Michal Pešta

Strongly consistent estimation in dependent errors-in-variables

Acta Universitatis Carolinae. Mathematica et Physica, Vol. 52 (2011), No. 1, 69--79

Persistent URL: http://dml.cz/dmlcz/143669

Strongly Consistent Estimation in Dependent Errors-in-Variables

MICHAL PEŠTA

Praha

Received April 15, 2010
Revised June 14, 2010

Errors-in-variables (EIV) model with *dependent* errors is considered. A strong consistency of the *total least squares* (TLS) estimate for *weakly dependent* (α- and ϕ-mixing) measurements – encumbered with errors which are not necessarily stationary and identically distributed – is proved.

1. Introduction

The main goal of this paper is to establish results concerning consistency in linear relations, where measurement errors in input and output data occur simultaneously. Due to the fact that in some situations these disturbances cannot be considered as independent by nature, a proper model is required and, consequently, a suitable statistical inference needs to be derived.

1.1 Errors-in-variables model

Errors-in-variables (EIV) model

$$\underset{n\times 1}{\mathbf{Y}} = \underset{n\times p}{\mathbf{Z}}\,\underset{p\times 1}{\boldsymbol{\beta}} + \underset{n\times 1}{\boldsymbol{\varepsilon}} \quad \text{and} \quad \underset{n\times p}{\mathbf{X}} = \underset{n\times p}{\mathbf{Z}} + \underset{n\times p}{\boldsymbol{\Theta}} \tag{E}$$

is assumed, where $\boldsymbol{\beta}$ is a vector of *regression parameters* to be estimated, $\mathbf{X}$ and $\mathbf{Y}$ consist of *observable random* variables ($\mathbf{X}$ are covariates and $\mathbf{Y}$ is a response), $\mathbf{Z}$ consists of *unknown constants* and has full rank, and $\boldsymbol{\varepsilon}$ and $\boldsymbol{\Theta}$ are composed of

MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín

The present work was supported by the grants GAUK B-MAT 49809, SVV 261315/2010, and GAČR 201/09/0755.

Key words and phrases. Errors-in-variables, dependent errors, strong consistency

E-mail address: pesta@karlin.mff.cuni.cz

random errors such that the joint distribution of the elements of $[\mathbf{\Theta}, \boldsymbol{\varepsilon}]$ is absolutely continuous with respect to the Lebesgue measure.

1.2 Weak dependence

We do not restrict our model to independent observations; therefore, the dependence between measurement errors needs to be specified. It is assumed that $\{\xi_n\}_{n=1}^{\infty}$ is a sequence of random variables on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. For sub-$\sigma$-fields $\mathscr{A}, \mathscr{B} \subseteq \mathscr{F}$, we define

$$\alpha(\mathscr{A}, \mathscr{B}) := \sup_{A\in\mathscr{A}, B\in\mathscr{B}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|,$$
$$\varphi(\mathscr{A}, \mathscr{B}) := \sup_{A\in\mathscr{A}, B\in\mathscr{B}, \mathbb{P}(A)>0} |\mathbb{P}(B|A) - \mathbb{P}(B)|.$$

Henceforth, let us define a filtration $\mathscr{F}_m^n := \sigma(\xi_i, m \le i \le n)$.

There are many ways to describe weak dependence or, in other words, asymptotic independence of random variables (see [1]). In this paper we concentrate on two approaches. A sequence $\{\xi_n\}_{n=1}^{\infty}$ of random elements (i.e., vectors) is said to be *strong mixing* (α-mixing) if

$$\alpha(n) := \sup_{k\in\mathbb{N}} \alpha(\mathscr{F}_1^k, \mathscr{F}_{k+n}^{\infty}) \to 0, \quad n \to \infty; \tag{1}$$

moreover, it is said to be *uniformly strong mixing* (φ-mixing) if

$$\varphi(n) := \sup_{k\in\mathbb{N}} \varphi(\mathscr{F}_1^k, \mathscr{F}_{k+n}^{\infty}) \to 0, \quad n \to \infty. \tag{2}$$

Uniformly strong mixing – presented by Rosenblatt in [12] – implies strong mixing (see [10]), which was introduced by Ibragimov in [8].

1.3 Error structure

Proper distributional assumptions of random errors in the EIV model need to be proposed. Two levels of the error structure have to be distinguished. The first level of error structure – *within-individual level* – is that each row $[\mathbf{\Theta}_{i,\bullet}, \varepsilon_i]$ has *zero mean* and non-singular covariance matrix $\sigma^2\mathbf{I}$, where $\sigma^2 > 0$ is unknown (for simplicity). This assumption can be straightforwardly generalized as discussed in Section 4. Relationships between individual observations are represented by the second level of error structure – *between-individual level*. Here, rows $[\mathbf{\Theta}_{i,\bullet}, \varepsilon_i]$ are *weakly dependent*, e.g., α- or φ-mixing. This assumption is based on an idea of mutual influence between those measurements which are “close to each other”, influence themselves somehow. Moreover, this influence decreases as the distance between observations increases.

It has to be emphasized that no form of the errors’ *stationarity* is necessary to be assumed. Hence we strengthen our results by omitting this sometimes restrictive assumption strengthen our results.

Additional *design assumption* is necessary for asymptotics even in the case of independent errors:

$$\mathbf{\Delta} := \lim_{n\to\infty} n^{-1}\mathbf{Z}^\top\mathbf{Z} \quad \text{exists and is positive definite.} \tag{D}$$

Importance of the previous design assumption has already been thoroughly discussed in [11].

2. TLS estimation

Total least squares (TLS) estimate of the unknown parameter $\boldsymbol{\beta}$ was proposed in [6] as

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^\top\mathbf{X} - \lambda\mathbf{I})^{-1}\mathbf{X}^\top\mathbf{Y}, \tag{3}$$

where $\lambda \equiv \lambda_{p+1}([\mathbf{X},\mathbf{Y}]^\top[\mathbf{X},\mathbf{Y}])$ and $\lambda_q(\mathbf{A})$ means the qth largest eigenvalue of a square positive semidefinite matrix $\mathbf{A}$.

Strong consistency of the TLS estimate for independent errors is proved in [5]; moreover, weak consistency – again for independent errors – is discussed in [3]. When a premise of independence cannot be assumed, consistency of the TLS estimate under weak dependence of errors has to be explored.

3. Strong consistency

First of all, the strong law of large numbers (SLLN) for α-dependent non-identically distributed variables should be recalled.

Lemma 1 (SLLN for α-mixing) *Let $\{X_n\}_{n=1}^\infty$ be a sequence of α-mixing random variables satisfying*

$$\sup_{n\in\mathbb{N}} \mathbb{E}|X_n|^q < \infty \tag{4}$$

for some $q > 1$. Suppose that there exists $\delta > 0$ such that as $n \to \infty$,

$$\alpha(n) = \begin{cases} \mathcal{O}\left(n^{-\frac{q}{2q-2}-\delta}\right) & \text{if} \quad 1 < q < 2, \\ \mathcal{O}\left(n^{-\frac{2}{q}-\delta}\right) & \text{if} \quad q \geq 2. \end{cases} \tag{5}$$

Then

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n (X_i - \mathbb{E}X_i)}{n} = 0, \quad \textit{a.s.}$$

Proof: See [2, Theorem 1]. □

Additionally, the SLLN for φ-dependent non-identically distributed variables will be used below.

Lemma 2 (SLLN for φ-mixing) *Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of zero mean φ-mixing random variables satisfying*

$$\sum_{n=1}^{\infty} \sqrt{\varphi(n)} < \infty \tag{6}$$

and let $\{b_n\}_{n=1}^{\infty}$ be a non-decreasing unbounded sequence of positive numbers. Assume that

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}X_n^2}{b_n^2} < \infty, \tag{7}$$

then

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} X_i}{b_n} = 0, \quad a.s.$$

Proof: See [14, Theorem 4.1]. □

For a given random sequence $\xi_\circ \equiv \{\xi_n\}_{n=1}^{\infty}$ of random elements, the dependence coefficients $\alpha(n)$ will be denoted $\alpha(\xi_\circ, n)$. Analogous notation is used for φ-mixing sequences. Moreover, three auxiliary lemmas for an application of the SLLN for non-identically distributed random variables are stated to be used later.

The following lemma describes an asymptotic behaviour of α- and φ-mixing coefficients of the corresponding random sequences after a transformation. More precisely, a Borel transformation preserves the property of α- and φ-mixing and, moreover, sustains the rate of the mixing coefficients.

Lemma 3 *Suppose that for each $m = 1, 2, \ldots$, there is given a sequence of random variables denoted by $\xi^{(m)} := \{\xi_k^{(m)}\}_{k\in\mathbb{Z}}$. Suppose that the sequences $\xi^{(m)}$, $m = 1, 2, \ldots$ are independent of each other and that $h_k : \mathbb{R} \times \mathbb{R} \times \ldots \to \mathbb{R}$ is a Borel function for each $k \in \mathbb{Z}$. Define the sequence $\xi := \{\xi_k\}_{k\in\mathbb{Z}}$ of random variables by*

$$\xi_k := h_k\left(\xi_k^{(1)}, \xi_k^{(2)}, \ldots\right), \quad k \in \mathbb{Z}.$$

Then for each $n \geq 1$, the following statements hold:

(1) $\alpha(\xi, n) \leq \sum_{m=1}^{\infty} \alpha(\xi^{(m)}, n)$,

(2) $\varphi(\xi, n) \leq \sum_{m=1}^{\infty} \varphi(\xi^{(m)}, n)$.

Proof: See [1, Theorem 5.2]. □

Design assumption (D) can be seen as a convergence of a specific sum in the Cauchy sense, e.g., a limit of the averaged partial sums. The following technical lemma enables us to derive various implications of design assumption (D).

Lemma 4 *If $\lim_{n\to\infty} n^{-2+\delta} \sum_{i=1}^{n} a_i$ exists and is finite for some $\delta > 0$, then $\sum_{n=1}^{\infty} \frac{a_n}{n^2}$ is convergent.*

Proof: Due to Abel's partial summation [9, p. 1412], we have

$$\sum_{i=1}^{n} \frac{a_i}{i^2} = \sum_{i=1}^{n-1} \left(\sum_{j=1}^{i-1} a_j\right)\left(\frac{1}{i^2} - \frac{1}{(i+1)^2}\right) + \frac{1}{n^2} \sum_{i=1}^{n} a_i, \quad \forall n > 1. \tag{8}$$

If n tends to infinity, the last term of (8) tends to zero due to the lemma's assumption. Moreover, the infinite sum formed from the first summand on the right hand side of (8) is convergent if and only if

$$\sum_{i=1}^{\infty} i^{-3} \sum_{j=1}^{i-1} a_j = \sum_{i=1}^{\infty} i^{-1-\delta} \left(i^{-2+\delta} \sum_{j=1}^{i-1} a_j \right)$$

is convergent, but the right hand side of previous equation is convergent according to the Abel's convergence criterion ($i^{-2+\delta} \sum_{j=1}^{i-1} a_j \xrightarrow{i\to\infty} 0$). Hence, $\sum_{n=1}^{\infty} \frac{a_n}{n^2}$ converges as well. □

A convergence of matrices in Frobenius norm implies spectral convergence, which can be mathematically formalized in the following lemma.

Lemma 5 *If* $\{\mathbf{A}_n\}_{n=1}^{\infty}$ *and* $\{\mathbf{B}_n\}_{n=1}^{\infty}$ *are sequences of* $m \times m$ *matrices, then* $\|\mathbf{A}_n - \mathbf{B}_n\|_F \to 0$ *as* $n \to \infty$ *implies* $\lambda_m(\mathbf{A}_n) - \lambda_m(\mathbf{B}_n) \to 0$ *as* $n \to \infty$, *where* $\|\cdot\|_F$ *denotes Frobenius matrix norm.*

Proof: See [4, Lemma 2.3]. □

The preliminary statistical machinery will now to be used for deriving the main results of this paper – strong consistency of the TLS estimate. Besides the main consistency results, an estimate of *nuisance* parameter σ^2 is defined as $\widehat{\sigma}^2 := \lambda/n$ and its strong consistency is proved as well.

First, the TLS estimate is strongly consistent, assuming α-mixing errors in the EIV model.

Theorem 3.1 (α-mixing TLS strong consistency) *Let the EIV model hold and assumption* (D) *be satisfied. Suppose*

$$\{\Theta_{n,1}\}_{n=1}^{\infty}, \ldots, \{\Theta_{n,p}\}_{n=1}^{\infty}, \quad \textit{and} \quad \{\varepsilon_n\}_{n=1}^{\infty} \tag{9}$$

are independent sequences of α-mixing random variables having

$$\alpha(\Theta_{\circ,j}, n) = \mathcal{O}(n^{-q_j/(2q_j-2)-\delta_j}), \quad j = 1, \ldots, p, \tag{10}$$

and

$$\alpha(\varepsilon_{\circ}, n) = \mathcal{O}(n^{-q_{p+1}/(2q_{p+1}-2)-\delta_{p+1}}), \tag{11}$$

as $n \to \infty$ for some $\delta_j > 0$ and $1 < q_j \leq 2$, $j \in \{1, \ldots, p+1\}$. If

$$\sup_{n\in\mathbb{N}} Z_{n,j}^2 < \infty, \tag{12}$$

$$\sup_{n\in\mathbb{N}} \mathbb{E}|\Theta_{n,j}|^{2q_j} < \infty, \quad \textit{and} \quad \sup_{n\in\mathbb{N}} \mathbb{E}|\varepsilon_n|^{2q_{p+1}} < \infty \tag{13}$$

for each $j \in \{1, \ldots, p\}$, then

$$\lim_{n\to\infty} \widehat{\boldsymbol{\beta}} = \boldsymbol{\beta} \quad \textit{a.s.}, \tag{14}$$

$$\lim_{n\to\infty} \frac{\lambda}{n} = \sigma^2 \quad \textit{a.s.} \tag{15}$$

Proof: The estimate of $\boldsymbol{\beta}$ from (3) can be expressed as

$$\widehat{\boldsymbol{\beta}} = \left\{\mathbf{I} + (\mathbf{Z}^\top\mathbf{Z})^{-1}(\mathbf{Z}^\top\boldsymbol{\Theta} + \boldsymbol{\Theta}^\top\mathbf{Z} + \boldsymbol{\Theta}^\top\boldsymbol{\Theta} - n\sigma^2\mathbf{I} + (n\sigma^2 - \lambda)\mathbf{I})\right\}^{-1} \times (\mathbf{Z}^\top\mathbf{Z})^{-1}\left(\mathbf{Z}^\top\mathbf{Z}\boldsymbol{\beta} + \mathbf{Z}^\top\boldsymbol{\varepsilon} + \boldsymbol{\Theta}^\top\mathbf{Z}\boldsymbol{\beta} + \boldsymbol{\Theta}^\top\boldsymbol{\varepsilon}\right). \quad (16)$$

In order to prove (14), it is sufficient to show that

(1) $n^{-1}\mathbf{Z}^\top\boldsymbol{\Theta} \to \mathbf{0}$ a.s., $n \to \infty$;
(2) $n^{-1}\boldsymbol{\Theta}^\top\mathbf{Z} \to \mathbf{0}$ a.s., $n \to \infty$;
(3) $n^{-1}(\boldsymbol{\Theta}^\top\boldsymbol{\Theta} - n\sigma^2\mathbf{I}) \to \mathbf{0}$ a.s., $n \to \infty$;
(4) $n^{-1}(n\sigma^2 - \lambda) \to 0$ a.s., $n \to \infty$;
(5) $n^{-1}\mathbf{Z}^\top\boldsymbol{\varepsilon} \to \mathbf{0}$ a.s., $n \to \infty$;
(6) $n^{-1}\boldsymbol{\Theta}^\top\boldsymbol{\varepsilon} \to \mathbf{0}$ a.s., $n \to \infty$.

Note that

$$\sup_{n\in\mathbb{N}} \mathbb{E}|Z_{n,j}\Theta_{n,k}|^2 = \sigma^2 \sup_{n\in\mathbb{N}} Z_{n,j}^2 < \infty, \quad \forall j,k \in \{1,\ldots,p\}.$$

Moreover, Lemma 3(i) implies that $\alpha(Z_{\circ,j}\Theta_{\circ,k}, n) = \mathscr{O}(n^{-q_k/(2q_k-2)-\delta_k})$, which implies $\alpha(Z_{\circ,j}\Theta_{\circ,k}, n) = \mathscr{O}(n^{-1-\delta_k})$ for all $j,k \in \{1,\ldots,p\}$. Applying SLLN for α-mixing (Theorem 1), we have

$$n^{-1}\sum_{i=1}^{n} Z_{i,j}\Theta_{i,k} \xrightarrow{a.s.} 0, \quad n \to \infty, \quad \forall j,k \in \{1,\ldots,p\}.$$

Therefore, (i) holds and the similar arguments demonstrate (ii) and (v).

Again, it follows from Lemma 3(i) that $\alpha(\Theta_{\circ,j}\Theta_{\circ,k}, n) = \mathscr{O}(n^{-1-\delta_j\wedge\delta_k})$ for all $j,k \in$ $\in \{1,\ldots,p\}$ such that $j \neq k$. The supremum assumption of Theorem 1 is straightforwardly satisfied, because the independence from (9) provides

$$\sup_{n\in\mathbb{N}} \mathbb{E}|\Theta_{n,j}\Theta_{n,k}|^2 = \sup_{n\in\mathbb{N}} \mathbb{E}\Theta_{n,j}^2\mathbb{E}\Theta_{n,k}^2 = [\sigma^2]^2 < \infty$$

for all $j,k \in \{1,\ldots,p\}$, $j \neq k$. Hence, the SLLN for α-mixing yields

$$n^{-1}\sum_{i=1}^{n} \Theta_{i,j}\Theta_{i,k} \xrightarrow{a.s.} 0, \quad n \to \infty, \quad \forall j,k \in \{1,\ldots,p\},\ j \neq k.$$

Thus the "nondiagonal" part of (iii) is satisfied and, furthermore, the analogous arguments demonstrate (vi).

Consequently, $\alpha(\Theta_{\circ,j}^2, n) = \mathscr{O}(n^{-q_j/(2q_j-2)-\delta_j})$ for all $j \in \{1,\ldots,p\}$ according to Lemma 3(i). Since $\sup_{n\in\mathbb{N}} \mathbb{E}|\Theta_{n,j}^2|^{q_j} < \infty$ for all $j \in \{1,\ldots,p\}$, the SLLN for α-mixing can be applied

$$n^{-1}\sum_{i=1}^{n} \Theta_{i,j}^2 \xrightarrow{a.s.} \sigma^2, \quad n \to \infty, \quad \forall j \in \{1,\ldots,p\},$$

and the "diagonal" part of (iii) holds as well.

Now,

$$n^{-1}(\lambda - n\sigma^2) = \lambda_{p+1}(n^{-1}[\mathbf{X},\mathbf{Y}]^\top[\mathbf{X},\mathbf{Y}] - \sigma^2\mathbf{I})$$

due to the eigendecomposition property. Let $\mathbf{B} := n^{-1}[\mathbf{I},\boldsymbol{\beta}]^\top \mathbf{Z}^\top \mathbf{Z}[\mathbf{I},\boldsymbol{\beta}]$. For each $n \in \mathbb{N}$, $\mathbf{B}$ is a positive semidefinite matrix of rank p. Thus it has p positive eigenvalues and the smallest one being zero. Note that

$$\begin{aligned}
&n^{-1}([\mathbf{X},\mathbf{Y}]^\top[\mathbf{X},\mathbf{Y}] - n\sigma^2\mathbf{I}) - \mathbf{B} \\
&\quad = n^{-1}\left\{[\mathbf{I},\boldsymbol{\beta}]^\top \mathbf{Z}^\top[\boldsymbol{\Theta},\boldsymbol{\varepsilon}] + [\boldsymbol{\Theta},\boldsymbol{\varepsilon}]^\top \mathbf{Z}[\mathbf{I},\boldsymbol{\beta}]\right\} + n^{-1}\left\{[\boldsymbol{\Theta},\boldsymbol{\varepsilon}]^\top[\boldsymbol{\Theta},\boldsymbol{\varepsilon}] - n\sigma^2\mathbf{I}\right\}. \qquad (17)
\end{aligned}$$

The first term on the right hand side of equation (17) converges almost surely to zero due to (i), (ii), and (v). The second one converges almost surely to zero as well, using similar arguments as in (iii) and (vi). Furthermore, it follows from Lemma 5 that

$$\lambda_{p+1}(n^{-1}[\mathbf{X},\mathbf{Y}]^\top[\mathbf{X},\mathbf{Y}] - \sigma^2\mathbf{I}) \xrightarrow[n\to\infty]{a.s.} \lambda_{p+1}(\mathbf{B}) = 0,$$

which demonstrates (iv).

Finally, (iv) directly implies (15) and completes the proof. □

Similar to the above-mentioned assumption, φ-mixing errors yield the TLS estimate's strong consistency as well, but under slightly different assumptions.

Theorem 3.2 (φ-mixing TLS strong consistency) *Let the EIV model hold and assumption* (D) *be satisfied. Suppose*

$$\{\Theta_{n,1}\}_{n=1}^{\infty},\ldots,\{\Theta_{n,p}\}_{n=1}^{\infty}, \quad \textit{and} \quad \{\varepsilon_n\}_{n=1}^{\infty} \qquad (18)$$

are independent sequences of φ-mixing random variables such that

$$\sum_{n=1}^{\infty}\sqrt{\varphi(\Theta_{\circ,j},n)} < \infty \quad \textit{and} \quad \sum_{n=1}^{\infty}\sqrt{\varphi(\varepsilon_\circ,n)} < \infty. \qquad (19)$$

If

$$\sum_{n=1}^{\infty}\frac{\mathbb{E}\Theta_{n,j}^4}{n^2} < \infty \quad \textit{and} \quad \sum_{n=1}^{\infty}\frac{\mathbb{E}\varepsilon_n^4}{n^2} < \infty \qquad (20)$$

for each $j \in \{1,\ldots,p\}$, then

$$\lim_{n\to\infty}\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta} \quad a.s., \qquad (21)$$

$$\lim_{n\to\infty}\frac{\lambda}{n} = \sigma^2 \quad a.s. \qquad (22)$$

Proof: The process of proving this theorem is analogous to the proof of Theorem 3.1. The only difference is that the SLLN for φ-mixing is applied instead of the SLLN for α-mixing. Therefore, one does not have to be concerned with the supremum condition (4) and the dependence coefficient assumption (5) from Theorem 1. On the other hand, the convergence condition (6) on sum of the square roots of dependence coefficients $\varphi(n)$ and the convergence assumption (7) from Theorem 2 needs to be fulfilled.

Let us consider six terms of (16) from the proof of Theorem 3.1. It follows from Lemma 3(ii) that $\{Z_{n,j}\Theta_{n,k}\}_{n=1}^{\infty}$ is also a φ-mixing sequence for all $j,k \in \{1,\ldots,p\}$. Moreover,

$$\sum_{n=1}^{\infty} \sqrt{\varphi(Z_{\circ,j}\Theta_{\circ,k},n)} \leq \sum_{n=1}^{\infty} \sqrt{\varphi(\Theta_{\circ,k},n)} < \infty, \quad \forall j,k \in \{1,\ldots,p\}.$$

Assumption (D) implies

$$0 < n^{-1}\sum_{i=1}^{n} Z_{i,j}^2 \to \Delta_{j,j} < \infty, \quad n \to \infty, \quad \forall j \in \{1,\ldots,p\}. \tag{23}$$

Due to Lemma 4,

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\{Z_{n,j}\Theta_{n,k}\}^2}{n^2} = \sigma^2 \sum_{n=1}^{\infty} \frac{Z_{n,j}^2}{n^2} < \infty, \quad \forall j,k \in \{1,\ldots,p\},$$

which allows us to apply the SLLN for φ-mixing. Hence, (i) holds and the similar arguments provide (ii) and (v).

The rest of the proof is now pretty straightforward. In order to show (iii), (iv), and (vi), one has to realize that Lemma 3(ii) yields $\varphi(\xi_\circ^2,n) \leq \varphi(\xi_\circ,n)$, $\varphi(\xi_\circ\zeta_\circ,n) \leq$ $\leq \varphi(\xi_\circ,n) + \varphi(\zeta_\circ,n)$. Furthermore,

$$\sum_{n=1}^{\infty} \sqrt{\varphi(\xi_\circ\zeta_\circ,n)} \leq \sum_{n=1}^{\infty} \sqrt{\varphi(\xi_\circ,n)} + \sum_{n=1}^{\infty} \sqrt{\varphi(\zeta_\circ,n)} < \infty$$

for $\xi_n,\zeta_n \in \{\Theta_{n,1},\ldots,\Theta_{n,p},\varepsilon_n\}$, $\xi_n \neq \zeta_n$. Moreover, (20) holds and, due to the independence from (18),

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}\{\xi_n\zeta_n\}^2}{n^2} = \sum_{n=1}^{\infty} \frac{\mathbb{E}\xi_n^2\mathbb{E}\zeta_n^2}{n^2} = \frac{\pi^2\sigma^4}{6} < \infty,$$

for $\xi_n,\zeta_n \in \{\Theta_{n,1},\ldots,\Theta_{n,p},\varepsilon_n\}$, $\xi_n \neq \zeta_n$, which completes the proof. □

4. Conclusions

A linear EIV model with its TLS solution is considered in this paper. An error structure of the EIV model with weakly dependent errors is introduced. Strong laws of large numbers for strong mixing and uniformly strong mixing are summarized. They allow us to derive and prove a *strong consistency* of the TLS estimate under both forms of errors' *asymptotic independence*. Furthermore, no form of stationarity is imposed on the errors. In these settings, the strong consistency of the nuisance variance parameter is proved as well.

4.1 Discussion

A homoscedastic covariance structure of the within-individual errors $[\boldsymbol{\Theta}_{i,\bullet}, \varepsilon_i]$ can be generalized by knowing the *heteroscedastic covariance matrix* $\boldsymbol{\Sigma} > \mathbf{0}$ in advance. Then, the observation data are just multiplied by the inverse of its square root as already discussed in [5] or [13], i.e. $\boldsymbol{\Sigma}^{-1/2}[\mathbf{X}, \mathbf{Y}]^\top$. This transformation of the original data is purely linear, which is not restrictive at all in our case. The only property that needs to be satisfied is an independence of the transformed errors. The assumptions of independence (9) and (18) between errors on the within-individual level are crucial and cannot be omitted. Incorporating an extra form of weak dependence on the within-individual error level may be considered as well, but this could unfortunately require very complicated additional assumptions.

If the covariance matrix $\boldsymbol{\Sigma}$ is unknown, it can be estimated using *repeated observations*, but, a more complicated design of the experiment will be necessary:

$$\underset{n\times 1}{\mathbf{Y}_\iota} = \underset{n\times p}{\mathbf{Z}}\ \underset{p\times 1}{\boldsymbol{\beta}} + \underset{n\times 1}{\boldsymbol{\varepsilon}_\iota} \quad \text{and} \quad \underset{n\times p}{\mathbf{X}_\iota} = \underset{n\times p}{\mathbf{Z}} + \underset{n\times p}{\boldsymbol{\Theta}_\iota}, \quad \iota = 1, \ldots, r; \tag{24}$$

where $r \in \mathbb{N}$ stands for the number of *replications*. Extra information – needed for estimation of the covariance matrix – is added by the replications. Then, a general covariance matrix $\boldsymbol{\Sigma}$ for the within-individual errors can be estimated as in [7], e.g., by

$$\widehat{\boldsymbol{\Sigma}} := \frac{1}{n(r-1)} \sum_{i=1}^{r} \sum_{j=1}^{r} [\mathbf{X}_i, \mathbf{Y}_i]^\top \left[\left(\delta_{ij} - \frac{1}{r}\right)\mathbf{I}\right] [\mathbf{X}_j, \mathbf{Y}_j], \tag{25}$$

where δ_{ij} denotes Kronecker delta. Previous equation (25) can, using the notations of replication model (24), be rewritten as

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{nr} \sum_{i=1}^{r} [\boldsymbol{\Theta}_i, \boldsymbol{\varepsilon}_i]^\top [\boldsymbol{\Theta}_i, \boldsymbol{\varepsilon}_i] - \frac{1}{nr(r-1)} \sum_{i=1}^{r} \sum_{\substack{j=1 \\ j\neq i}}^{r} [\boldsymbol{\Theta}_i, \boldsymbol{\varepsilon}_i]^\top [\boldsymbol{\Theta}_j, \boldsymbol{\varepsilon}_j],$$

which illustrates the meaning of the estimate. Under some additional assumptions on "replicated" errors $[\boldsymbol{\Theta}_\iota, \boldsymbol{\varepsilon}_\iota] \in \mathbb{R}^{n\times(p+1)}$, $\iota = 1, \ldots, r$, the appropriate SLLN can be applied on $\widehat{\boldsymbol{\Sigma}}$ and, consequently, its consistency can be proved as well.

Heteroscedastic covariance structure of the within-individual errors can even be estimated without possessing repeated observations for each "individual," but a structure of the covariance matrix has to be predefined in advance according to some prior knowledge about the data dependence. E.g., if there is no reason to suppose that the error structure is changing over particular covariates and response, Toeplitz or AR1 covariance models are reasonable choices.

Moreover, if we *compare the assumptions* for α- and φ-mixing in our EIV model, α-mixing has weaker assumptions on dependence of the errors (every φ-mixing is

α-mixing, see e.g., [1]), but stronger on the design (α-mixing requires bounded moments of the errors). For φ-mixing, it is the other way around. Indeed, assumption (19) implies

$$\varphi(\Theta_{\circ,j}, n) = o(n^{-2}) \quad \text{and} \quad \varphi(\varepsilon_{\circ}, n) = o(n^{-2}).$$

Taking into account $\alpha(n) \leq \varphi(n)$ and supposing

$$\varphi(\Theta_{\circ,j}, n) = \mathcal{O}(n^{-2-\delta_j}), \quad \varphi(\varepsilon_{\circ}, n) = \mathcal{O}(n^{-2-\delta_{p+1}}),$$

assumptions (10) and (11) are satisfied for some $4/3 \leq q_i \leq 2$ $j \in \{1, \ldots, p+1\}$. On the other hand, assumption (13) with $q_i = 2$, $j \in \{1, \ldots, p+1\}$ clearly implies assumption (20). The choice of q_i is essential as well. Smaller q_is make assumption (13) more restrictive, but then assumptions (10) and (11) become more realizable.

Additional design assumption (12), which is necessary for proving strong consistency for α-mixing errors, may be viewed as a competitive one to the "basic" design assumption (D). These assumptions are not equivalent and neither of them implies the other one. On the other hand, assumption (12) can be considered as a *supplementary* assumption to assumption (D) in the following sense: (D) implies (23). Hence, Lemma 4 yields $Z_{n,j}^2 = o(n^2)$, $n \to \infty$ for all $j \in \{1, \ldots, p\}$, which is a weaker condition than the equiboundedness of $Z_{n,j}^2$ over all $n \in \mathbb{N}$ for all $j \in \{1, \ldots, p\}$ from (12).

Finally, if identically distributed rows of errorsare taken into account together with existence of their suitable moments, assumptions (13) and (20) are trivially satisfied. Then, a *strict stationarity* of the between-individual errors with an existence of the appropriate moments has to imply these assumptions as well. In other words, moment assumptions (13) and (20) cannot be considered as unattainable. Moreover, for strictly stationary errors even the supremum in definitions (1) and (2) can simply be avoided.

References

[1] Bradley, R. C.: *Basic properties of strong mixing conditions. A survey and some open questions.* Probab. Surv. **2** (2005), 107–144.

[2] Chen, X., Wu, Y.: *Strong law for mixing sequence.* Acta Math. Appl. Sin. **5** (4) (1989), 367–371.

[3] Gallo, P. P.: *Consistency of regression estimates when some variables are subject to error.* Comm. Statist. Theory Methods **11** (1982), 973–983.

[4] Gallo, P. P.: *Properties of estimators in Errors-in-variables models.* Ph.D. thesis, University of North Carolina, Chapel Hill, NC (1982).

[5] Gleser, L. J.: *Estimation in a multivariate "Errors in Variables" regression model: Large sample results.* Ann. Statist. **9** (1981), 24–44.

[6] Golub, G. H., Van Loan, C. F.: *An analysis of the total least squares problem.* SIAM J. Numer. Anal. **17** (6) (1980), 883–893.

[7] Healy, J. D.: *Estimation and tests for unknown linear restrictions in multivariate linear models.* Ph.D. Thesis, Purdue University (1975).

[8] Ibragimov, I. A.: *Some limit theorems for stochastic processes stationary in the strict sense.* Doklady Akademii Nauk SSSR **125** (1959), 711–714, (In Russian).

[9] Itō, K.: Encyclopedic Dictionary of Mathematics: The Mathematical Society of Japan. 2nd Edition, MIT Press, Cambridge, MA (1993).
[10] Lin, Z., Lu, C.: Limit Theory for Mixing Dependent Random Variables. Springer Verlag, New York (1997).
[11] Pešta, M.: *Total least squares and bootstrapping with application in calibration.* Submitted (2009).
[12] Rosenblatt, M.: *A central limit theorem and a strong mixing condition.* Proc. Natl. Acad. Sci. USA **42** (1956), 43–47.
[13] Van Huffel, S., Vandewalle, J.: The Total Least Squares Problem: Computational Aspects and Analysis. SIAM, Philadelphia, PA (1993).
[14] Xuejun, W., Shuhe, H., Yan, S., Wenzhi, Y.: *Moment inequalities for φ-mixing sequences and its applications*. J. Inequal. Appl. **2009** (2009), Article ID 379743, 12 pages.