Journal of the Korean
Statistical Society
Vol. 23, No. 1, 1994

Estimation of the Polynomial Errors-in-variables Model with Decreasing Error Variances†

Myung-Sang Moon [1] and R. F. Gunst [2]

ABSTRACT

Polynomial errors-in-variables model with one predictor variable and one response variable is defined and an estimator of model is derived following the Booth's linear model estimation procedure. Since polynomial model is nonlinear function of the unknown regression coefficients and error-free predictors, it is nonlinear model in errors-in-variables model. As a result of applying linear model estimation method to nonlinear model, some additional assumptions are necessary. Hence, an estimator is derived under the assumption that the error variances are decreasing as sample size increases. Asymptotic properties of the derived estimator are provided. A simulation study is presented to compare the small sample properties of the derived estimator with those of OLS estimator.

KEYWORDS: Errors-in-variables, Polynomial model, Error-free predictors, Decreasing error variance, OLS estimator.

[1] Department of Statistics, Yonsei University, Wonju-Kun, Kangwon-Do , 222-701, Korea.
[2] Department of Statistical Science, Southern Methodist University, Dallas, TX, 75275, USA.
† This Research was supported by the 1993 Yonsei Maeji Research Fund

1. INTRODUCTION

In traditional regression setting in which a response variable is modeled by one or more predictor variables, non-stochastic and error-free predictor variables are assumed. When predictor variables are subject to error, the usual least squares parameter estimators (OLS estimators) used with traditional regression models are biased and inconsistent. Major research efforts since the early 1870's have been directed towards finding alternative estimators that have desirable properties.

Polynomial errors-in-variables model is a nonlinear function of the unknown regression coefficients and the (unknown) error-free predictiors. Generalized LS estimation, nonlinear errors-in-variables model estimation, and approaches based on score functions are alternatives that can be used to fit this model. The focus of this work is on the polynomial functional relationship in which jointly normally distributed measurement errors with completely known covariance matrices are assumed throughout. A modified generalized LS estimator of the polynomial errors-in-variables model is derived under the assumption that the error variances decrease with increasing sample size. Booth(1973) develops a pseudo maximum likelihood estimator of the linear errors-in-variables model under the assumption that the covariance matrices of the error vectors are unequal but completely known. Although the developments presented in this work are similar to those given in Booth, the assumptions made are different. First, the polynomial model considered in this work is a nonlinear model. Second, decreasing error variances are assumed. The third difference will be described in Section 2 after introducing necessary notation.

The polynomial errors-in-variables model and assumptions that are necessary in subsequent sections are given in Section 2. A modified generalized LS estimator is derived in Section 3 and properties of that estimator are introduced in Section 4. Simulation results comparing the derived estimator and OLS estimator are presented in Section 5.

2. POLYNOMIAL FUNCTIONAL RELATIONSHIP

Let $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$ be sequences of positive real numbers such that $n = a_n b_n$ for $n = 1, 2, \cdots, \infty$. Assume the existence of a sequence of experiments indexed by n, and let b_n denote the number of observations in the n-th experiment. Thus, $(y_{ni},\ x_{ni} : i = 1, 2, \cdots, b_n)$ are observed in the n-th experiment, where

$$y_{ni} = \psi_i + v_{ni}, \qquad x_{ni} = \pi_i + u_{ni}.$$

The random variables $(v_{ni},\ u_{ni})$ denote errors of measurment obtained in observing unknown error-free $(\psi_i,\ \pi_i)$. To condense the notation, the subscript n will be suppressed. The k-th order polynomial functional relationship in one predictor is given by the model (2.1). Note that in this work, bold-face letters denote vectors or matrices and all vectors are column vectors.

$$\begin{aligned}\psi_i &= \beta_0 + \beta_1\pi_i + \beta_2\pi_i^2 + \ldots + \beta_k\pi_i^k \\ &= \boldsymbol{\pi}_i'\boldsymbol{\beta}, \qquad i = 1, 2, \cdots, b_n, \end{aligned} \tag{2.1}$$

where $\boldsymbol{\pi}_i = (1,\ \pi_i,\ \pi_i^2,\ \cdots,\ \pi_i^k)'$ and $\boldsymbol{\beta} = (\beta_0,\ \beta_1,\ \beta_2,\ \cdots,\ \beta_k)'$. Let $\boldsymbol{z}_i = (y_i,\ x_i)'$; i.e., $\boldsymbol{z}_i = \boldsymbol{\xi}_i + \boldsymbol{w}_i$ with $\boldsymbol{\xi}_i = (\psi_i, \pi_i)'$ denoting the vector of error-free variates and $\boldsymbol{w}_i = (v_i, u_i)'$ the vector of measurement errors. The vector of measurement errors $\boldsymbol{w}_i$ are assumed to be i.i.d. N(0, $\sum_{ww}$) with completely known covariance matrix

$$\textstyle\sum_{ww} = \begin{bmatrix} \sigma_{vv} & \sigma_{vu} \\ \sigma_{uv} & \sigma_{uu} \end{bmatrix} \tag{2.2}$$

For further development, a property of Hermite polynomials in normal variates is needed. Let $H_m(Z)$ be the m-th Hermite polynomial : $H_0(Z) = 1$, $H_1(Z) = Z$, and $H_m(Z) = ZH_{m-1}(Z) - (m-1)H_{m-2}(Z)$. Define $P_m(Z) = \sigma^m H_m\left(\frac{Z}{\sigma}\right)$, where σ^2 is the variance of Z. If Z is normally distributed with mean μ and variance σ^2, then (Stulajter 1978)

$$E\{P_m(Z)\} = \mu^m. \tag{2.3}$$

The following definitions are used frequently in subsequent sections.

Definition 2.1. Define the $(k+1) \times 1$ polynomial vector $\boldsymbol{p}_i$ by

$$\boldsymbol{p}_i = \boldsymbol{\pi}_i + \boldsymbol{f}_i = (1,\ P_1(x_i),\ P_2(x_i),\ \cdots,\ P_k(x_i))',$$

where $\boldsymbol{f}_i = (0,\ P_1(x_i) - \pi_i,\ P_2(x_i) - \pi_i^2, \cdots, P_k(x_i) - \pi_i^k)'$.

Define the $(k+2) \times 1$ vectors $\boldsymbol{\zeta}_i, \boldsymbol{r}_i, \boldsymbol{g}_i$ and $\boldsymbol{\theta}$ by

$$\boldsymbol{\zeta}_i = (\psi_i, \boldsymbol{\pi}_i')',\ \boldsymbol{r}_i = (v_i, \boldsymbol{f}_i')',\ \boldsymbol{g}_i = \boldsymbol{\zeta}_i + \boldsymbol{r}_i = (y_i, \boldsymbol{p}_i')' \text{ and } \boldsymbol{\theta} = (1, -\boldsymbol{\beta}')'. \quad \square$$

Definition 2.2. Define $b_n \times (k+2)$ matrix $\boldsymbol{G}$ by

$$\boldsymbol{G}' = [\boldsymbol{g}_1\ \ \boldsymbol{g}_2\ \ \cdots\ \ \boldsymbol{g}_{b_n}],$$

and define $b_n \times (k+1)$ matrices $\boldsymbol{T}, \boldsymbol{\Pi}$, and $\boldsymbol{F}$ by

$$\boldsymbol{T}' = [\,\boldsymbol{p}_1\ \ \boldsymbol{p}_2\ \ \cdots\ \ \boldsymbol{p}_{b_n}\,],\quad \boldsymbol{\Pi}' = [\,\boldsymbol{\pi}_1\ \ \boldsymbol{\pi}_2\ \ \cdots\ \ \boldsymbol{\pi}_{b_n}\,] \text{ and}$$

$$\boldsymbol{F}' = [\,\boldsymbol{f}_1\ \ \boldsymbol{f}_2\ \ \cdots\ \ \boldsymbol{f}_{b_n}\,]\text{ , respectively,}$$

where $\boldsymbol{g}_i$, $\boldsymbol{p}_i$, $\boldsymbol{\pi}_i$ and $\boldsymbol{f}_i$ are given in Definition 2.1. Finally, define $b_n \times 1$ vector $\boldsymbol{y}' = (y_1, y_2, \cdots, y_{b_n})$. $\square$

Note that $\boldsymbol{p}_i$ denotes a vector of k-th order Hermit polynomials in the observable predictor x_i and $\boldsymbol{f}_i$ denotes the vector of deviations from the corresponding powers in the error-free predictor. Also, observe that $\boldsymbol{\zeta}_i$ is the vector of error-free response and polynomial predictor variates, $\boldsymbol{g}_i$ is the corresponding vector of observable variates, and $E(\boldsymbol{r}_i) = \mathbf{o}$ by (2.3). From Definition 2.1,

$$y_i = \boldsymbol{\pi}_i'\boldsymbol{\beta} + v_i = \boldsymbol{p}_i'\boldsymbol{\beta} + e_i,$$

where $e_i = v_i - \boldsymbol{f}_i'\boldsymbol{\beta}$. This is traditional errors-in-variables model in which the vector of Hermite polynomials $\boldsymbol{p}_i$ is a vector of observed predictor variables, $\boldsymbol{\pi}_i$ is the vector of error-free predictors, and $\boldsymbol{f}_i$ is its measurement error vector. Therefore, $\boldsymbol{r}_i$ is the vector of measurement errors for the observable variates $\boldsymbol{g}_i$. Utilizing the measurement error vector $\boldsymbol{r}_i$ and the vectors in Definition 2.1, model (2.1) can be rewritten in the form

$$\boldsymbol{g}_i = \boldsymbol{\zeta}_i + \boldsymbol{r}_i, \quad \boldsymbol{\zeta}_i'\boldsymbol{\theta} = 0, \quad i = 1, 2, \cdots, b_n. \tag{2.4}$$

The symmetric $(k+2) \times (k+2)$ covariance matrix of the i-th measurement error vector $\boldsymbol{r}_i$ is

$$\boldsymbol{\Omega}_i = Cov(\boldsymbol{r}_i)$$

$$= \begin{bmatrix} \sigma_{vv} & 0 & Cov(v_i, P_1(x_i) - \pi_i) & \cdots & Cov(v_i, P_k(x_i) - \pi_i^k) \\ & 0 & 0 & \cdots & 0 \\ Symmetric & & Cov\{P_j(x_i) - \pi_i^j, P_m(x_i) - \pi_i^m\} & & \end{bmatrix}. \quad (2.5)$$

The elements of the covariance matrix $\boldsymbol{\Omega}_i$ can be computed using the following expressions.

(i) $Cov\{P_l(x_i) - \pi_i^l, v_i\} = l\pi_i^{l-1}\sigma_{uv}, \quad l = 1, 2, \cdots, k,$

(ii) $E\{P_{2m}(x_i)P_{2n}(x_i)\}$

$$= \sum_{c=0}^{m+n}\sum_{j=c}^{m+n} d_j \frac{(2j)!}{(2c)!\,(j-c)!\,2^{m+n-c}}\sigma_{uu}^{m+n-c}\pi_i^{2c},$$

where $d_j = \sum_{t=t_1}^{t_2}(-1)^{m+n-j}\dfrac{(2m)!\,(2n)!}{(2t)!\,(m-t)!\{2(j-t)\}!\,(n-j+t)!}$

with $t_1 = \max\{0, j-n\}$ and $t_2 = \min\{j, m\}$.

(iii) $E\{P_{2m-1}(x_i)P_{2n-1}(x_i)\}$

$$= \sum_{c=0}^{m+n-1}\sum_{j=c+1}^{m+n} d_j \frac{\{2(j-1)\}!}{(2c)!\,(j-c-1)!\,2^{m+n-c-1}}\sigma_{uu}^{m+n-c-1}\pi_i^{2c},$$

where $d_j = \sum_{t=t_1}^{t_2}(-1)^{m+n-j}\dfrac{(2m-1)!\,(2n-1)!}{(2t-1)!\,(m-t)!\{2(j-t)-1\}!\,(n-j+t)!}$

with $t_1 = \max\{1, j-n\}$ and $t_2 = \min\{j-1, m\}$.

(iv) $E\{P_{2m-1}(x_i)P_{2n}(x_i)\}$

$$= \sum_{c=1}^{m+n}\sum_{j=c}^{m+n} d_j \frac{(2j-1)!}{(2c-1)!\,(j-c)!\,2^{m+n-c}}\sigma_{uu}^{m+n-c}\pi_i^{2c-1},$$

where $d_j = \sum_{t=t_1}^{t_2} (-1)^{m+n-j} \frac{(2m-1)!\,(2n)!}{(2t-1)!\,(m-t)!\,\{2(j-t)\}!\,(n-j+t)!}$ with $t_1 = \max\{1, j-n\}$ and $t_2 = \min\{j, m\}$.

Since $\boldsymbol{\Omega}_i$'s contain powers of the π_i, they are unequal and unknown although $\boldsymbol{\Sigma}_{\mathbf{ww}}$ is known, which comprises the third different assumption from that of Booth. The construction of $\tilde{\boldsymbol{\theta}}$ is based on the assumption that $a_n^{-1} = o(n^{-1/3})$ and there is a preliminary estimator of $\boldsymbol{\theta}$, say $\hat{\boldsymbol{\theta}}$, satisfying $\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} = O_p\{\max(a_n^{-1}, n^{-1/2})\}$. For general nonlinear models, Wolter & Fuller (1982b) proved that $\hat{\boldsymbol{\beta}}_{OLS} - \boldsymbol{\beta} = O_p\{\max(a_n^{-1}, n^{-1/2})\}$ when $a_n^{-1} = o(n^{-1/3})$, where $\hat{\boldsymbol{\beta}}_{OLS}$ is OLS estimator of $\boldsymbol{\beta}$. It is also assumed that an estimator of $\boldsymbol{\Omega}_i$, say $\hat{\boldsymbol{\Omega}}_i$, is available satisfying $\hat{\boldsymbol{\Omega}}_i = \boldsymbol{\Omega}_i + O_p(a_n^{-3/2})$. An unbiased estimator of $\boldsymbol{\Omega}_i$ is obtained by replacing π_i^j with $P_j(x_i)$, $j = 1, 2, \cdots, 2(k-1)$, by (2.3). It is shown is Section 4 that an unbiased estimator of $\boldsymbol{\Omega}_i$ satisfies the mentioned assumption. In the next section, the derivation of $\tilde{\boldsymbol{\theta}}$ is detailed and it is demonstrated in Theorem 4.1 that $n^{1/2}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta})$ converges in distribution to a normal random variable.

3. ESTIMATOR

Suppose that model (2.1) and (2.2) hold. Utilizing $\boldsymbol{r}_i$ as the measurement error vector component of the observable vector $\boldsymbol{g}_i' = (y_i, \boldsymbol{p}_i')$, the generalized LS estimator of $\boldsymbol{\beta}$ minimizes the weighted sum of squares

$$\sum_{i=1}^{b_n} \boldsymbol{r}_i^{(2)\prime} \boldsymbol{\Omega}_i^{(2)^{-1}} \boldsymbol{r}_i^{(2)} \qquad \text{subject to } \boldsymbol{\zeta}_i'\boldsymbol{\theta} = 0, \tag{3.1}$$

where $\boldsymbol{r}_i^{(2)}$ denotes $\boldsymbol{r}_i$ with second element deleted and $\boldsymbol{\Omega}_i^{(2)}$ denotes $\boldsymbol{\Omega}_i$ with second row and second column deleted. This notation is used so zero row and zero column corresponding to the constant term(β_0) can be removed, thereby making $\boldsymbol{\Omega}_i^{(2)}$ nonsingular.

Let $\boldsymbol{g}_i^{(2)}, \boldsymbol{\zeta}_i^{(2)}$ and $\boldsymbol{\theta}^{(2)}$ denote $\boldsymbol{g}_i, \boldsymbol{\zeta}_i$ and $\boldsymbol{\theta}$ respectively with second element deleted as before. Using Lagrangian multipliers, minimizing (3.1) is equivalent to minimizing

$$Q = \sum_{i=1}^{b_n} (\boldsymbol{g}_i^{(2)'} \boldsymbol{\Omega}_i^{(2)^{-1}} \boldsymbol{g}_i^{(2)} - \boldsymbol{g}_i^{(2)'} \boldsymbol{\Omega}_i^{(2)^{-1}} \boldsymbol{\zeta}_i^{(2)} - \boldsymbol{\zeta}_i^{(2)'} \boldsymbol{\Omega}_i^{(2)^{-1}} \boldsymbol{g}_i^{(2)} + \boldsymbol{\zeta}_i^{(2)'} \boldsymbol{\Omega}_i^{(2)^{-1}} \boldsymbol{\zeta}_i^{(2)})$$
$$-2\sum_{i=1}^{b_n} \alpha_i^* (\boldsymbol{\zeta}_i^{(2)'} \boldsymbol{\theta}^{(2)} + \beta_0) \qquad (3.2)$$

with respect to $\boldsymbol{\zeta}_i^{(2)}, \boldsymbol{\theta}$ and α_i^*. Differentiating (3.2) with respect to $\boldsymbol{\zeta}_i^{(2)}$ results in the system of equations

$$\frac{\partial Q}{\partial \boldsymbol{\zeta}_i^{(2)}} = 2\boldsymbol{\Omega}_i^{(2)^{-1}} (\boldsymbol{\zeta}_i^{(2)} - \boldsymbol{g}_i^{(2)}) - 2\alpha_i^* \boldsymbol{\theta}^{(2)}, \quad i = 1, 2, \cdots, b_n, \qquad (3.3)$$

which gives

$$\alpha_i^* = \frac{-\beta_0 - \boldsymbol{\theta}^{(2)'} \boldsymbol{g}_i^{(2)}}{\boldsymbol{\theta}^{(2)'} \boldsymbol{\Omega}_i^{(2)} \boldsymbol{\theta}^{(2)}}, \qquad (3.4)$$

using $\boldsymbol{\zeta}_i' \boldsymbol{\theta} = 0$. Substituting (3.4) into (3.3) and setting equal to zero yields

$$\begin{aligned} \boldsymbol{r}_i^{(2)'} \boldsymbol{\Omega}_i^{(2)^{-1}} \boldsymbol{r}_i^{(2)} &= -\frac{1}{\boldsymbol{\theta}^{(2)'} \boldsymbol{\Omega}_i^{(2)} \boldsymbol{\theta}^{(2)}} \{\beta_0 (\boldsymbol{\zeta}_i - \boldsymbol{g}_i)' \boldsymbol{\theta} + (\boldsymbol{\theta}' \boldsymbol{g}_i - \beta_0)(\boldsymbol{\zeta}_i - \boldsymbol{g}_i)' \boldsymbol{\theta}\} \\ &= \frac{\boldsymbol{\theta}' \boldsymbol{g}_i \boldsymbol{g}_i' \boldsymbol{\theta}}{\boldsymbol{\theta}' \boldsymbol{\Omega}_i \boldsymbol{\theta}}, \end{aligned}$$

since the second element of $\boldsymbol{\zeta}_i - \boldsymbol{g}_i$ is 0 and $\boldsymbol{\zeta}_i' \boldsymbol{\theta} = 0$. Thus, by defining $\boldsymbol{M}_i = \boldsymbol{g}_i \boldsymbol{g}_i'$, the quantity to be minimized is

$$\sum_{i=1}^{b_n} \boldsymbol{r}_i^{(2)'} \boldsymbol{\Omega}_i^{(2)^{-1}} \boldsymbol{r}_i^{(2)} = \sum_{i=1}^{b_n} \frac{\boldsymbol{\theta}' \boldsymbol{M}_i \boldsymbol{\theta}}{\boldsymbol{\theta}' \boldsymbol{\Omega}_i \boldsymbol{\theta}}. \qquad (3.5)$$

By taking the derivative of (3.5) with respect to $\boldsymbol{\theta}$ and setting the result equal to zero, the value of $\boldsymbol{\theta}$, say $\boldsymbol{\theta}^0$, which minimizes (3.5) subject to $\boldsymbol{\zeta}_i' \boldsymbol{\theta} = 0$ satisfies

$$\sum_{i=1}^{b_n} \frac{(\boldsymbol{M}_i - \sigma_i \boldsymbol{\Omega}_i) \boldsymbol{\theta}^0}{\sigma_{e_i}^2} = \mathbf{o}, \qquad (3.6)$$

where $\alpha_i = \dfrac{\boldsymbol{\theta}^{0'} \boldsymbol{M}_i \boldsymbol{\theta}^0}{\sigma_{e_i}^2} \geq 0$ and $\sigma_{e_i}^2 = Var(e_i) = \boldsymbol{\theta}^{0'} \boldsymbol{\Omega}_i \boldsymbol{\theta}^0 > 0$. Since $\sigma_{e_i}^2$ and α_i involve $\boldsymbol{\theta}^0$, the solution to (3.6) is difficult to obtain. Therefore, an alternative estimator of $\boldsymbol{\theta}$ for which asymptotic properties can be obtained is needed.

Define $\hat{\sigma}^2_{e_i} = \hat{\boldsymbol{\theta}}'\hat{\boldsymbol{\Omega}}_i\hat{\boldsymbol{\theta}}$ and $\hat{\alpha}_i = \dfrac{\hat{\boldsymbol{\theta}}'\boldsymbol{M}_i\hat{\boldsymbol{\theta}}}{\hat{\sigma}^2_{e_i}}$, where $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\Omega}}_i$ are estimators of $\boldsymbol{\theta}$ and $\boldsymbol{\Omega}_i$ satisfying assumptions mentioned in the previous section. If $\sigma^2_{e_i}, \boldsymbol{\Omega}_i$ and α_i in (3.6) are replaced by $\hat{\sigma}^2_{e_i}, \hat{\boldsymbol{\Omega}}_i$ and $\hat{\alpha}_i$ respectively, then the relationship

$$\sum_{i=1}^{b_n} \frac{(\boldsymbol{M}_i - \hat{\alpha}_i\hat{\boldsymbol{\Omega}}_i)\tilde{\boldsymbol{\theta}}}{\hat{\sigma}^2_{e_i}} = \mathbf{o}, \tag{3.7}$$

can be used to determine $\tilde{\boldsymbol{\beta}}$. Note that $\tilde{\boldsymbol{\theta}}$ will not in general equal to $\boldsymbol{\theta}^0$. For further developments, define $b_n \times b_n$ diagonal matrix $\boldsymbol{W} = \mathtt{diag}\{\sigma^2_{e_i}\}$ and let $\hat{\boldsymbol{W}} = \mathtt{diag}\{\hat{\sigma}^2_{e_i}\}$. By Definition 2.2, (3.7) can be expressed as

$$(\frac{1}{b_n}\boldsymbol{G}'\hat{\boldsymbol{W}}^{-1}\boldsymbol{G} - \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{1}{\hat{\sigma}^2_{e_i}}\hat{\alpha}_i\hat{\boldsymbol{\Omega}}_i)\tilde{\boldsymbol{\theta}} = \mathbf{o}. \tag{3.8}$$

Thus, using partitioned matrices and vectors in (3.8),

$$\tilde{\boldsymbol{\beta}} = (\frac{1}{b_n}\boldsymbol{T}'\hat{\boldsymbol{W}}^{-1}\boldsymbol{T} - \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\hat{\alpha}_i}{\hat{\sigma}^2_{e_i}}\hat{\boldsymbol{\Omega}}_{ff(i)})^{-1}(\frac{1}{b_n}\boldsymbol{T}'\hat{\boldsymbol{W}}^{-1}\boldsymbol{y} - \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\hat{\alpha}_i}{\hat{\sigma}^2_{e_i}}\hat{\boldsymbol{\Omega}}_{fv(i)}), \tag{3.9}$$

which is an explicit representation of the estimator.

4. PROPERTIES OF THE ESTIMATOR

In this section, theorems will be given that describe the asymptotic properties of $\tilde{\boldsymbol{\beta}}$. Since $e_i = v_i - \boldsymbol{f}'_i\boldsymbol{\beta}$, we have $\hat{\boldsymbol{\Omega}}_{fe(i)} = \hat{\boldsymbol{\Omega}}_{fv(i)} - \hat{\boldsymbol{\Omega}}_{ff(i)}\boldsymbol{\beta}$ and $\boldsymbol{y} = \boldsymbol{T}\boldsymbol{\beta} + \boldsymbol{e}$, where $\boldsymbol{e}' = (e_1, e_2, \cdots, e_{b_n})$. Define $\hat{\boldsymbol{H}} = (\frac{1}{b_n}\boldsymbol{T}'\hat{\boldsymbol{W}}^{-1}\boldsymbol{T} - \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\hat{\alpha}_i}{\hat{\sigma}^2_{e_i}}\hat{\boldsymbol{\Omega}}_{ff(i)})$. Then, from (3.9)

$$\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{H}}^{-1}(\hat{\boldsymbol{H}}\boldsymbol{\beta} + \hat{\boldsymbol{N}}),$$

where $\hat{\boldsymbol{N}} = \frac{1}{b_n}\boldsymbol{\Pi}'\hat{\boldsymbol{W}}^{-1}\boldsymbol{e} + \frac{1}{b_n}\boldsymbol{F}'\hat{\boldsymbol{W}}^{-1}\boldsymbol{e} - \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\hat{\alpha}_i}{\hat{\sigma}^2_{e_i}}\hat{\boldsymbol{\Omega}}_{fe(i)}$. Hence,

$$\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta} = \hat{\boldsymbol{H}}^{-1}\hat{\boldsymbol{N}}. \tag{4.1}$$

The next theorem uses (4.1) and establishes a relationship which can be used to prove asymptotic normality of $\tilde{\boldsymbol{\beta}}$.

Theorem 4.1. Let the model (2.1) define a polynomial errors-in- variables model and assume:

(i) The measurement errors $\boldsymbol{w}_i$ are i.i.d. $\mathrm{N}(\mathbf{o}, \Sigma_{ww})$ with Σ_{ww} completely known.

(ii) The measurement error vectors $\boldsymbol{r}_i$ are independent and $\mathrm{E}(|\boldsymbol{r}_i|^4) = La_n^{-2}$, for some real $L < \infty$ and all i.

(iii) The elements of $\boldsymbol{\zeta}_i$ are bounded by a finite constant.

(iv) $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \infty$.

(v) The elements of the error covariance matrix satisfy $\Sigma_{ww} = O(a_n^{-1})$.

(vi) The elements of the sequence $\{a_n\}_{n=1}^{\infty}$ satisfy $a_n^{-1} = o(n^{-1/3})$.

(vii) A preliminary estimator of $\boldsymbol{\theta}$, say $\hat{\boldsymbol{\theta}}$, exists and satisfies

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} = O_p\{\max(a_n^{-1},\ n^{-1/2})\}.$$

(viii) An estimator of $\boldsymbol{\Omega}_i$, say $\hat{\boldsymbol{\Omega}}_i$, is available for $i = 1, 2, \cdots, b_n$, such that $\hat{\boldsymbol{\Omega}}_i$ and $\hat{\boldsymbol{\Omega}}_j$ are independent for $i \neq j$, and $\hat{\boldsymbol{\Omega}}_i = \hat{\boldsymbol{\Omega}}_i + O_p(a_n^{-3/2})$.

(ix) For any column vector, $\boldsymbol{z}$, in an open set containing the true parameter $\boldsymbol{\theta}$,
$0 < K_L < \boldsymbol{z}'\boldsymbol{\Omega}_i\boldsymbol{z} < K_U < \infty$ and $0 < K_L^* < \boldsymbol{z}'\hat{\boldsymbol{\Omega}}_i\boldsymbol{z} < K_U^* < \infty, i = 1, 2, \cdots,$
where $K_L(K_L^*)$ and $K_U(K_U^*)$ are fixed constants.

Then, $\sqrt{n}(\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}) = \sqrt{n}\{\frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}\boldsymbol{\Pi}\}^{-1}\tilde{\boldsymbol{\phi}}_\beta + o_p(1),$

where $\tilde{\boldsymbol{\phi}}_\beta = \frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}e + \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{1}{a_n\sigma^2_{e_i}}(\boldsymbol{f}_i e_i - \frac{a_n e_i^2}{a_n\sigma^2_{e_i}}\boldsymbol{\Omega}_{fe(i)})$.

Proof. By the assumptions (ii), (v), (vii), (viii) and the relation $\boldsymbol{\zeta}_i'\boldsymbol{\theta} = 0$,

$$\begin{aligned}\hat{\boldsymbol{\theta}}'(a_n\boldsymbol{M}_i)\hat{\boldsymbol{\theta}} &= \{\boldsymbol{\theta}' + O_p\{\max(a_n^{-1},\ n^{-1/2})\}\}\{a_n(\boldsymbol{\zeta}_i + \boldsymbol{r}_i)(\boldsymbol{\zeta}_i' + \boldsymbol{r}_i')\} \\ &\quad \{\boldsymbol{\theta} + O_p\{\max(a_n^{-1},\ n^{-1/2})\}\} \\ &= \boldsymbol{\theta}'(a_n\boldsymbol{r}_i\boldsymbol{r}_i')\boldsymbol{\theta} + O_p\{a_n^{1/2}\ \max(a_n^{-1},\ n^{-1/2})\}, \text{and}\end{aligned}$$

$$\begin{aligned}a_n\hat{\sigma}^2_{e_i} &= \{\boldsymbol{\theta}' + O_p\{\max(a_n^{-1},\ n^{-1/2})\}\}[a_n\boldsymbol{\Omega}_i + O_p(a_n^{-1/2})] \\ &\quad \{\boldsymbol{\theta} + O_p\{\max(a_n^{-1},\ n^{-1/2})\}\} \\ &= a_n\sigma^2_{e_i} + O_p(a_n^{-1/2}).\end{aligned} \tag{4.2}$$

An application of assumption (ii) and (4.2) yields

$$\begin{aligned}\hat{\alpha}_i &= \{\boldsymbol{\theta}'(a_n\boldsymbol{r}_i\boldsymbol{r}_i')\boldsymbol{\theta} + O_p\{a_n^{1/2}\max(a_n^{-1},\ n^{-1/2})\}\}\{(a_n\sigma^2_{e_i})^{-1} + O_p(a_n^{-1/2})\} \\ &= \frac{\boldsymbol{\theta}'\{E(\boldsymbol{r}_i\boldsymbol{r}_i') + O_p(a_n^{-1})\}\boldsymbol{\theta}}{\sigma^2_{e_i}} + O_p\{a_n^{1/2}\max(a_n^{-1},\ n^{-1/2})\} \\ &= 1 + O_p(1).\end{aligned} \tag{4.3}$$

Now, write the right-hand side of (4.1) as $(a_n^{-1}\hat{\boldsymbol{H}})^{-1}(a_n^{-1}\hat{\boldsymbol{N}})$ and consider two terms separately.

$$\begin{aligned}a_n^{-1}\hat{\boldsymbol{H}} &= \frac{1}{b_n}\boldsymbol{T}'(a_n\hat{\boldsymbol{W}})^{-1}\boldsymbol{T} - \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\hat{\alpha}_i}{a_n\hat{\sigma}^2_{e_i}}\hat{\boldsymbol{\Omega}}_{ff(i)} \\ &= \frac{1}{b_n}(\boldsymbol{\Pi}'+\boldsymbol{F}')(a_n\boldsymbol{W})^{-1}(\boldsymbol{\Pi}+\boldsymbol{F}) - \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\hat{\alpha}_i}{a_n\hat{\sigma}^2_{e_i}}\hat{\boldsymbol{\Omega}}_{ff(i)} + O_p(a_n^{-1/2}), \\ &= \frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}\boldsymbol{\Pi} + O_p(a_n^{-1/2}).\end{aligned} \tag{4.4}$$

The above relation (4.4) holds since we have

$$(a_n\hat{\boldsymbol{W}})^{-1} = (a_n\boldsymbol{W})^{-1} + O_p(a_n^{-1/2}),$$

$$\frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}\boldsymbol{F} = O_p(n^{-1/2}),$$

$$\frac{1}{b_n}\boldsymbol{F}'(a_n\boldsymbol{W})^{-1}\boldsymbol{\Pi} = O_p(n^{-1/2}),$$

and
$$\frac{1}{b_n}\boldsymbol{F}'(a_n\boldsymbol{W})^{-1}\boldsymbol{F} = \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{1}{a_n\sigma_{e_i}^2}\boldsymbol{\Omega}_{ff(i)} + O_p(n^{-1/2}a_n^{-1/2}).$$

To find $a_n^{-1}\hat{\boldsymbol{N}}$, consider its individual terms separately. By (4.2) and assumption (ii),

$$\begin{aligned}\frac{1}{b_n}\boldsymbol{\Pi}'(a_n\hat{\boldsymbol{W}})^{-1}\boldsymbol{e} &= \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\boldsymbol{\pi}_i e_i}{a_n\sigma_{e_i}^2} + O_p(n^{-1/2}a_n^{-1/2}) \\ &= \frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}\boldsymbol{e} + O_p(n^{-1/2}a_n^{-1/2}), \\ \frac{1}{b_n}\boldsymbol{F}'(a_n\hat{\boldsymbol{W}})^{-1}\boldsymbol{e} &= \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\boldsymbol{f}_i e_i}{a_n\sigma_{e_i}^2} + O_p(a_n^{-3/2}),\end{aligned}$$

and
$$\begin{aligned}&\frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\hat{\alpha}_i}{a_n\hat{\sigma}_{e_i}^2}\hat{\boldsymbol{\Omega}}_{fe(i)} \\ &= \frac{1}{b_n}\sum_{i=1}^{b_n}\{\{\boldsymbol{\theta}'(a_n\boldsymbol{r}_i\boldsymbol{r}_i')\boldsymbol{\theta}\}(a_n\sigma_{e_i}^2)^{-1}+O_p\{a_n^{1/2}\ max(a_n^{-1},\ n^{-1/2})\}\} \\ &\quad \times \{(a_n\sigma_{e_i}^2)^{-1} + O_p(a_n^{-1/2})\}\{\boldsymbol{\Omega}_{fe(i)} + O_p(a_n^{-3/2})\} \\ &= \frac{1}{b_n}\sum_{i=1}^{b_n}\{\frac{\boldsymbol{\theta}'(a_n\boldsymbol{r}_i\boldsymbol{r}_i')\boldsymbol{\theta}}{(a_n\sigma_{e_i}^2)^2}\boldsymbol{\Omega}_{fe(i)}\} + O_p(a_n^{-3/2}),\end{aligned}$$

since $\max(a_n^{-3/2},\ a_n^{-1/2}n^{-1/2}) = a_n^{-3/2}$ by assumption (vi). Therefore

$$\begin{aligned}a_n^{-1}\hat{\boldsymbol{N}} &= \frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}\boldsymbol{e} + \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{1}{a_n\sigma_{e_i}^2}(\boldsymbol{f}_i e_i - \frac{a_n e_i^2}{a_n\sigma_{e_i}^2}\boldsymbol{\Omega}_{fe(i)}) + O_p(a_n^{-3/2}) \\ &= \tilde{\boldsymbol{\phi}}_\beta + O_p(a_n^{-3/2}). \qquad (4.5)\end{aligned}$$

From (4.4) and (4.5), it follows that

$$\begin{aligned}\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} &= (a_n^{-1}\hat{\boldsymbol{H}})^{-1}(a_n^{-1}\hat{\boldsymbol{N}}) \\ &= \{\frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}\boldsymbol{\Pi}\}^{-1}\tilde{\boldsymbol{\phi}}_\beta + O_p(a_n^{-3/2})\end{aligned}$$

$$= \{\frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}\boldsymbol{\Pi}\}^{-1}\tilde{\boldsymbol{\phi}}_\beta + o_p(n^{-1/2}),$$

by assumption (vi). By multiplying $\sqrt{n}$ on both sides of the above equation, the conclusion of the theorem is obtained. □

It is evident that $E(\tilde{\boldsymbol{\phi}}_\beta)$ and $E(\tilde{\boldsymbol{\phi}}_\beta\tilde{\boldsymbol{\phi}}'_\beta)$ are necessary to know the mean and variance of the limiting distribution of $\tilde{\boldsymbol{\beta}}$. They are stated here without proof.

Theorem 4.2. Given the same model and assumptions as in Theorem 4.1, $\tilde{\boldsymbol{\phi}}_\beta$ has the following properties.

(i) $E(\tilde{\boldsymbol{\phi}}_\beta) = \mathbf{o}$,

(ii) $$E(\tilde{\boldsymbol{\phi}}_\beta\tilde{\boldsymbol{\phi}}'_\beta) = \frac{1}{n^2}\boldsymbol{\Pi}'\boldsymbol{W}^{-1}\boldsymbol{\Pi} + \frac{1}{n^2}\sum_{i=1}^{b_n}\frac{1}{(\sigma^2_{e_i})^2}\{\boldsymbol{\pi}_i E(e_i^2\boldsymbol{f}'_i) - \frac{\boldsymbol{\pi}_i E(e_i^3)}{\sigma^2_{e_i}}\boldsymbol{\Omega}'_{fe(i)}\}$$
$$+\frac{1}{n^2}\sum_{i=1}^{b_n}\frac{1}{(\sigma^2_{e_i})^2}\{E(e_i^2\boldsymbol{f}_i)\boldsymbol{\pi}'_i - \boldsymbol{\Omega}_{fe(i)}\frac{E(e_i^3)\boldsymbol{\pi}'_i}{\sigma^2_{e_i}}\}$$
$$+\frac{1}{n^2}\sum_{i=1}^{b_n}E\{\frac{1}{(\sigma^2_{e_i})^2}(\boldsymbol{f}_i e_i - \frac{e_i^2}{\sigma^2_{e_i}}\boldsymbol{\Omega}_{fe(i)})(\boldsymbol{f}'_i e_i - \frac{e_i^2}{\sigma^2_{e_i}}\boldsymbol{\Omega}'_{fe(i)})\}.$$ □

Now, using Theorem 4.1 and Theorem 4.2, a result concerning the asymptotic distribution of $\tilde{\boldsymbol{\beta}}$ is presented.

Theorem 4.3. Suppose that the same model and assumptions as in Theorem 4.1 hold. Then,

$$\sqrt{n}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta}) \xrightarrow{d} N(\mathbf{o}, \lim_{n\to\infty}[n\{\frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}\boldsymbol{\Pi}\}^{-1}E(\tilde{\boldsymbol{\phi}}_\beta\tilde{\boldsymbol{\phi}}'_\beta)\{\frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}\boldsymbol{\Pi}\}^{-1}]).$$

Proof. Let $\boldsymbol{\lambda}$ be an arbitrary $(k+1)$-dimensional real-valued vector, then

$$\begin{aligned}\sqrt{n}\boldsymbol{\lambda}'\{\frac{1}{b_n}\boldsymbol{\Pi}'(a_n\boldsymbol{W})^{-1}\boldsymbol{\Pi}\}(\tilde{\boldsymbol{\beta}}-\boldsymbol{\beta}) &= \sqrt{n}\boldsymbol{\lambda}'\tilde{\boldsymbol{\phi}}_\beta + o_p(1) \\ &= \sqrt{n}\frac{1}{b_n}\sum_{i=1}^{b_n}Q_i + o_p(1),\end{aligned}$$

where $Q_i = \sum_{j=1}^{k+1}\{\frac{\lambda_j \pi_{i_j} e_i}{a_n \sigma_{e_i}^2} + \frac{\lambda_j f_{i_j} e_i}{a_n \sigma_{e_i}^2} - \frac{\lambda_j (a_n e_i^2)}{(a_n \sigma_{e_i}^2)^2}\Omega_{fe(i)_j}\}$. The subscript j denotes the j-th element of a vector in the expression for Q_i. Let $Q_i^* = a_n Q_i$. The random variables $Q_i^*(i = 1, 2, \cdots, b_n)$ are independent with zero means and obviously each term of Q_i^* is a polynomial function of normally distributed random variables divided by a power of $\sigma_{e_i}^2$. Now, since $\sigma_{e_i}^2$ and $\boldsymbol{\pi}_i$ are bounded by assumptions (iii) and (ix) of Theorem 4.1, all moments of Q_i^* are bounded. So for $\nu > 2$,

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{b_n} E|Q_i|^\nu}{\{\sum_{i=1}^{b_n} E(Q_i^2)\}^{\nu/2}} = \lim_{n\to\infty} \frac{\sum_{i=1}^{b_n} E|Q_i^*|^\nu}{\{\sum_{i=1}^{b_n} E(Q_i^{*2})\}^{\nu/2}} = 0.$$

The result of the theorem follows by applying Liapounov CLT, multivariate CLT and Theorem 4.2 □

As a last step in this section, it is shown that an unbiased estimator $\hat{\boldsymbol{\Omega}}_i$ of $\boldsymbol{\Omega}_i$ defined in Section 2 satisfies assumption (viii).

Lemma 4.1. Under assumptions (i) - (ix) of Theorem 4.1,

$$\hat{\boldsymbol{\Omega}}_i = \boldsymbol{\Omega}_i + O_p(a_n^{-3/2}).$$

Proof. i) For the last k elements of the first row and column of $\hat{\boldsymbol{\Omega}}_i$. From the expression given in Section 2,

$$\begin{aligned} l\,P_{l-1}(x_i)\sigma_{uv} &= l\,(x_i^{l-1} + a_{l-3}x_i^{l-3}\sigma_{uu} + a_{l-5}x_i^{l-5}\sigma_{uu}^2 + \cdots)\sigma_{uv} \\ &= l\,x_i^{l-1}\sigma_{uv} + O_p(a_n^{-2}) \\ &= l\,\pi_i^{l-1}\sigma_{uv} + l\sum_{s=1}^{l-1}\binom{l-1}{s} u_i^s \pi_i^{l-1-s}\sigma_{uv} + O_p(a_n^{-2}) \\ &= l\,\pi_i^{l-1}\sigma_{uv} + O_p(a_n^{-3/2}), \qquad l = 1, 2, \cdots, k. \end{aligned}$$

ii) For the remaining nonzero elements of $\hat{\boldsymbol{\Omega}}_i$.
Again from the expression given in Section 2,

$$(\hat{\Omega}_i)_{2m+2,2n+2}$$

$$= \sum_{c=0}^{m+n-1} \sum_{j=c}^{m+n-1} d_j \frac{(2j)!}{(2c)!\,(j-c)!\,2^{m+n-c}} \sigma_{uu}^{m+n-c} P_{2c}(x_i)$$
$$= \sum_{c=0}^{m+n-1} \sum_{j=c}^{m+n-1} d_j \frac{(2j)!}{(2c)!\,(j-c)!\,2^{m+n-c}} \sigma_{uu}^{m+n-c} x_i^{2c} + O_p(a_n^{-2})$$
$$= \sum_{c=0}^{m+n-1} \sum_{j=c}^{m+n-1} d_j \frac{(2j)!}{(2c)!\,(j-c)!\,2^{m+n-c}} \sigma_{uu}^{m+n-c} \pi_i^{2c} + O_p(a_n^{-3/2})$$
$$= (\boldsymbol{\Omega}_i)_{2m+2,2n+2} + O_p(a_n^{-3/2}).$$

The proofs for $(\hat{\boldsymbol{\Omega}}_i)_{2m+1,2n+1}$ and $(\hat{\boldsymbol{\Omega}}_i)_{2m+1,2n+2}$ are similar.
Therefore, from i) and ii), $\hat{\boldsymbol{\Omega}}_i = \boldsymbol{\Omega}_i + O_p(a_n^{-3/2})$. □

5. SIMULATION RESULTS

In section 3, an estimator of the unknown $\boldsymbol{\beta}$ of the k-th order polynomial errors-in-variables model was constructed. Although asymptotic properties of the derived $\tilde{\boldsymbol{\beta}}$ were also obtained, the question arises as to how well this estimator performs with small sample sizes. In this section, $\tilde{\boldsymbol{\beta}}$ is considered from a numerical point of view using the cubic errors-in-variables model ($k = 3$) and is compared with the performance of $\hat{\boldsymbol{\beta}}_{OLS}$. Before stating the simulation procedure in detail, it is necessary to mention on the modification of $\tilde{\boldsymbol{\beta}}$ derived earlier.

An estimator of the k-th order polynomial model is constructed using OLS estimator as an initial estimator in Section 3. It is given in (3.9) and rewritten here for convenience. Specifically,

$$\tilde{\boldsymbol{\beta}} = \Big(\frac{1}{b_n}\sum_{i=1}^{b_n} \frac{1}{\hat{\sigma}_{e_i}^2} \boldsymbol{p}_i \boldsymbol{p}_i' - \frac{1}{b_n}\sum_{i=1}^{b_n} \frac{\hat{\alpha}_i}{\hat{\sigma}_{e_i}^2} \hat{\boldsymbol{\Omega}}_{ff(i)}\Big)^{-1}\Big(\frac{1}{b_n}\sum_{i=1}^{b_n} \frac{1}{\hat{\sigma}_{e_i}^2} \boldsymbol{p}_i \boldsymbol{y}_i - \frac{1}{b_n}\sum_{i=1}^{b_n} \frac{\hat{\alpha}_i}{\hat{\sigma}_{e_i}^2} \hat{\boldsymbol{\Omega}}_{fv(i)}\Big). \qquad (5.1)$$

Unfortunately, $\hat{\boldsymbol{\Omega}}_i$ introduced in Lemma 4.1 is not always positive semidefinite (Fuller 1987, p.213) and it was found that $\hat{\sigma}_{e_i}^2$ can take negative values in preliminary work. To eliminate that problem, another estimator of $\boldsymbol{\Omega}_i$, which is positive semidefinite, is defined. Let

$$\phi_i = \begin{bmatrix} 0 \\ \sigma_{uu}^{1/2} \\ 2\sigma_{uu}^{1/2}\{\overline{P}_1 + a^{1/2}\{P_1(x_i) - \overline{P}_1\}\} \\ 3\sigma_{uu}^{1/2}\{\overline{P}_2 + b^{1/2}\{P_2(x_i) - \overline{P}_2\}\} \end{bmatrix},$$

where $\overline{P}_1 = b_n^{-1}\sum_{i=1}^{b_n} P_1(x_i), \qquad \overline{P}_2 = b_n^{-1}\sum_{i=1}^{b_n} P_2(x_i),$

$$a = (\sum_{i=1}^{b_n}\{P_1(x_i) - \overline{P}_1\}^2)^{-1}(\sum_{i=1}^{b_n}\{P_1(x_i) - \overline{P}_1\}^2 - \frac{1}{2}b_n\sigma_{uu}) \text{ and}$$

$$b = (\sum_{i=1}^{b_n}\{P_2(x_i)-\overline{P}_2\}^2)^{-1}(\sum_{i=1}^{b_n}\{P_2(x_i)-\overline{P}_2\}^2-\frac{4}{3}b_n\sigma_{uu}^2-2\sigma_{uu}\sum_{i=1}^{b_n}P_2(x_i)).$$

Define $\hat{\boldsymbol{\Omega}}_i = \boldsymbol{\phi}_i\boldsymbol{\phi}_i'$ and $\hat{\sigma}_{e_i}^2 = \hat{\boldsymbol{\theta}}'\boldsymbol{\phi}_i\boldsymbol{\phi}_i'\hat{\boldsymbol{\theta}}$ where $\hat{\boldsymbol{\theta}}$ is OLS estimator of $\boldsymbol{\theta}$. This $\hat{\boldsymbol{\Omega}}_i$ is used in the simulation because it is positive semidefinite.

From (4.3), $\hat{\alpha}_i$ is not a consistent estimator of 1 and as a result, it was found that $\hat{\alpha}_i$ took very unstable values (for example, $10^{-4} \sim 10^2$) during the preliminary work. These unstable values of $\hat{\alpha}_i$ result in the extremely unbalanced weights for each i. Hence, $\hat{\alpha}_i$ is set equal to 1 for each i in the first term of (5.1) so that $\tilde{\boldsymbol{\beta}}$ takes the same form as that of the quadratic model estimator (Wolter & Fuller 1982a). From (4.4), it is evident that the asymptotic properties of $\tilde{\boldsymbol{\beta}}$ remain the same even with this modification. Therefore, $\tilde{\boldsymbol{\beta}}$ is modified to

$$\tilde{\boldsymbol{\beta}}_M = (\frac{1}{b_n}\sum_{i=1}^{b_n}\frac{1}{\hat{\sigma}_{e_i}^2}\boldsymbol{p}_i\boldsymbol{p}_i' - \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{1}{\hat{\sigma}_{e_i}^2}\hat{\boldsymbol{\Omega}}_{ff(i)})^{-1}(\frac{1}{b_n}\sum_{i=1}^{b_n}\frac{1}{\hat{\sigma}_{e_i}^2}\boldsymbol{p}_i\boldsymbol{y}_i - \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\hat{\alpha}_i}{\hat{\sigma}_{e_i}^2}\hat{\boldsymbol{\Omega}}_{fv(i)}). \tag{5.2}$$

Booth(1973) suggests a further modification of the estimator (5.1). To present his revised estimator, the following definition is necessary.

Definition 5.1. Define $(k+2)\times(k+2)$ matrices $\boldsymbol{M}_{gg}^*$ and $\hat{\boldsymbol{\Omega}}^*$ by

$$\boldsymbol{M}_{gg}^* = \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{1}{\hat{\sigma}_{e_i}^2}\boldsymbol{g}_i\boldsymbol{g}_i' = \begin{bmatrix} M_{yy}^* & \boldsymbol{M}_{yp}^* \\ \boldsymbol{M}_{py}^* & \boldsymbol{M}_{pp}^* \end{bmatrix}$$

and

$$\hat{\boldsymbol{\Omega}}^* = \begin{bmatrix} \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{1}{\hat{\sigma}^2_{e_i}}\sigma_{vv} & \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\hat{\alpha}_i}{\hat{\sigma}^2_{e_i}}\hat{\boldsymbol{\Omega}}_{vf(i)} \\ \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{\hat{\alpha}_i}{\hat{\sigma}^2_{e_i}}\hat{\boldsymbol{\Omega}}_{fv(i)} & \frac{1}{b_n}\sum_{i=1}^{b_n}\frac{1}{\hat{\sigma}^2_{e_i}}\hat{\boldsymbol{\Omega}}_{ff(i)} \end{bmatrix} = \begin{bmatrix} \sigma^*_{vv} & \hat{\boldsymbol{\Omega}}^*_{vf} \\ \hat{\boldsymbol{\Omega}}^*_{fv} & \hat{\boldsymbol{\Omega}}^*_{ff} \end{bmatrix}.$$

□

Using Definition 5.1, his suggested estimator $\boldsymbol{\theta}^*$ satisfies

$$(\boldsymbol{M}^*_{gg} - \alpha\hat{\boldsymbol{\Omega}}^*)\boldsymbol{\theta}^* = 0, \tag{5.3}$$

where α is the smallest root of

$$|\boldsymbol{M}^*_{gg} - \alpha\hat{\boldsymbol{\Omega}}^*| = 0. \tag{5.4}$$

Therefore, from Fuller(1987, p.125-126), the modified estimator of $\tilde{\boldsymbol{\beta}}$ given in (3.9) is

$$\tilde{\boldsymbol{\beta}}^* = (\boldsymbol{M}^*_{pp} - \alpha\hat{\boldsymbol{\Omega}}^*_{ff})^{-1}(\boldsymbol{M}^*_{py} - \alpha\hat{\boldsymbol{\Omega}}^*_{fv}),$$

where α is defined in (5.4). Finally, a small-order modification introduced by Fuller(1980) results in

$$\tilde{\boldsymbol{\beta}}^*(h) = \{\boldsymbol{M}^*_{pp} - (\alpha - \frac{h}{b_n})\hat{\boldsymbol{\Omega}}^*_{ff}\}^{-1}\{\boldsymbol{M}^*_{py} - (\alpha - \frac{h}{b_n})\hat{\boldsymbol{\Omega}}^*_{fv}\}, \tag{5.5}$$

where $h \geq 0$ is a fixed number. In simulation, estimator (5.5) is uesd with $h = 0, 1, 4$.

The cubic measurement error model considered in the simulation study is defined by $\psi_i = \pi_i - \pi_i^3$, with $\boldsymbol{w}_i = (v_i, u_i)' \sim N(\mathbf{o}, \boldsymbol{\Sigma}_{ww})$. Simulations are conducted in which 50 replications of samples of size $n = 30$ and 100 are generated from the given cubic model for each parameter set, with bivariate normal variates generated by IMSL subroutine GGNSM on an IBM 3081D mainframe computer. For each sample size, the error-free predictors are equally spaced from -5 to 5. Sample statistics such as the mean, variance and mean squared

(MSE) are computed and the evaluation of the small sample behavior of the derived estimator and OLS estimator is made on the basis of these sample statistics.

The four error covariance matrices used in the simulation are :

$$\underset{\text{I}}{\begin{bmatrix} 0.50 & 0.00 \\ 0.00 & 0.50 \end{bmatrix}} \quad \underset{\text{II}}{\begin{bmatrix} 1.50 & 0.00 \\ 0.00 & 0.50 \end{bmatrix}} \quad \underset{\text{III}}{\begin{bmatrix} 0.50 & 0.25 \\ 0.25 & 0.50 \end{bmatrix}} \quad \underset{\text{IV}}{\begin{bmatrix} 0.10 & 0.00 \\ 0.00 & 0.10 \end{bmatrix}}.$$

Covariance matrix II was included to examine the effect of changing the ratio $\lambda = \sigma_{vv}/\sigma_{uu}$ from 1 to 3 relative to covariance matrix I. Covariance matrix III was included in order to assess the performance of the estimators when the errors are correlated with $\rho = 0.5$. Covariance matrix IV decreases the size of the error variances by a factor of 5 from those for covariance matrix I. For each parameter set, $a_n = 1$ was used in calculating the value of derived estimator.

Table 1 and Table 2 summarize the results of the simmulation study. Total squared error (TSE) is the sum of four MSE's. In order to standardize the changes in the size of error variances, the ratios of TSE for each estimator reported rather than the raw TSE values. Of course, all the raw TSE values increase with the size of the error variances. The ratios of TSE for each estimator are obtained relative to that of OLS estimator and is presented in Table 1 ; i.e., each tabled value is $100(TSE_{\tilde{\boldsymbol{\beta}}^*(h)}/TSE_{OLS}), h = 0, 1, 4$. From Table 1, it is seen that all $\tilde{\boldsymbol{\beta}}^*(h), h = 0, 1, 4$, have smaller TSE than OLS estimator except $\tilde{\boldsymbol{\beta}}^*(0)$ for parameter set I with $n = 30$. And, as the sample size is increased from 30 to 100, TSE of all $\tilde{\boldsymbol{\beta}}^*(h)$ (relative to OLS estimator),$h = 0, 1, 4$, have been reduced and this reduction is at least 55%. No big difference can be found between $\tilde{\boldsymbol{\beta}}^*(1)$ and $\tilde{\boldsymbol{\beta}}^*(4)$ for all parameter sets.

Table 2 and Table 3 contain more detailed information on the estimators. These tables include the average biases and the sample variances of the estimated regression coefficients for the parameter set III with $n = 30, 100$, respectively. It is found that bias in $\hat{\boldsymbol{\beta}}_{OLS}$ is quite large, especially for the second element $\hat{\boldsymbol{\beta}}_1$, for this parameter set (In fact, for all parameter sets). The results of $\hat{\boldsymbol{\beta}}_{OLS}$ for $n = 30$ and $n = 100$ show no big difference. The above results agree with the fact that OLS estimator is biased and inconsistent in errors-in-variables model. It is seen that $\tilde{\boldsymbol{\beta}}^*(h)$'s defined in (5.5) show much better results than those of OLS estimator in both bias and variance aspects with no big difference among them.

As a conclusion of simulation study, the derived estimator $\tilde{\boldsymbol{\beta}}^*(h)$ turns out to be superior to OLS estimator in bias and TSE except one case. (Parameter set I with $n = 30$) Although there is no clear preference among $\tilde{\boldsymbol{\beta}}^*(h)$'s ($h = 0, 1, 4$) when measurement error variances are relatively small, $\tilde{\boldsymbol{\beta}}^*(1)$ or $\tilde{\boldsymbol{\beta}}^*(4)$ is considered as good choice, otherwise.

REFERENCES

(1) Booth, G. D. (1973). The Errors-in-variables Model When the Covariance Matrix is Not Constant. Ph. D. Dissertation, Iowa State University.

(2) Fuller, W. A. (1980). Properties of Some Estimators for the Errors-in-variables Model. *The Annals of Statistics, Vol. 8*, 407-422.

(3) Fuller, W. A. (1987). *Measurement Error Models.* New York : Wiley and Sons, Inc.

(4) IMSL User's Manual (1984). IMSL inc., Houston, TX.

(5) Stulajter, F. (1978). Nonlinear Estimators of Polynomials in Mean Values of a Gaussian Stochastic Process. *Kybernetika, Vol. 14*, 206-220.

(6) Wolter, K. M. and Fuller, W. A. (1982a). Estimation of the Quadratic Errors-in-variables Model. *Biometrika, Vol. 69,* 175-182.

(7) Wolter, K. M. and Fuller, W. A. (1982b). Estimation of Nonlinear Errors-in-variables Models. *The Annals of Statistics, Vol. 10,* 539-548.

Table 1. Ratios of TSE [1]

	Parameter Set							
	I		II		III		IV	
Estimator n :	30	100	30	100	30	100	30	100
$\tilde{\beta}^*(0)$	181	21	47	17	43	20	16	8
$\tilde{\beta}^*(1)$	85	22	45	17	40	20	16	8
$\tilde{\beta}^*(4)$	54	23	43	18	38	21	15	8

Table 2. Estimator Bias and Variance, Parameter Set III, $n = 30$

Estimator	β_0	β_1	β_2	β_3	Absolute Sum
			(a) Bias		
OLS	.19	-5.73	.02	.46	6.40
$\tilde{\beta}^*(0)$	.02	-3.46	.04	-.09	3.61
$\tilde{\beta}^*(1)$	.01	-3.52	.05	-.07	3.65
$\tilde{\beta}^*(4)$	-.02	-3.70	.06	.00	3.78
			(b) Variance		
OLS	9.73	7.11	.22	.04	17.10
$\tilde{\beta}^*(0)$	1.31	7.54	.57	.08	9.50
$\tilde{\beta}^*(1)$	1.20	6.05	.53	.06	7.84
$\tilde{\beta}^*(4)$	1.00	3.87	.42	.04	5.33

[1] $100(TSE_{\tilde{\beta}^*(h)}/TSE_{OLS}), h = 0, 1, 4.$

Table 3. Estimator Bias and Variance, Parameter Set III, $n = 100$

Estimator	β_0	β_1	β_2	β_3	Absolute Sum
			(a) Bias		
OLS	-.04	-6.41	-.01	.53	6.99
$\tilde{\beta}^*(0)$	-.02	-2.62	-.04	-.15	2.83
$\tilde{\beta}^*(1)$	-.02	-2.67	-.03	-.14	2.86
$\tilde{\beta}^*(4)$	-.03	-2.79	-.03	-.11	2.96
			(b) Variance		
OLS	2.38	1.83	.05	.01	4.27
$\tilde{\beta}^*(0)$	.44	1.51	.15	.03	2.13
$\tilde{\beta}^*(1)$	.43	1.44	.15	.02	2.04
$\tilde{\beta}^*(4)$	.40	1.26	.14	.02	1.82