eta_1, ..., \eta_n) \tag{16}$$

$$= nlog(2\pi) - \frac{n}{2}log(\sigma_\delta^2) + log(\sigma_\epsilon^2 - \frac{\sum_{i=1}^n (x_i\ - \xi_i)}{2\sigma_\delta^2}\ - \frac{\sum_{i=1}^n (y_i\ -\ \alpha\ -\ \beta\xi_i)^2}{2\sigma_\epsilon^2} \tag{17}$$

However, the above likelihood function is unbounded. To see this, let us put and this results in σ_δ^2 approaching 0. The likelihood function will then approach infinity, irrespective of the values of α, β and σ_ε^2. An additional assumption is required about the variance ratio $\lambda = \frac{\sigma_\varepsilon^2}{\sigma_\delta^2}$, and it is assumed to be known. As a matter of fact, in the case of λ is unknown, the solution is a saddle point rather than maximum of likelihood surface.

Consider the simple linear structural MEM, given in (3); then under the regular assumption, this model can be estimated by solving a system of five equations as follows:
$E(X) = E\xi = \mu; EY = \mu_y = E\eta = \beta_0 + \beta_1\mu;$
$Var(X) = \sigma_x^2 = var(\xi) + \sigma_\delta^2 = \sigma^2 + \sigma_\delta^2;$
$Var(Y) = \sigma_y^2 = var(\eta) + \sigma_\epsilon^2 = \beta_1^2\sigma^2 + \sigma_\epsilon^2;$
$cov(X, Y) = \sigma_{xy}^2 = cov(\xi, \eta) = \beta_1\sigma^2.$

Meanwhile, one can use the first and second order moments to find the optimal estimators of the structural MEM, that is:

$\hat{\mu} = \bar{\chi} = \frac{1}{n}\sum_{i=1}^n x_i;$
$\hat{\mu}_y = \bar{y} = \frac{1}{n}\sum_{i=1}^n y_i;$

$$\hat{\sigma}_x^2 = s_{xx} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{\chi})^2;$$
$$\hat{\sigma}_y^2 = s_{yy} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{\chi})^2;$$
$$\hat{\sigma}_{xy}^2 = s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{\chi})(y_i - \bar{y}).$$

Let's replace $\hat{\mu}_y$, $\widehat{\sigma}_x^2$, $\hat{\sigma}_y^2$ and $\hat{\sigma}_{xy}^2$ by these values:

$$\bar{\chi} = \hat{\mu} \qquad ;$$
$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1\hat{\mu}; \qquad ;$$
$$s_{xx} = \hat{\sigma}^2 + \hat{\sigma}_\delta^2; \qquad ;$$
$$s_{yy} = \hat{\beta}_1^2\hat{\sigma}^2 + \hat{\sigma}_\varepsilon^2; \qquad ;$$
$$s_{xy} = \hat{\beta}_1\hat{\sigma}^2; \qquad .$$

Since $\hat{\sigma}^2 \geq 0$, $\hat{\sigma}_\delta^2 \geq 0$ and $\hat{\sigma}_\varepsilon^2 \geq 0$, then the following inequalities hold

$$s_{xx} \geq \frac{s_{yy}}{\hat{\beta}_1};$$
$$s_{yy} \geq \hat{\beta}_1 s_{xy};$$
$$s_{xx} \geq \hat{\sigma}_\delta^2;$$
$$s_{yy} \geq \hat{\sigma}_\varepsilon^2;$$
$$sign(s_{xy}) = sign\left(\hat{\beta}_1\right).$$

Since the number of equations is less than the number of parameters, additional information is needed to solve the system, which appears on six sides in the literature:

- The ratio of the error variances, $\lambda = \frac{\sigma_\epsilon^2}{\sigma_\delta^2}$, is known.
- The reliability ratio, K_ξ which is equal to $= \frac{\sigma^2}{\left(\sigma^2+\sigma_\delta^2\right)}$, is known.
- The variance σ_δ^2 is known.
- The variance σ_ε^2 is known.
- Both variances σ_ε^2 and σ_δ^2 are known.
- The intercept β_0 is known.

In this method, for example based on the first prior assumption, then the slops can be computed as follows:

$$\hat{\beta} = \frac{(Syy - \lambda Sxx) + \sqrt{(}Syy - \lambda Sxx^2 + 4\lambda Sxy^2}{2Sxy} \tag{18}$$

and

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \tag{19}$$

3.3 Method of Moments

Many mathematical statistics books, including (Casella and Berger (1990)), describe the method of moments (MOM) technique, though the treatment is brief here as it is elsewhere. They, like many other mathematical statistical texts, emphasized the maximum likelihood method. MOM estimators have been criticized because they are not uniquely defined, which means that if the method is used, it must be chosen from among possible estimators to find those that best suit the data being analyzed. This is demonstrated by applying the method to errors in variable regression theory. On the other hand, the MOM has the advantage of being simple, with the only assumptions being that low-order moments of the distributions describing the observations exist (Gillard and Iles (2005)).

In this method, the slops are computed as follows:

$$\hat{\beta}_1 = \frac{s_{xy}}{s_{xx} - \sigma_\delta^2}; s_{xx} > \sigma_\delta^2, \qquad s_{yy} > \frac{(S_{xy})^2 - \sigma_{\delta\varepsilon}}{s_{xx} - \sigma_\delta^2} \text{ and } \sigma_\delta^2 \text{ is known.} \tag{20}$$

where:
$s_{xy} = \sum xy - n \sum \bar{x}\bar{y}$,
$S_{xx} = \sum x_i^2 - n\bar{x}^2$
$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$

4 The Proposed Estimation Procedure

This paper proposes two new extensions of Wald-type procedures to improve the estimation of the structural MEM. The general idea of these procedures can be summarized as follows:

- Sort the x's values in ascending order from smallest to largest with their associated y's values $(x(i), y[i])$, i=1, 2, ..., n.
- Divide the sample into r-subgroups of equal size (i.e., say the subsample size is k) such that $r \leq [\frac{n}{2}]$.
- Compute the mean for each subgroup $\bar{x}_i, \bar{y}_i; i = 1, 2, ..., r$.

Next, the estimation of the unknown parameters involves the following two procedures:

Procedure 1

• The idea in this procedure is to compute all possible slopes (Figure.1). This can be done by defining the jth slope iteratively as:

$$\hat{\beta}_j = \frac{\bar{y}_{mj} - \bar{y}_{(m-1)j}}{\bar{x}_{mj} - \bar{x}_{(m-1)j}}, \ j = 1, 2, \ldots\ldots\ldots, \binom{r}{2}, \ m = 1, 2, \ldots\ldots\ldots r. \tag{21}$$

Accordingly the unknown parameters can be estimated as:

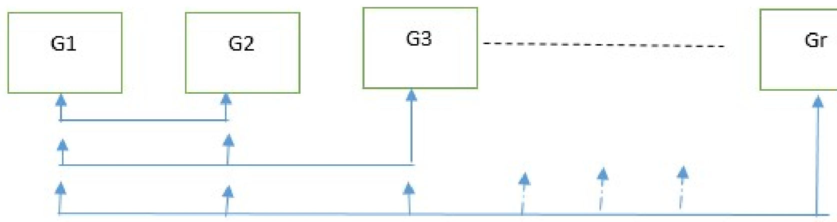

Figure 1: All possible slopes between the subgroups

$$\hat{\beta}_{p1} = \frac{1}{\binom{r}{2}} \sum_j \hat{\beta}_j \quad and \quad \hat{\alpha}_{p1} = \bar{y} - \hat{\beta}_{p1}\bar{x} \tag{22}$$

Procedure 2 In this procedure the idea is to compute the pairwise slopes continuously and gradually from one subgroup to another subgroup (Figure 2). Therefore, the jth slope in this case can be computed as:

$$\hat{\beta}_j = \frac{\bar{y}_{mj} - \bar{y}_{(m-1)j}}{\bar{x}_{mj} - \bar{x}_{(m-1)j}}, j \;=\; 1,\; 2,\; \ldots\ldots\ldots, (r-1)\,, \qquad m = 1, 2,\; \ldots,\; r-1 \tag{23}$$

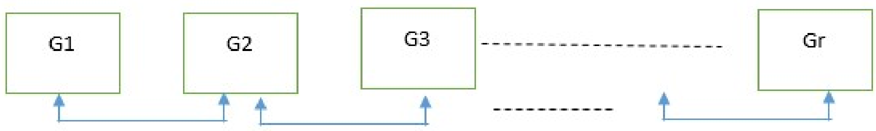

Figure 2: Pairwise slope between the subgroups

- Then the unknown parameters of the MEM can be estimated as

$$\hat{\beta}_{p2} = \frac{1}{(r-1)} \sum_j \hat{\beta}_j \quad and \quad \hat{\alpha}_{p2} = \bar{y} - \hat{\beta}_{p2}\bar{x} \tag{24}$$

Theorem: Assuming that the model in Eq.1-2 is satisfied, then the estimators given in Eq. 16 and Eq.18 are unbiased estimators. Proof:

$$E\left(\hat{\beta}_j\right) = E\left(\frac{\bar{y}_{mj} - \bar{y}_{(m-1)j}}{\bar{x}_{mj} - \bar{x}_{(m-1)j}}\right) \tag{25}$$

with associated variance given as:

$$Var(\hat{\beta}) = \frac{1}{\bar{x}_{mj} - \bar{x}^2_{(m-1)j}} Var(\bar{y}_{mj}, \bar{y}_(m-1)j) \tag{26}$$

$$= \frac{var\left(\bar{y}_{mj}\right) + \ var\left(\bar{y}_{(m-1)j}\right) - 2cov\left(\bar{y}_{mj}\ , \bar{y}_{(m-1)j}\right)}{\left(\bar{x}_{mj} - \bar{x}_{(m-1)j}\right)^2} \tag{27}$$

Also,

$$E\ (\hat{\alpha}) = E\ \left(\bar{y} - \ \hat{\beta}_{p2}\ \bar{x}\right)$$

$$= \left(\alpha + \ \hat{\beta}_{pi}\ \bar{x} - \ \hat{\beta}_{pi}\ \bar{x}\right) = \ \alpha \quad i = 1,\ 2 \tag{28}$$

with

$$var\ \ (\hat{\alpha}) = var\ \left(\bar{y} - \ \hat{\beta}_{pi}\ \bar{x}\right)$$

$$= var\left(\bar{y}\right) + \ var\left(\hat{\beta}_{pi}\right) - 2\ cov\ \left(\bar{y}_m\ , \bar{y}_{(m-1)}\right) \tag{29}$$

5 Monte Carlo Experiment

Monte Carlo simulations and experiments were carried out to evaluate the performance of the proposed procedures. In these experiments, 10,000 random samples were generated from the structural MEM, each of size n = 50, 100, 200, and 500. Two cases were considered within the sample by assuming the data were either with or without some outliers. The simulations were done with the following set-up:

- Order the data from the smallest to the largest with their respective associated $Yi'si = 1, 2, ..., n$, where:

$$\begin{aligned} \eta_i = &\ \alpha + \beta_1 \xi_i & i = 1,\ 2, \ldots, n \\ y_i = &\ \eta_i + \ \epsilon_i & \\ x_i = &\ \xi_i + \ \delta_i & i = 1,\ 2, \ldots, n \end{aligned}$$

- Set the initial values for the model parameters are $\alpha = 0, \beta = 1, \sigma^2_\epsilon = 1, \lambda = 1, \sigma^2_\delta = 1$ and $\sigma^2_\xi = 1$.
- Divide each sample into r =3 ,4 subsamples when dealing with the classical and proposed Wald-type procedures. Noted that for MOM and MLE procedures we didn't divide the sample for subgroups.
- In the case where outlier exists, the data are considered impure. In each step, a set number of observations (10%) were extracted and replaced with outliers.

An impure data set was formed based on the following specifications:

- Outliers exist only in y. In this case, the variance of the response error is assumed to be $(\epsilon_i N(0, \sigma_\epsilon^2), \sigma_\epsilon^2 = 16$.
- Outliers exist only in x. In this case, the variance of the response error is assumed to be $(\delta_i N(0, \sigma_\delta^2), \sigma_\delta^2 = 16$.
- Outliers exist both in y and x. Then the error terms were generated from normal distribution with variances $(\sigma_\delta^2, \sigma_\epsilon^2) = (16, 16)$.

The performances of these estimators were measured by computing the simulated bias and mean square error, which are represented as:

$$Bias = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{\mu}_i - \mu) \;\;\&\;\; MSE = \frac{1}{10000} \sum_{i=1}^{10000} (\hat{\mu}_i - \mu)^2 \tag{30}$$

with $\hat{\mu}_i$is the estimates given by one of the proposed estimators for the ith sample. The results of the Monte Carlo experiment are presented in Table 1 for the inliers cases. Table 2, Table 3, and Table 4 each have outliers in x only, y only, and in both x and y, respectively. The simulated results as reported in Tables 2-4 suggested that the proposed estimation method is better than the classical Wald-type procedure for each different sample size, when the data contains outliers. Also, as the sample size increases, the proposed procedures outperform the classical Wald-type procedure in terms of bias and MSE for both parameters. On the other hand, for the inlier cases shown in Table 1, the classical Wald-type procedure is more efficient compared to the proposed estimation method.

6 Real Data Application

In the past, a nation's overall development level was determined by its national income because it was believed that the more a nation produced, the more progress it would make both economically and socially. However, we acknowledge that there may be significant differences between societal progress or overall development and GDP growth. Over the past two decades, there has been much discussion about the limitations of using GDP as a gauge of a country's quality of life or social well-being. The fact that a large portion of the population's quality of life has not improved despite a high GDP growth rate has led some people to believe that the GDP measure should be expanded to consider human well-being and life quality. The Human Development Index (HDI), a multidimensional indicator of development, has proven to be more reasonable in comparison to the measure of GDP growth, which is one-dimensional in income. This is in line with the general belief that well-being is a multidimensional concept that cannot be measured by market production or GDP alone, so the value of all goods produced in a nation during a fiscal year is used to define its GDP. It is discovered to be one of the economic growth and production indicators and to play a crucial strategic role in employment, development, and the balance of payments.

Table 1: The Bias and MSE of α and β for samples without outlier.

n	Parameter	Statistic	Classical Procedures	Proposed Procedures						
			MLE	MOM	Wald-type	Procedure 1 (Repetitive)	Procedure 2 (Iterative)			
					r = 2	r = 3	r = 3	r = 2	r = 3	r = 4
50	alpha	Bias	-0.0001	-0.0099	-0.0019	-0.0018	0.4937	0.0007	0.0019	-0.0023
		MSE	0.0619	0.0757	0.0154	0.0238	0.268	0.0149	0.0312	0.0313
	beta	Bias	-0.4708	0.1866	-0.5069	-0.4964	-0.4975	-0.5105	-0.5004	-0.4994
		MSE	0.286	0.5442	0.2688	0.2552	0.2573	0.2701	0.2723	0.2735
100	alpha	Bias	-0.0008	0.0078	0.0052	0.0001	0.5007	-0.0041	-0.0006	-0.0008
		MSE	0.0154	0.0392	0.0075	0.0121	0.2641	0.0077	0.0152	0.0153
	beta	Bias	-0.4937	0.0845	-0.5001	-0.5041	-0.4999	-0.5025	-0.4984	-0.5008
		MSE	0.2524	0.4401	0.2563	0.2546	0.2533	0.2571	0.2585	0.2609
200	alpha	Bias	0.0008	-0.0003	-0.0006	-0.0009	-0.0004	-0.0005	0.0001	-0.0007
		MSE	0.0076	0.0258	0.0077	0.0114	0.0075	0.0075	0.0076	0.0075
	beta	Bias	-0.496	0.0692	-0.499	-0.5003	-0.4992	-0.5001	0.4992	-0.5004
		MSE	0.25	0.3875	0.255	0.2532	0.2522	0.2548	0.2541	0.2532
500	alpha	Bias	-0.0005	-0.0033	0.0001	-0.0002	0.0005	-0.0001	-0.0003	-0.0002
		MSE	0.003	0.0121	0.0031	0.0045	0.0031	0.003	0.003	0.003
	beta	Bias	-0.4978	0.2379	-0.4996	-0.5001	-0.5002	-0.4992	-0.4996	-0.5002
		MSE	0.2494	0.3335	0.2519	0.2521	0.2501	0.2511	0.2510	0.2500

Table 2: The Bias and MSE for α and β when σ_δ^2 with outliers in x.

n	Parameter	Statistic	Classical Procedures	Proposed Procedures						
			MLE	MOM	Wald-type	Procedure 1 (Repetitive)	Procedure 2			
(Iterative)										
					r = 2	r = 3	r = 3	r = 2	r = 3	r = 4
50	alpha	Bias	0.0167	-0.0213	-0.0057	-0.0234	0.5546	-0.0048	0.002	0.0006
		MSE	0.4811	0.5086	0.0204	0.0349	0.3367	0.0214	0.0678	0.0699
	beta	Bias	-0.5187	-0.3391	-0.5445	-0.5517	-0.5638	-0.5507	-0.5	-0.4976
		MSE	0.6769	0.5464	0.3284	0.3149	0.3071	0.3130	0.2906	0.2915
100	alpha	Bias	0.0027	-0.0029	0.0008	0.0066	0.5322	-0.004	-0.0	0.0031
		MSE	0.0881	0.027	0.0086	0.0133	0.2958	0.008	0.0252	0.026
	beta	Bias	-0.5055	-0.2506	-0.5241	-0.5248	-0.5300	-0.5256	-0.4999	-0.5036
		MSE	0.4165	0.4123	0.2807	0.2894	0.2754	0.2812	0.2664	0.2699
200	alpha	Bias	0.0027	-0.0118	-0.0011	0.0014	-0.0003	-0.0007	0.0009	-0.001
		MSE	0.0881	0.0718	0.0147	0.0274	0.0146	0.0147	0.01	0.0103
	beta	Bias	-0.5055	-0.2369	-0.5276	-0.5297	-0.5291	-0.5268	-0.5007	-0.4989
		MSE	0.4113	0.3987	0.2889	0.291	0.2696	0.2771	0.2575	0.2554
500	alpha	Bias	-0.0022	-0.0118	0.0004	-0.0004	-0.0006	-0.0014	0.0004	0.0004
		MSE	0.0044	0.0131	0.0042	0.0072	0.0043	0.0043	0.0036	0.0034
	beta	Bias	-0.5325	-0.1784	-0.5111	-0.5118	-0.5119	-0.5113	-0.5001	-0.4996

Table 3: The Bias and MSE for α and β when σ_ε^2 with outliers in y.

n	Parameter	Statistic	Classical Procedures	Proposed Procedures						
			MLE	MOM	Wald-type	Procedure 1 (Repetitive)	Procedure 2			
(Iterative)										
					r = 2	r = 3	r = 3	r = 2	r = 3	r = 4
50	α	Bias	-0.6594	0.0265	-0.007	0.0079	0.4995	0.0015	-0.0023	0.0035
		MSE	0.7014	0.1987	0.0385	0.0577	0.3420	0.0404	0.1373	0.1365
	β	Bias	0.1081	0.2705	-0.5076	-0.4936	-0.4955	-0.4955	-0.5056	-0.5009
		MSE	0.959	0.5318	0.288	0.2663	0.2731	0.2721	0.212	0.2823
100	α	Bias	0.0349	-0.0433	-0.0085	-0.0083	0.4987	0.0074	0.0008	-0.0003
		MSE	0.7663	0.0683	0.0135	0.019	0.2783	0.0146	0.0409	0.0412
	β	Bias	0.7394	0.6033	-0.4999	-0.5087	-0.5007	-0.496	-0.5013	-0.4987
		MSE	0.8381	0.4528	0.2611	0.2577	0.2594	0.2552	0.2777	0.2755
200	α	Bias	-0.0011	-0.0238	0.0002	-0.0004	0.0004	0.0002	0.0002	-0.0004
		MSE	0.036	0.0397	0.0088	0.0072	0.0088	0.0088	0.0142	0.014
	β	Bias	0.0481	0.0766	-0.4724	-0.5118	-0.4732	-0.4749	-0.5003	-0.4989
		MSE	0.4445	0.3763	0.2289	0.2647	0.2587	0.2502	0.2500	0.2461
500	α	Bias	0.0017	-0.0229	-0.0002	-0.0005	-0.0012	0.0003	-0.0008	0.0005
		MSE	0.0057	0.0322	0.0033	0.0051	0.0033	0.0033	0.004	0.004
	β	Bias	-0.3289	0.0643	-0.4896	-0.4883	-0.4885	-0.4884	-0.5008	-0.4996

Table 4: The Bias and MSE for α and β when $(\sigma_\delta^2, \sigma_\varepsilon^2) = (16, 16)$ with outliers in both (x, y).

n	Parameter	Statistic	Classical Procedures	Proposed Procedures						
			MLE	MOM	Wald-type	Procedure 1 (Repetitive)	Procedure 2 (Iterative)			
					r = 2	r = 3	r = 3	r = 2	r = 3	r = 4
50	α	Bias	0.0415	-0.0059	-0.0005	-0.0048	0.5307	0.0076	0.0033	0.0002
		MSE	.9337	0.1172	0.0421	0.0723	0.3509	0.0381	0.1338	0.1405
	β	Bias	-0.2438	-0.292	-0.5592	-0.5425	-0.5219	-0.5496	-0.5024	-0.4986
		MSE	0.9197	0.5558	0.3403	0.3232	0.3213	0.3076	0.3411	0.331
100	α	Bias	0.0021	0.0014	0.0095	-0.0038	0.5286	-0.0125	0.0028	-0.0014
		MSE	0.341	0.0391	0.0141	0.0277	0.3065	0.0146	0.0444	0.0461
	β	Bias	-0.1181	-0.2787	-0.5208	-0.5245	-0.5303	-0.5247	-0.5036	-0.4968
		MSE	0.3138	0.7082	0.284	0.286	0.2908	0.2816	0.2853	0.2778
200	α	Bias	0.0087	0.0003	-0.0007	0.0008	-0.0001	0.0005	0.0008	-0.0003
		MSE	0.1106	0.0154	0.0159	0.0304	0.0161	0.0162	0.0161	0.0161
	β	Bias	-0.4506	-0.2296	-0.5003	-0.5004	-0.5002	-0.5019	-0.5014	-0.4988
		MSE	0.3722	0.4959	0.262	0.2625	0.2627	0.2623	0.2625	0.2599
500	α	Bias	-0.0011	0.0004	0.0001	0.0008	0.0014	-0.0004	0.0004	-0.0007
		MSE	0.0104	0.0049	0.0044	0.0076	0.0044	0.0045	0.0045	0.0044
	β	Bias	-0.4496	-0.1748	-0.4996	-0.5001	-0.5005	-0.5003	-0.4999	-0.4998
		MSE	0.3407	0.4764	0.253	0.2532	0.2536	0.2534	0.253	0.2529

To show the application of the proposed SMEM, two different datasets were collected to interpret the relationship between GDP and HDI. The data were from Jordan's economic report for 1990-2019 (The World Bank. Jordan — Data (worldbank.org), Country Economy. Jordan - Human Development Index - HDI 2019 — countryeconomy.com) and presented in Table 5.
Table 6 reports the descriptive analysis of the data. It indicates that the smallest GDP was in 1991 (GDP = 1155.2), while the largest was in 2019 (GDP = 4405.487). Meanwhile, the lowest HDI is 0.625, as reported in 1990, and the highest is in 2008 which is reported at 0.745. Overall, the mean GDP of Jordan is 2618.4 (SD = 1207.3) and the mean HDI is 0.709 (SD = 0.035). It is worth noting that GDP and HDI have a strong positive significant correlation in Jordan ($r = 0.741$, $p < 0.001$), and the trend of both variables within the study period are given in Figure 1 and 2.

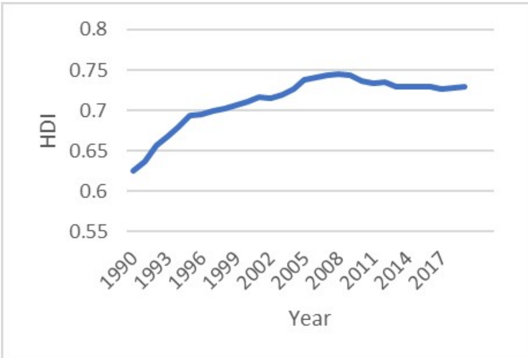

Figure 3: The trend of the Jordanian HDI within 1990-2019

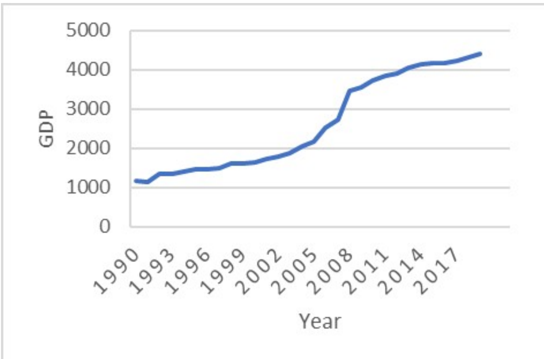

Figure 4: The trend of the National GDP within 1990-2019

Moreover, the scatter plot (Figure 5) suggests the type of the relationships to be almost linear.

GDP and HDI can be written as a linear relationship model. However, it is believed

Table 5: Yearly Dataset of HDI and GDP of Jordan (1990-2021)

Year	HDI	GDP
1990	0.625	1166.611
1991	0.636	1155.234
1992	0.657	1335.288
1993	0.668	1334.229
1994	0.679	1414.339
1995	0.693	1466.045
1996	0.695	1463.888
1997	0.699	1494.511
1998	0.702	1600.398
1999	0.706	1619.536
2000	0.711	1651.622
2001	0.717	1720.361
2002	0.715	1802.055
2003	0.72	1876.259
2004	0.726	2044.964
2005	0.738	2183.395
2006	0.741	2513.029
2007	0.744	2735.379
2008	0.745	3455.77
2009	0.743	3559.692
2010	0.737	3736.645
2011	0.734	3852.89
2012	0.735	3910.347
2013	0.729	4044.427
2014	0.729	4131.447
2015	0.73	4164.109
2016	0.729	4175.357
2017	0.726	4231.518
2018	0.728	4308.151
2019	0.729	4405.487

Table 6: Descriptive Statistics

Variable	Min	Max	Mean	SD	Correlation	P.
GDP	1155.2	4405.5	2618.4	1207.3	0.741	¡ 0.001
HDI	0.625	0.745	0.709	0.035		

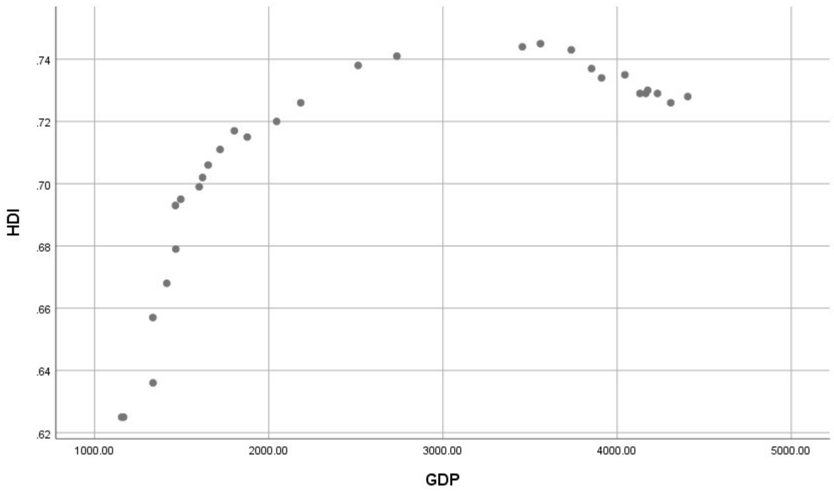

Figure 5: The scatter plot of HDI and GDP

that both variables are subject to error because several sub-factors determine the final value of each of them. As a result, MEM is the best model to study the relationship between HDI and GDP, which can be rewritten as

$HDI = \alpha + \beta \times (GDP - \delta) + \epsilon$

Accordingly, Table 7 shows the results of the parameter estimation according to the three methods discussed previously in this article. The results indicate, based on the mean square error (MSE), that the proposed procedures with r = 3 and r = 4 gave more accurate estimators than the other estimation methods. The outcomes are shown by the residual line, such that the line that is nearest to zero is the best-fit line. The residual plots in Figure 6 indicate that the proposed procedures (yellow, green, and gray lines) are better than the classical procedures (light blue and orange lines).

7 Concluding Remarks

To address the problems in estimating the SMEM, this study proposed a new nonparametric estimation procedure. In the new procedure, an iterative Wald-type estimation method was used. For medium or large sample size data, Monte Carlo simulations extend strong evidence of the prevalence of the proposed estimation procedures over

Table 7: Parameter Estimation of HDI vs GDP

Procedure	Method	Criterion	α	β	MSR
Classical	Wald-type	r = 2	1.77673 $\times$10-5	0.66	0.008475
		r = 3	1.90293 $\times$10-5	0.756	0.009427
	MLE		2.32E-05	0.647	0.019723
	MOM		2.22E-05	0.637	0.018722
Proposed	Procedure.1	r = 3	2.06822 $\times$10-5	0.655	0.000692
		r = 4	2.715 $\times$10-5	0.639	0.000442
	Procedure.2	r = 3	1.434 $\times$10-5	0.672	0.00049
		r = 4	1.748 $\times$10-5	0.666	0.000462

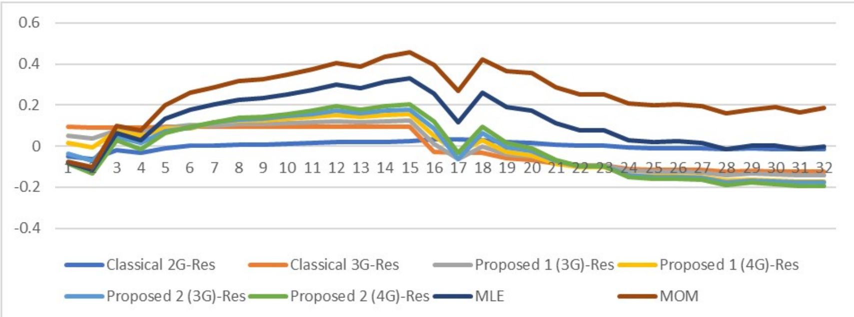

Figure 6: Residual Comparisons of the estimation methods

classical methods. Moreover, an estimation method was applied to real data to study the effect of GDP on HDI. According to the data analysis, there is a strong positive relationship between the two variables. The optimal r value of the proposed procedure will be determined in future work.

8 Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgment

The authors would like to thank the editorial team and the reviewers for their inputs and value comments which improve the contents of this article.

References

Al-Nasser, A. D., Al-Sliti, A., and Edous, M. (2016). New iterative am estimation procedure for fitting the simple linear measurement error models. *Electronic Journal of Applied Statistical Analysis*, 9(3):491–501.

Bartlett, M. S. (1949). Fitting a straight line when both variables are subject to error. *Biometrics*, 5(3):207–212.

Carroll, R. J., Ruppert, D., and Stefanski, L. A. (1995). *Measurement error in nonlinear models*, volume 105. CRC press.

Casella, G. and Berger, R. L. (1990). *Statistical Inference*. Wadsworth and Brooks.

Cheng, C. L. and Van Ness, J. W. (1999). *Statistical Regression with Measurement Error*. London: Arnold.

Fuller, W. A. (1987). *Measurement error models*. New York: Wiley.

Gillard, J. (2010). An overview of linear structural models in errors in variables regression. *REVSTAT-Statistical Journal*, 8(1):57–80.

Gillard, J. and Iles, T. (2005). Method of moments estimation in linear regression with errors in both variables. *Cardiff University School of Mathematics Technical Paper*, pages 3–43.

Gillard, J. W. (2006). An historical overview of linear regression with errors in both variables. *Cardiff University School of Mathematics Technical Report*.

Griliches, Z. (1974). Errors in variables and other unobservables. *Econometrica: Journal of the Econometric Society*, pages 971–998.

Gupta, Y. P. and Amanull, A. (1970). note on the moments of the wald's estimator. *STATISTICA NEERLANDICA*, 24.

Gustafson, P. (2021). *Handbook of measurement error models*, volume 16. CRC Press.

Lindley, D. V. (1947). Regression lines and the linear functional relationship. *Supplement to the Journal of the Royal Statistical Society*, 9(2):218–244.

Madansky, A. (1959). The fitting of straight lines when both variables are subject to error. *Journal of the american statistical association*, 54(285):173–205.

Nair, K. and Shrivastava, M. (1942). On a simple method of curve fitting. *Sankhyā: The Indian Journal of Statistics*, pages 121–132.

Stefanski, L. A. (2000). Measurement error models. *Journal of the American Statistical Association*, 95(452):1353–1358.

Wald, A. (1940). The fitting of straight lines if both variables are subject to error. *The annals of mathematical statistics*, 11(3):284–300.