

ASYMPTOTIC BEHAVIOR OF SIMEX ESTIMATORS IN AN ERRORS-IN-VARIABLES LINEAR STRUCTURAL REGRESSION MODEL

UDC 519.21

O. B. GONTAR

ABSTRACT. We study the SIMEX estimators in an errors-in-variables linear structural regression model. The SIMEX method was introduced by Cook and Stefanski (1994). We investigate the behavior of adjusted SIMEX estimators proposed by Polzehl and Zwanzig (2005) as the size of a sample increases. We prove that the estimators are consistent and asymptotically normal.

1. INTRODUCTION

We consider an errors-in-variables linear structural regression model. It is known that the method of moments estimators as well as least squares estimators are consistent for this model, while the “naive” estimator is not, since the latter ignores the errors in variables. The idea of the SIMulation and EXtrapolation (SIMEX) method is to determine the measurement errors experimentally via simulation and then to exclude the errors. To achieve the result, the variables are mixed with random variables with a known variance described by a parametric model. As estimates of the parameters of the model, one can use the naive estimators. Then an extrapolation procedure allows one to exclude the errors in variables.

There are three modifications of the SIMEX estimator proposed in the paper by Polzehl and Zwanzig [6].

1. SIMEX (Ordinary SIMEX). This estimator is used if the variance of the errors in variables is known. This variance influences the extrapolation procedure essentially.
2. ASIMEX (Adaptive SIMEX). The estimator uses an estimate of the variance of errors in variables instead of its unknown value.
3. SYMEX (Symmetrical SIMEX). The estimator applies to an implicit model where all variables are symmetric.

The construction of any of the above estimators consists of the following three steps:

1. (Simulation). New samples are simulated with an essentially larger variance of errors in variables by adding new errors to the actual observations.
2. (Fitting a model). A model is found to describe the dependence between an estimator and the variance of errors in variables.
3. (Extrapolation). A parameter corresponding to the case of zero variance of errors in variables is chosen for the extrapolation procedure.

2000 *Mathematics Subject Classification*. Primary 62J05, 62F12, 62-07.

Key words and phrases. SIMEX estimators, errors-in-variables models.

Polzehl and Zwanzig [6] study the estimators if the size of the sample of observations is fixed, while the number of simulated samples is growing. We consider the case where the number of simulated samples is fixed, while the size of the sample of observations is growing.

The paper is organized as follows. Section 2 contains the description of the structural errors-in-variables regression model. The construction of the SIMEX and ASIMEX estimators is described in Section 3. We prove that these estimators are consistent and asymptotically normal in Section 4. The construction and properties of the SYMEX estimator are given in Section 5. Results of the simulation of SIMEX estimators are described in Section 6.

2. THE MODEL

Consider the following errors-in-variables linear structural model:

$$(1) \quad y_i = \eta_i + \varepsilon_i, \quad x_i = \xi_i + \delta_i, \quad \eta_i = \beta^t \xi_i + \beta_0,$$

where $y_i \in \mathbb{R}$ and $x_i \in \mathbb{R}^p$ are observed variables, $i = 1, 2, \dots, n$, while $\beta \in \mathbb{R}^p$ is the parameter to be estimated. Hidden variables $\xi_i \in \mathbb{R}^p$ are independent and identically distributed random vectors. The unknown mean value of ξ_i is denoted by $\mathbf{E} \xi_i = \mu$. Assume that the covariance matrix of ξ_i is such that $\text{Cov}(\xi_i) = \sigma_\xi^2 I_p$, where σ_ξ^2 is a positive unknown number. The errors $\varepsilon_i \in \mathbb{R}$ and $\delta_i \in \mathbb{R}^p$ are independent random variables with zero mean value and covariance matrix

$$\text{Cov} \begin{pmatrix} \varepsilon_i \\ \delta_i \end{pmatrix} = \begin{pmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_\delta^2 I_p \end{pmatrix},$$

where σ_ε^2 and σ_δ^2 are positive numbers, σ_ε^2 is unknown, and I_p is the unit $p \times p$ matrix. We consider both cases of the known and unknown variance σ_δ^2 . The errors ε_i and δ_i do not depend on ξ_i .

For arbitrary $\{x_i\}$, $i = 1, 2, \dots, n$, put

$$x_i^c = x_i - \frac{1}{n} \sum_{i=1}^n x_i.$$

We also write

$$(2) \quad M_{XY} = \frac{1}{n} \sum_{i=1}^n x_i^c (y_i^c)^t$$

for arbitrary $\{x_i\}$ and $\{y_i\}$, $i = 1, 2, \dots, n$.

Let $z_i^t = (y_i, x_i^t)$. The matrix $M_{ZZ} = n^{-1} \sum_{i=1}^n z_i^c z_i^{ct}$ can be written in the block form

$$(3) \quad M_{ZZ} = \begin{pmatrix} M_{YY} & M_{YX} \\ M_{XY} & M_{XX} \end{pmatrix}.$$

The naive estimator $\widehat{\beta}_{OLS} = (M_{XX})^{-1} M_{XY}$ of the unknown parameter β is inconsistent for the above regression model, since its construction does not include the errors in variables. Consistent estimators in this case are represented by the method of moments estimators

$$(4) \quad \widehat{\beta}_{MME} = (M_{XX} - \sigma_\delta^2 I_p)^{-1} M_{XY}$$

and least squares estimators

$$(5) \quad \widehat{\beta}_{TLS} = (M_{XX} - \lambda_{\min}(M_{ZZ}) I_p)^{-1} M_{XY};$$

see [5].

3. THE SIMEX PROCEDURE

- *The simulation of auxiliary samples.* Let $\Lambda := \{0, \lambda_1, \dots, \lambda_K\}$, where $\lambda_k > 0$ for $k \geq 1$. For all $i = 1, \dots, n$ and $b = 1, \dots, B$, standard random vectors $\varepsilon_{i,b}^* \simeq N(0, I_p)$ are simulated, where B is an arbitrary large number. Further, for every $\lambda \in \Lambda$, a sample

$$\begin{aligned} x_{i,b}(\lambda) &= x_i + \sqrt{\lambda} \varepsilon_{i,b}^*, & y_{i,b}(\lambda) &= y_i, \\ z_{i,b}^t(\lambda) &= (y_{i,b}(\lambda), x_{i,b}^t(\lambda)), & i &= 1, \dots, n, \quad b = 1, \dots, B, \end{aligned}$$

is simulated. Then the matrices

$$M_{ZZ}(\lambda) = \frac{1}{nB} \sum_{i,b} z_{i,b}^c(\lambda) (z_{i,b}^c(\lambda))^t$$

are constructed by using these samples.

- *Estimation.* For all $\lambda \in \Lambda$, the naive estimator of the parameter β is constructed from matrices $M_{ZZ}(\lambda)$:

$$\hat{\beta}(\lambda) = (M_{XX}(\lambda))^{-1} M_{XY}(\lambda).$$

- *The parametric model for the naive estimator.* We assume that $\hat{\beta}(\lambda_k)$ depends on $\lambda = \lambda_k$ as follows:

$$\hat{\beta}(\lambda, \theta) = (M_{XX} + \lambda I_p)^{-1} \theta, \quad \theta \in \mathbb{R}^p.$$

The parameter θ can be found by the least squares method as follows:

$$(6) \quad \hat{\theta} = \arg \min_{\theta} \sum_{k=0}^K \|\hat{\beta}(\lambda_k) - \hat{\beta}(\lambda_k, \theta)\|^2.$$

- *Extrapolation.* If σ_{δ}^2 is known, then we put $\lambda = -\sigma_{\delta}^2$ and

$$\hat{\beta}_{\text{SIMEX}} := \hat{\beta}(-\sigma_{\delta}^2, \hat{\theta}).$$

Otherwise, if σ_{δ}^2 is unknown, then we put $\lambda = -\lambda_{\min}(M_{ZZ})$ and

$$\hat{\beta}_{\text{ASIMEX}} := \hat{\beta}(-\lambda_{\min}(M_{ZZ}), \hat{\theta}).$$

Remark 3.1. If $K = 0$, then $\Lambda = \{0\}$, whence $\hat{\theta} = M_{XY}$. Thus the SIMEX estimator for $K = 0$ becomes the method of moments estimator, while the ASIMEX estimator becomes the total least squares estimator.

Remark 3.2. The matrices of observations are averaged with respect to b in this paper, while the naive estimators are averaged with respect to B in [2].

4. PROPERTIES OF THE SIMEX AND ASIMEX ESTIMATORS

We find an explicit form of estimator (6) by applying the formula for the least squares estimator:

$$\hat{\theta} = \left(\sum_{k=0}^K (M_{XX} + \lambda_k I_p)^{-2} \right)^{-1} \left(\sum_{k=0}^K (M_{XX} + \lambda_k I_p)^{-1} (M_{XX}(\lambda_k))^{-1} (M_{XY}(\lambda_k)) \right).$$

Note that

$$(7) \quad \begin{aligned} M_{XX} &\xrightarrow{P1} (\sigma_\xi^2 + \sigma_\delta^2) I_p, & M_{XY} &\xrightarrow{P1} \sigma_\xi^2 \beta, \\ M_{XX}(\lambda_k) &\xrightarrow{P1} (\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k) I_p, & M_{XY}(\lambda_k) &\xrightarrow{P1} \sigma_\xi^2 \beta, \\ M_{YY} &\xrightarrow{P1} \sigma_\xi^2 \|\beta\|^2 + \sigma_\varepsilon^2 \end{aligned}$$

as $n \rightarrow \infty$. The symbol $\xrightarrow{P1}$ stands for the convergence with probability one.

Thus $\hat{\theta} \xrightarrow{P1} \sigma_\xi^2 \beta$ and every entry of the matrix M_{ZZ} converges as $n \rightarrow \infty$ to that of the matrix

$$(8) \quad M_{ZZ} \xrightarrow{P1} \begin{pmatrix} \sigma_\xi^2 \|\beta\|^2 + \sigma_\varepsilon^2 & \sigma_\xi^2 \beta^t \\ \sigma_\xi^2 \beta & (\sigma_\xi^2 + \sigma_\delta^2) I_p \end{pmatrix}.$$

Since the matrix M_{ZZ} is symmetric, its minimal eigenvalue converges to the minimal eigenvalue of the limit matrix. The latter is an almost sure solution of the following equation:

$$\begin{aligned} &\det \begin{pmatrix} \sigma_\xi^2 \|\beta\|^2 + \sigma_\varepsilon^2 - \lambda & \sigma_\xi^2 \beta^t \\ \sigma_\xi^2 \beta & (\sigma_\xi^2 + \sigma_\delta^2 - \lambda) I_p \end{pmatrix} = 0, \\ &(\sigma_\xi^2 \|\beta\|^2 + \sigma_\varepsilon^2 - \lambda) (\sigma_\xi^2 + \sigma_\delta^2 - \lambda)^p \\ &+ \sum_{i=1}^p (-1)^p \sigma_\xi^2 \beta_i \det \begin{pmatrix} \sigma_\xi^2 \beta_1 \dots \sigma_\xi^2 \beta_{i-1} & \sigma_\xi^2 \beta_i & \sigma_\xi^2 \beta_{i+1} \dots \sigma_\xi^2 \beta_p \\ (\sigma_\xi^2 + \sigma_\delta^2 - \lambda) I_{i-1} & 0 & 0 \\ 0 & 0 & (\sigma_\xi^2 + \sigma_\delta^2 - \lambda) I_{p-i} \end{pmatrix} = 0, \\ &(\sigma_\xi^2 \|\beta\|^2 + \sigma_\varepsilon^2 - \lambda) (\sigma_\xi^2 + \sigma_\delta^2 - \lambda)^p - \sigma_\xi^4 \|\beta\|^2 (\sigma_\xi^2 + \sigma_\delta^2 - \lambda)^{p-1} = 0. \end{aligned}$$

The minimal solution of the latter equation is

$$(9) \quad c := \frac{1}{2} \left(\sigma_\xi^2 + \sigma_\varepsilon^2 + \sigma_\xi^2 \beta^t \beta + \sigma_\delta^2 - \sqrt{(\sigma_\xi^2 + \sigma_\delta^2 - \sigma_\xi^2 \beta^t \beta - \sigma_\varepsilon^2)^2 + 4\sigma_\xi^4 \beta^t \beta} \right).$$

The number c is such that

$$\min(\sigma_\varepsilon^2, \sigma_\delta^2) \leq c \leq \max(\sigma_\varepsilon^2, \sigma_\delta^2).$$

Note that $c < \sigma_\xi^2 + \sigma_\delta^2$. If $\sigma_\varepsilon^2 = \sigma_\delta^2 =: \sigma^2$, then $c = \sigma^2$.

4.1. The convergence of estimators.

Theorem 4.1. *Let σ_δ^2 be known. The estimator $\hat{\beta}_{SIMEX}$ is consistent for the model (1) and is asymptotically normal if the errors are normal. In this case, the covariance matrix of the limit distribution is given by*

$$\begin{aligned} \Xi_{SIMEX} &= \frac{1}{\sigma_\xi^4} (\sigma_\xi^2 \sigma_\delta^2 \|\beta\|^2 + \sigma_\xi^2 \sigma_\varepsilon^2 + \sigma_\delta^2 \sigma_\varepsilon^2 + 2\sigma_\delta^4 \|\beta\|^2) I_p \\ &+ \frac{1}{\sigma_\xi^4} \frac{1}{B} \left(\sum_{k=0}^K \frac{1}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^{-2} \left(\sum_{k=0}^K \frac{\sqrt{\lambda_k}}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^2 \\ &\times (\sigma_\varepsilon^2 I_p + \sigma_\xi^2 \|\beta\|^2) I_p. \end{aligned}$$

Proof. Here and in what follows the symbol \xrightarrow{d} stands for the convergence in distribution.

The strong consistence of $\hat{\beta}_{SIMEX}$ follows from

$$\hat{\beta}_{SIMEX} = (M_{XX} - \sigma_\delta^2 I_p)^{-1} \hat{\theta} \xrightarrow{P1} \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 - \sigma_\delta^2} \beta = \beta.$$

To prove the asymptotic normality of the estimator $\widehat{\beta}_{\text{SIMEX}}$ consider the asymptotic covariance matrix of the method of moments estimator $\widehat{\beta}_{\text{MME}}$ defined by (4):

$$\sqrt{n}(\widehat{\beta}_{\text{MME}} - \beta) \xrightarrow{d} \nu_{\text{MME}} \simeq N(0, \Sigma_{\text{MME}}),$$

where

$$\Sigma_{\text{MME}} = \frac{1}{\sigma_\xi^4} (\sigma_\xi^2 \sigma_\delta^2 \|\beta\|^2 + \sigma_\xi^2 \sigma_\varepsilon^2 + \sigma_\delta^2 \sigma_\varepsilon^2 + 2\sigma_\delta^4 \|\beta\|^2) I_p$$

(see [3, p. 147]). Then

$$\begin{aligned} \sqrt{n}(\widehat{\beta}_{\text{MME}} - \beta) &= \sqrt{n} (M_{XX} - \sigma_\delta^2)^{-1} (M_{XY} - (M_{XX} - \sigma_\delta^2) \beta) \\ &= \sqrt{n} \frac{M_{XY} - (M_{XX} - \sigma_\delta^2) \beta}{\sigma_\xi^2 + o(1)}. \end{aligned}$$

To prove the asymptotic normality of $\widehat{\beta}_{\text{SIMEX}}$ we introduce the notation

$$\eta_k = \sqrt{n} ((M_{XX}(\lambda_k))^{-1} M_{XY}(\lambda_k) - (M_{XX} + \lambda_k I_p)^{-1} (M_{XX} - \sigma_\delta^2 I_p) \beta).$$

After simple algebra we get

$$\sqrt{n}(\widehat{\beta}_{\text{SIMEX}} - \beta) = \left(\sum_{k=0}^K (M_{XX} + \lambda_k I_p)^{-2} \right)^{-1} \sum_{k=0}^K (M_{XX} + \lambda_k I_p)^{-1} (M_{XX} - \sigma_\delta^2 I_p)^{-1} \eta_k.$$

First we prove the asymptotic normality of η_k . Note that

$$\begin{aligned} M_{XX} &= M_{\xi\xi} + 2M_{\xi\delta} + M_{\delta\delta}, \\ M_{XY}(\lambda_k) &= M_{\xi\xi}\beta + M_{\xi\delta}\beta + M_{\xi\varepsilon} + M_{\varepsilon\delta} + \sqrt{\lambda_k} M_{\varepsilon\varepsilon^*} + \sqrt{\lambda_k} M_{\varepsilon^*\xi}\beta \\ &= M_{XY} + \sqrt{\lambda_k} M_{\varepsilon\varepsilon^*} + \sqrt{\lambda_k} M_{\varepsilon^*\xi}\beta. \end{aligned}$$

Here

$$M_{\varepsilon\varepsilon^*} = \frac{1}{n} \sum_{i=1}^n \varepsilon_i^c \varepsilon_i^{*c}, \quad M_{\varepsilon^*\xi} = \frac{1}{n} \sum_{i=1}^n \xi_i^c \varepsilon_i^{*c}, \quad \varepsilon_i^{*c} = \frac{1}{B} \sum_{i=1}^B \varepsilon_{i,b}^*.$$

Now

$$\begin{aligned} \eta_k &= \sqrt{n} \frac{1}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k + o(1)} \left(M_{XY} - (M_{XX} - \sigma_\delta^2 I_p) \beta + \sqrt{\lambda_k} M_{\varepsilon\varepsilon^*} + \sqrt{\lambda_k} M_{\varepsilon^*\xi} \beta \right) \\ &= \sqrt{n} \frac{M_{XY} - (M_{XX} - \sigma_\delta^2 I_p) \beta}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k + o(1)} + \sqrt{n} \frac{\sqrt{\lambda_k} M_{\varepsilon\varepsilon^*} + \sqrt{\lambda_k} M_{\varepsilon^*\xi} \beta}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k + o(1)}. \end{aligned}$$

Furthermore,

$$\begin{aligned} &(M_{XX} - \sigma_\delta^2 I_p)^{-1} \eta_k \\ &= \sqrt{n} \frac{M_{XY} - (M_{XX} - \sigma_\delta^2 I_p) \beta}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k + o(1))(\sigma_\xi^2 + o(1))} + \sqrt{n} \frac{\sqrt{\lambda_k} M_{\varepsilon\varepsilon^*} + \sqrt{\lambda_k} M_{\varepsilon^*\xi} \beta}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k + o(1))(\sigma_\xi^2 + o(1))} \\ &= \sqrt{n} (\widehat{\beta}_{\text{MME}} - \beta) \frac{1}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k + o(1)} + \sqrt{n} \frac{\sqrt{\lambda_k} M_{\varepsilon\varepsilon^*} + \sqrt{\lambda_k} M_{\varepsilon^*\xi} \beta}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k + o(1))(\sigma_\xi^2 + o(1))}. \end{aligned}$$

One can show that

$$\sqrt{n} M_{\varepsilon^*\varepsilon} \xrightarrow{d} \nu_1 \simeq N\left(0, \frac{1}{B} \sigma_\varepsilon^2 I_p\right), \quad \sqrt{n} M_{\varepsilon^*\xi} \beta \xrightarrow{d} \nu_2 \simeq \left(0, \frac{1}{B} \sigma_\xi^2 \|\beta\|^2 I_p\right).$$

Then

$$\sqrt{n} \begin{pmatrix} \widehat{\beta}_{\text{MME}} - \beta \\ M_{\varepsilon\varepsilon^*} \\ M_{\varepsilon^*\xi} \beta \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \nu_{\text{MME}} \\ \nu_1 \\ \nu_2 \end{pmatrix} =: \nu \simeq N(0, \Sigma),$$

where Σ is a block diagonal matrix whose diagonal entries are the corresponding covariance matrices.

Applying the Cramér–Wald device [1] we get

$$\begin{aligned} (M_{XX} - \sigma_\delta^2 I_p)^{-1} \eta_k &\xrightarrow{d} \eta_k^* = \frac{1}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k} \left(I_p, \frac{\sqrt{\lambda_k}}{\sigma_\xi^2} I_p, \frac{\sqrt{\lambda_k}}{\sigma_\xi^2} I_p \right) \nu, \\ \eta_k^* &\simeq N \left(0, \frac{1}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \left(\Sigma_{MME} + \frac{1}{\sigma_\xi^4} \left(\frac{1}{B} \lambda_k \sigma_\varepsilon^2 + \frac{1}{B} \lambda_k \sigma_\xi^2 \|\beta\|^2 \right) I_p \right) \right). \end{aligned}$$

Note that

$$\sum_{k=0}^K (M_{XX} + \lambda_k I_p)^{-1} (M_{XX} - \sigma_\delta^2 I_p)^{-1} \eta_k \xrightarrow{d} \sum_{k=0}^K \frac{1}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k} \eta_k^* =: \zeta \simeq N(0, \Xi),$$

where

$$\Xi = \sum_{k,m=0}^K \frac{1}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k} \mathbf{E} \eta_k^* (\eta_m^*)^t \frac{1}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda_m}.$$

We have

$$\begin{aligned} \mathbf{E} \eta_k^* (\eta_m^*)^t &= \frac{1}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_m)} \\ &\quad \times \left(\Sigma_{MME} + \frac{1}{\sigma_\xi^4} \frac{1}{B} \sqrt{\lambda_m} \sqrt{\lambda_k} (\sigma_\varepsilon^2 + \sigma_\xi^2 \|\beta\|^2) I_p \right). \end{aligned}$$

Then

$$\begin{aligned} \Xi &= \left(\sum_{k=0}^K \frac{1}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^2 \Sigma_{MME} \\ &\quad + \frac{1}{\sigma_\xi^4} \frac{1}{B} \left(\sum_{k=0}^K \frac{\sqrt{\lambda_k}}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^2 (\sigma_\varepsilon^2 I_p + \sigma_\xi^2 \|\beta\|^2 I_p). \end{aligned}$$

Turning back to the initial expression we get

$$\begin{aligned} &\sqrt{n}(\widehat{\beta}_{\text{SIMEX}} - \beta) \\ &= \left(\sum_{k=0}^K (M_{XX} + \lambda_k I_p)^{-2} \right)^{-1} \sum_{k=0}^K (M_{XX} + \lambda_k I_p)^{-1} (M_{XX} - \sigma_\delta^2 I_p)^{-1} \eta_k \\ &\xrightarrow{d} \left(\sum_{k=0}^K \frac{1}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^{-1} \zeta \simeq N(0, \Sigma_{\text{SIMEX}}), \end{aligned}$$

where

$$\begin{aligned} \Sigma_{\text{SIMEX}} &= \Sigma_{MME} \\ &\quad + \frac{1}{\sigma_\xi^4} \frac{1}{B} \left(\sum_{k=0}^K \frac{1}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^{-2} \left(\sum_{k=0}^K \frac{\sqrt{\lambda_k}}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^2 \\ &\quad \times (\sigma_\varepsilon^2 + \sigma_\xi^2 \|\beta\|^2) I_p. \quad \square \end{aligned}$$

Remark 4.1. The matrices Σ_{SIMEX} and Σ_{MME} coincide if $K = 0$ ($\lambda_0 = 0$), since the estimators themselves coincide in this case.

Theorem 4.2. a) Assume that the variances of errors σ_ε^2 and σ_δ^2 are equal for the model (1). Then the estimator $\widehat{\beta}_{ASIMEX}$ is strictly consistent. Moreover the estimator is asymptotically normal if the errors are normal. In this case, the covariance matrix of the limit distribution is given by

$$\begin{aligned} \Xi_{ASIMEX} &= \frac{1}{\sigma_\xi^4} (\sigma_\xi^2 \sigma_\delta^2 \|\beta\|^2 + \sigma_\xi^2 \sigma_\varepsilon^2 + \sigma_\delta^2 \sigma_\varepsilon^2) I_p \\ &\quad + \frac{1}{\sigma_\xi^4} \frac{1}{B} \left(\sum_{k=0}^K \frac{1}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^{-2} \left(\sum_{k=0}^K \frac{\sqrt{\lambda_k}}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^2 \\ &\quad \times (\sigma_\varepsilon^2 + \sigma_\xi^2 \|\beta\|^2) I_p. \end{aligned}$$

b) If the size of the sample is growing, the asymptotic behavior of the ASIMEX estimator is the same as that of the least squares estimator $\widehat{\beta}_{TLS}$ (5), namely

$$\|\widehat{\beta}_{ASIMEX} - \widehat{\beta}_{TLS}\| \rightarrow 0 \quad \text{almost surely}$$

(we do not assume that the variances of errors σ_ε^2 and σ_δ^2 are equal for this result).

Proof. Recall that $\lambda_{\min}(M_{ZZ}) \rightarrow \sigma^2$ almost surely if $\sigma_\varepsilon^2 = \sigma_\delta^2 = \sigma^2$.

a) The consistence of the estimator $\widehat{\beta}_{ASIMEX}$ follows from the relation

$$\widehat{\beta}_{ASIMEX} = (M_{XX} - \lambda_{\min}(M_{ZZ}))^{-1} \widehat{\theta} \rightarrow \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma^2 - \sigma^2} \beta = \beta \quad \text{almost surely.}$$

To prove that the estimator $\widehat{\beta}_{ASIMEX}$ is asymptotically normal we use the asymptotic covariance matrix of the method of moments estimator $\widehat{\beta}_{TLS}$ ([3, p. 148]):

$$\sqrt{n}(\widehat{\beta}_{TLS} - \beta) \xrightarrow{d} \nu_{TLS} \simeq N(0, \Sigma_{TLS}),$$

where

$$\Sigma_{TLS} = \frac{1}{\sigma_\xi^4} (\sigma_\xi^2 \sigma_\delta^2 \|\beta\|^2 + \sigma_\xi^2 \sigma_\varepsilon^2 + \sigma_\delta^2 \sigma_\varepsilon^2) I_p.$$

Further, reasoning similarly to the preceding proof we obtain

$$\sqrt{n}(\widehat{\beta}_{ASIMEX} - \beta) \rightarrow N(0, \Sigma_{ASIMEX}),$$

where

$$\begin{aligned} \Sigma_{ASIMEX} &= \Sigma_{TLS} \\ &\quad + \frac{1}{\sigma_\xi^4} \frac{1}{B} \left(\sum_{k=0}^K \frac{1}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^{-2} \left(\sum_{k=0}^K \frac{\sqrt{\lambda_k}}{(\sigma_\xi^2 + \sigma_\delta^2 + \lambda_k)^2} \right)^2 \\ &\quad \times (\sigma_\varepsilon^2 I_p + \sigma_\xi^2 \|\beta\|^2) I_p. \end{aligned}$$

b) We have

$$\begin{aligned} &\|\widehat{\beta}_{ASIMEX} - \widehat{\beta}_{TLS}\| \\ &= \left\| (M_{XX} - \lambda_{\min}(M_{ZZ}) I_p)^{-1} \widehat{\theta} - (M_{XX} - \lambda_{\min}(M_{ZZ}) I_p)^{-1} M_{XY} \right\| \\ &\leq \left\| (M_{XX} - \lambda_{\min}(M_{ZZ}) I_p)^{-1} \right\| \|\widehat{\theta} - M_{XY}\| \rightarrow 0 \end{aligned}$$

almost surely for all σ_ε^2 and σ_δ^2 . \square

5. THE SYMEX ESTIMATOR

5.1. **Implicit model.** Consider the implicit model. We have

$$(10) \quad 0 = -\eta + \beta^t \xi + \beta_0 = \alpha^t \zeta + \beta_0,$$

where $\alpha = (-1, \beta_1, \beta_2, \dots, \beta_p)^t$, $\zeta = (\eta, \xi_1, \xi_2, \dots, \xi_p)^t$,

$$z = \zeta + \bar{\varepsilon},$$

$\bar{\varepsilon}^t = (\varepsilon, \delta^t)$, and $\text{Cov}(\bar{\varepsilon}) = \sigma^2 I_{p+1}$.

We assume that

$$\min_{l=0, \dots, p} |\beta_l| > 0.$$

Dividing (10) by $-\beta_l$, $l = 1, \dots, p$, yields:

$$0 = \frac{-\eta + \beta^t \xi + \beta_0}{-\beta_l} = (\alpha^l)^t \zeta + \beta_0^l,$$

where $\alpha^l = -\alpha/\beta_l$, $l = 1, \dots, p$, $\alpha^0 = \alpha$, and $\beta_0^l = -\beta_0/\beta_l$. It is clear that $\alpha_l^l = -1$, where α_l^l is the coordinate l of the vector α^l .

Denote by $\alpha_{(-l)}^l$ the vector α^l whose coordinate l is excluded. Let the $n \times p$ matrix $Z_{(-l)}$ be formed by deleting the column Z_l in the matrix Z ,

$$(11) \quad Z = \begin{pmatrix} Y_1^c & X_{11}^c & \dots & X_{l-11}^c & X_{l1}^c & X_{l+11}^c & \dots & X_{p1}^c \\ Y_2^c & X_{12}^c & \dots & X_{l-12}^c & X_{l2}^c & X_{l+12}^c & \dots & X_{p2}^c \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ Y_p^c & X_{1p}^c & \dots & X_{l-1p}^c & X_{lp}^c & X_{l+1p}^c & \dots & X_{pp}^c \end{pmatrix},$$

$$(12) \quad Z_l = \begin{pmatrix} X_{l1}^c \\ X_{l2}^c \\ \dots \\ X_{lp}^c \end{pmatrix}.$$

Let

$$M_{(-l)(-l)} = \frac{1}{n} Z_{(-l)}^t Z_{(-l)}, \quad M_{(-l)(l)} = \frac{1}{n} Z_{(-l)}^t Z_l = M_{(l)(-l)}^t, \quad M_{(l)(l)} = \frac{1}{n} Z_l^t Z_l.$$

Here $M_{(-l)(-l)}$ is a $p \times p$ matrix and $M_{(-l)(l)}$ and $M_{(l)(-l)}$ are $p \times 1$ and $1 \times p$ vectors, respectively; $M_{(l)(l)}$ is a number.

If the $1 \times (p-1)$ vector $\zeta_{(-l)}$ is formed from the coordinates of the vector ζ with excluded coordinate l , then

$$\zeta_l = (\alpha_{(-l)}^l)^t \zeta_{(-l)} + \beta_0^l.$$

5.2. **SYMEX procedure.** Denote by P_l the diagonal matrix whose entry ll is equal to 1, while all other entries are zero.

- *Simulation of auxiliary samples.* For every $\lambda \in \{\lambda_k, k = 0, \dots, K\}$, a sample

$$z_{ib}(\lambda) = z_i + (I - P_l) \lambda^{1/2} \varepsilon_{ib}^*$$

is simulated, where $\varepsilon_{ib}^* \simeq N_{p+1}(0, I_{p+1})$, $b = 1, \dots, B$, $i = 1, \dots, n$.

- *Estimation.*

$$\hat{\alpha}_{(-l)}^l(\lambda) = M_{(-l)(-l)}^{-1}(\lambda) M_{(-l)(l)}(\lambda), \quad \hat{\alpha}_{(l)}^l(\lambda) = -1.$$

- *Approximation.* We assume that $\hat{\alpha}_{(-l)}^l(\lambda)$ depends on λ in the following way:

$$\hat{\alpha}_{(-l)}^l(\lambda, \theta_l) = (M_{(-l)(-l)} + \lambda I_p)^{-1} \theta_l, \quad \hat{\alpha}_{(l)}^l(\lambda, \theta_l) = -1.$$

Here the symbol θ_l stands for the $1 \times p$ vector defined by

$$\widehat{\theta}_l = \arg \min_{\theta_l} \sum_{k=0}^K \|\widehat{\alpha}_{(-l)}^l(\lambda_k) - \widehat{\alpha}_{(-l)}^l(\lambda_k, \theta_l)\|^2.$$

Turning back to the initial coordinates gives

$$\widehat{\beta}_j^l(\lambda) = -\frac{\widehat{\alpha}_j^l(\lambda, \theta_l)}{\widehat{\alpha}_0^l(\lambda, \theta_l)}, \quad j = 1, \dots, p.$$

- *Optimal value of λ .* Put $\lambda_{\min} = \min_l \lambda_{\min}(M_{(-l)(-l)})$ and

$$\lambda^* := \arg \min_{\lambda \in [-\lambda_{\min}, 0]} \sum_{l=0}^p \sum_{m=0}^p \|\widehat{\beta}^l(\lambda) - \widehat{\beta}^m(\lambda)\|^2.$$

- *SYMEX estimator:*

$$\widehat{\beta}_{\text{SYMEX}} := \widehat{\beta}(\lambda^*) = \frac{1}{p+1} \sum_{l=0}^p \widehat{\beta}^l(\lambda^*).$$

5.3. Properties of the estimator.

Theorem 5.1. *a) The estimator $\widehat{\beta}_{\text{SYMEX}}$ is strongly consistent for the model (10) if $\sigma_\varepsilon^2 = \sigma_\delta^2$.*

b) If $p = 1$ and σ_ε^2 and σ_δ^2 are arbitrary, then the asymptotic behavior of the estimator $\widehat{\beta}_{\text{SYMEX}}$ is the same as that of the total least squares estimator $\widehat{\beta}_{\text{TLS}}$ defined by (5); namely,

$$\|\widehat{\beta}_{\text{SYMEX}} - \widehat{\beta}_{\text{TLS}}\| \rightarrow 0 \quad \text{almost surely.}$$

Proof. The matrix $M_{(-l)(-l)}$ coincides with M_{XX} , and $M_{(-l)(l)}$ coincides with M_{XY} if $l = 0$. It follows from the above computations that

$$\widehat{\beta}^0(\lambda) \rightarrow \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \beta \quad \text{almost surely.}$$

Let $l \neq 0$. Consider the almost sure limits as $n \rightarrow \infty$:

$$\begin{aligned} M_{(-l)(-l)} &\rightarrow \begin{pmatrix} \sigma_\eta^2 + \sigma_\varepsilon^2 & 0 \\ 0 & (\sigma_\xi^2 + \sigma_\delta^2)I_{p-1} \end{pmatrix}, \\ M_{(-l)(-l)}(\lambda) &\rightarrow \begin{pmatrix} \sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda & 0 \\ 0 & (\sigma_\xi^2 + \sigma_\delta^2 + \lambda)I_{p-1} \end{pmatrix}, \\ M_{(-l)(l)} &\rightarrow (\alpha_0^l \sigma_\eta^2, \alpha_1^l \sigma_\xi^2, \dots, \alpha_{l-1}^l \sigma_\xi^2, \alpha_{l+1}^l \sigma_\xi^2, \dots, \alpha_p^l \sigma_\xi^2)^t, \\ (13) \quad M_{(-l)(l)}(\lambda) &\rightarrow (\alpha_0^l \sigma_\eta^2, \alpha_1^l \sigma_\xi^2, \dots, \alpha_{l-1}^l \sigma_\xi^2, \alpha_{l+1}^l \sigma_\xi^2, \dots, \alpha_p^l \sigma_\xi^2)^t. \end{aligned}$$

We rewrite $\widehat{\theta}_l$ in an explicit form:

$$\begin{aligned} \widehat{\theta}_l &= \left(\sum_{k=0}^K (M_{(-l)(-l)}(\lambda_k))^{-2} \right)^{-1} \\ &\quad \times \left(\sum_{k=0}^K (M_{(-l)(-l)} + \lambda_k I_p)^{-1} (M_{(-l)(-l)}(\lambda_k))^{-1} M_{(-l)(l)}(\lambda_k) \right). \end{aligned}$$

Passing to the limit as $n \rightarrow \infty$ we get

$$\widehat{\theta}_l \xrightarrow{P1} (\alpha_0^l \sigma_\eta^2, \alpha_1^l \sigma_\xi^2, \dots, \alpha_{l-1}^l \sigma_\xi^2, \alpha_{l+1}^l \sigma_\xi^2, \dots, \alpha_p^l \sigma_\xi^2)^t =: h.$$

Further

$$\begin{aligned} \widehat{\alpha}_{(-l)}^l(\lambda) &\xrightarrow{P_1} \begin{pmatrix} \sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda & 0 \\ 0 & (\sigma_\xi^2 + \sigma_\delta^2 + \lambda)I_{p-1} \end{pmatrix}^{-1} h \\ &= \begin{pmatrix} \alpha_0^l \frac{\sigma_\eta^2}{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda} \\ \alpha_1^l \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ \dots \\ \alpha_{l-1}^l \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ \alpha_{l+1}^l \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ \dots \\ \alpha_p^l \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \end{pmatrix} =: a_l(\lambda). \end{aligned}$$

Turning back to the initial coordinates yields

$$\widehat{\beta}^l(\lambda) \xrightarrow{P_1} \begin{pmatrix} -\frac{\alpha_l^l}{\alpha_0^l} \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ \dots \\ -\frac{\alpha_{l-1}^l}{\alpha_0^l} \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ \frac{1}{\alpha_0^l} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda} \\ -\frac{\alpha_{l+1}^l}{\alpha_0^l} \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ \dots \\ -\frac{\alpha_p^l}{\alpha_0^l} \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \end{pmatrix} = \begin{pmatrix} \beta_1 \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ \dots \\ \beta_{l-1} \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ \beta_l \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\eta^2} \\ \beta_{l+1} \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ \dots \\ \beta_p \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \end{pmatrix} =: b_l(\lambda).$$

Put

$$\begin{aligned} K_n(\lambda) &= \sum_{l=0}^p \sum_{m=0}^p \|\widehat{\beta}^l(\lambda) - \widehat{\beta}^m(\lambda)\|^2 \\ &= 2 \sum_{m=1}^p \|\widehat{\beta}^0(\lambda) - \widehat{\beta}^m(\lambda)\|^2 + \sum_{l=1}^p \sum_{m=1}^p \|\widehat{\beta}^l(\lambda) - \widehat{\beta}^m(\lambda)\|^2, \\ K_n(\lambda) &\xrightarrow{P_1} 2 \sum_{m=1}^p \left\| \begin{pmatrix} \beta_1 \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} - \beta_1 \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ \dots \\ \beta_m \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} - \beta_m \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\eta^2} \\ \dots \\ \beta_p \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} - \beta_p \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \end{pmatrix} \right\|^2 \\ &\quad + \sum_{l=1}^p \sum_{m=1}^p \left\| \begin{pmatrix} 0 \\ \beta_l \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\eta^2} - \beta_l \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \\ 0 \\ \beta_m \frac{\sigma_\xi^2}{\sigma_\eta^2} \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} - \beta_m \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\eta^2} \\ 0 \end{pmatrix} \right\|^2 \end{aligned}$$

$$\begin{aligned}
&= 2(p-1)\|\beta\|^2 \left(\frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\eta^2} \left(\frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} - 1 \right) \right)^2 \\
&\quad + 2(p-1)\|\beta\|^2 \left(\left(\frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\eta^2} - 1 \right) \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \right)^2 \\
&\quad + 2\|\beta\|^2 \left(\frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\eta^2} - \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \right)^2 \\
&=: f(\lambda).
\end{aligned}$$

It is clear that $f(\lambda) \geq 0$ for all real $\lambda \neq -\sigma_\xi^2 - \sigma_\delta^2$.

Since $\lambda_{\min} \rightarrow \min(\sigma_\xi^2 + \sigma_\delta^2, \sigma_\eta^2 + \sigma_\varepsilon^2)$ almost surely as $n \rightarrow \infty$, we consider λ belonging to the compact set $B = [-\min(\sigma_\xi^2 + \sigma_\delta^2, \sigma_\eta^2 + \sigma_\varepsilon^2) - \varepsilon, 0]$, where $0 < \varepsilon < \frac{1}{2} \min\{\sigma_\xi^2, \sigma_\eta^2\}$. We show that $K_n(\lambda) \rightrightarrows f(\lambda)$ in B .

First we show that $\widehat{\alpha}_{-l}^l(\lambda) \rightrightarrows a_l(\lambda)$ in B . It is easy to prove that $\{\widehat{\alpha}_{-l}^l(\lambda), n \geq 1\}$ is a sequence of uniformly bounded and equicontinuous functions. To check this property one can prove that the sequence of functions $\{\widehat{\alpha}_{-l}^l(\lambda), n \geq 1\}$ and the sequence of their derivatives $\{\widehat{\alpha}_{-l}^l(\lambda)', n \geq 1\}$ are bounded with probability one and the upper bound is uniform. By Arzelà–Ascoli’s theorem, $\{\widehat{\alpha}_{-l}^l(\lambda), n \geq 1\}$ is a precompact sequence in $C(B)$ almost surely.

Since this sequence converges pointwise to $a_l(\lambda)$, we get

$$\widehat{\alpha}_{-l}^l(\lambda) \rightrightarrows a_l(\lambda) \quad \text{and} \quad \widehat{\beta}_l(\lambda) \rightrightarrows b_l(\lambda)$$

almost surely. Thus $K_n(\lambda) \rightrightarrows f(\lambda)$ in B . Therefore λ^* , as a point of minimum of $K_n(\lambda)$, converges almost surely to a point of minimum of $f(\lambda)$ in the compact set B .

a) Consider the case of $\sigma_\delta^2 = \sigma_\varepsilon^2 = \sigma^2$. Then $\min(\sigma_\xi^2 + \sigma_\delta^2, \sigma_\eta^2 + \sigma_\varepsilon^2) = \min(\sigma_\xi^2, \sigma_\eta^2) + \sigma^2$ and $B = [-\min(\sigma_\xi^2, \sigma_\eta^2) - \sigma^2 - \varepsilon, 0]$. There is only one zero $\lambda = -\sigma^2$ of the function $f(\lambda)$ in this compact set; thus it is the limit of the point of minimum of $K_n(\lambda)$. Thus

$$\widehat{\beta}_{\text{SYMEX}} = \frac{1}{p+1} \sum_{l=0}^p \widehat{\beta}^l(\lambda^*) \rightarrow \frac{1}{p+1} \sum_{l=0}^p \beta = \beta \quad \text{almost surely.}$$

b) Let $p = 1$ and let σ_δ^2 and σ_ε^2 be arbitrary. Then

$$f(\lambda) = 2\|\beta\|^2 \left(\frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\eta^2} - \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda} \right)^2.$$

The function $f(\lambda)$ has a unique zero $\lambda = -c$ in the compact set B , where c is given by equality (9). We introduce the following notation:

$$g(\lambda) = \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\delta^2 + \lambda}, \quad h(\lambda) = \frac{\sigma_\eta^2 + \sigma_\varepsilon^2 + \lambda}{\sigma_\eta^2}.$$

Then $f(\lambda) = (g(\lambda) - h(\lambda))^2$. It is clear that $g(-c) = h(-c)$. We represent the limit value of $\widehat{\beta}_{\text{SYMEX}}$ in terms of these functions:

$$\widehat{\beta}_{\text{SYMEX}} = \frac{1}{2}(\widehat{\beta}^0(\lambda) + \widehat{\beta}^1(\lambda)) \rightarrow \frac{1}{2}(g(\lambda) + h(\lambda))\beta.$$

Substituting $\lambda = -c$ to the latter relation we get

$$\widehat{\beta}_{\text{SYMEX}} \xrightarrow{P1} g(-c)\beta.$$

Note that

$$\widehat{\beta}_{\text{TLS}} \xrightarrow{P1} g(-c)\beta.$$

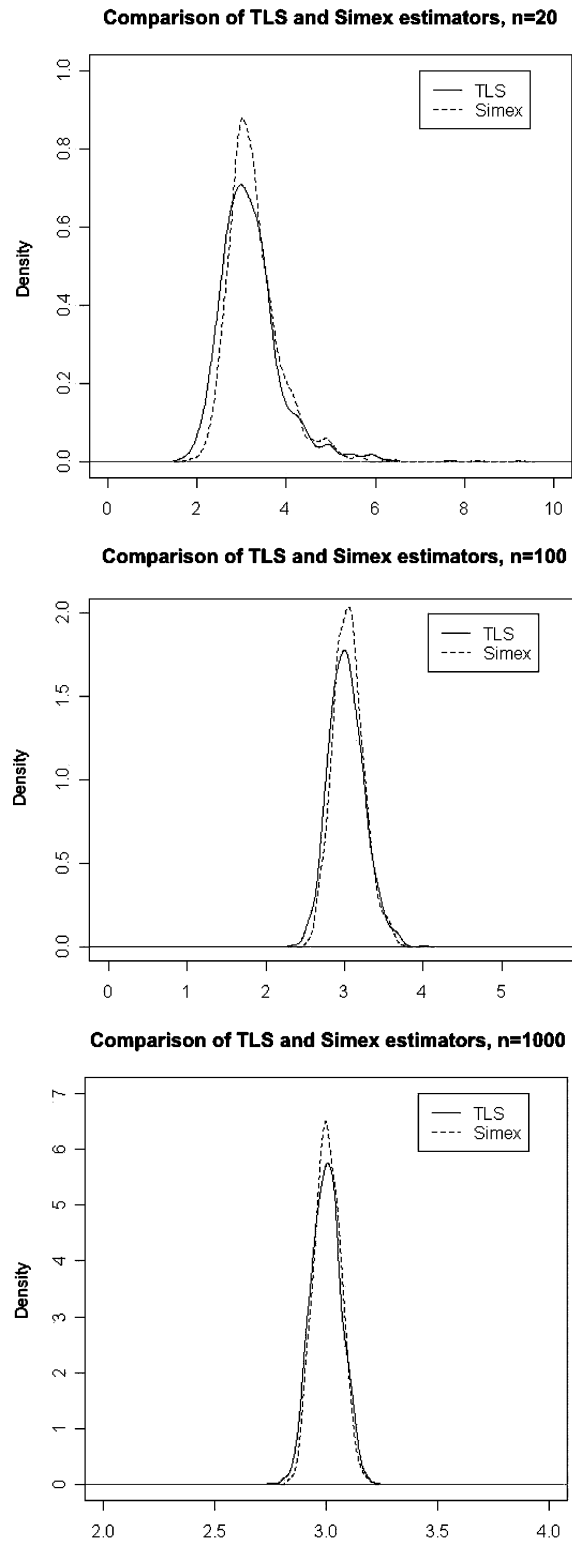


FIGURE 1. Comparison of SIMEX and total least squares estimators

Therefore the limit values of the SYMEX and TLS estimators coincide if $p = 1$. \square

Remark 5.1. The minimum of the function $K_n(\lambda)$ always exists in the compact set B . The point of minimum can be either the boundary point of B or its interior point. The point of minimum not necessarily coincides with the minimal eigenvalue of the limit matrix of observations.

6. SIMULATION

The simulation is performed with the help of the statistical software R. The slope in the linear model $y = 5 + 3x$ is estimated with the help of the SIMEX estimator. The variables are simulated as the normal random variables with zero mean and unit variance. The errors in variables and responses are simulated as centered normal random variables with the variance 0.25. We consider the behavior of the SIMEX estimator for small ($n = 20$), moderate ($n = 100$), and large ($n = 1000$) sizes of a sample of observations. The size of the bootstrap sample is $B = 100$. To estimate the parameter θ defined by (6), we use the following values of λ : 0, 0.25, 0.5, \dots , 3.5, 3.75, 4. The number of iterations is 1000 (see Figure 1).

The results of the simulation are exhibited by the empirical probability densities of SIMEX and the total least squares estimators. As seen from Figure 1, the distribution of the SIMEX estimator has lighter tails as compared to that of the total least squares estimator.

7. CONCLUDING REMARKS

We considered three modifications of the SIMEX method for the linear structural errors-in-variables model. We showed that these estimators are close to the known consistent estimators; namely, the estimator $\hat{\beta}_{\text{SIMEX}}$ has a structure similar to the method of moments estimator $\hat{\beta}_{\text{MME}}$, while the estimator $\hat{\beta}_{\text{ASIMEX}}$ behaves similarly to the total least squares estimator $\hat{\beta}_{\text{TLS}}$. It would be an interesting problem to study the SIMEX estimator for nonlinear models.

BIBLIOGRAPHY

1. P. Billingsley, *Convergence of Probability Measures*, John Wiley, New York, 1968. MR0233396 (38:1718)
2. R. J. Carroll, D. Ruppert, and L. A. Stefanski, *Measurement Error in Nonlinear Models*, Chapman and Hall, London, 1996. MR1630517 (2000c:62001)
3. C.-L. Cheng and John W. Van Ness, *Statistical Regression with Measurement Error*, Oxford University Press, New York, 1999. MR1719513 (2001k:62001)
4. J. R. Cook and L. A. Stefanski, *Simulation-extrapolation estimation in parametric measurement error models*, J. Amer. Stat. Assoc. **89** (1994), 1314–1328.
5. W. A. Fuller, *Measurement Error Models*, Wiley, New York, 1987. MR898653 (89a:62160)
6. J. Polzehl and S. Zwanzig, *SIMEX and TLS: An Equivalence Result*, Technical Report 999, WIAS, Berlin, 2005.

DEPARTMENT OF PROBABILITY THEORY AND MATHEMATICAL STATISTICS, FACULTY FOR MECHANICS AND MATHEMATICS, NATIONAL TARAS SHEVCHENKO UNIVERSITY, ACADEMICIAN GLUSHKOV AVENUE 6, KYIV 03127, UKRAINE

E-mail address: gontaro@ukr.net

Received 23/OCT/2006

Translated by N. SEMENOV