AN ESTIMATION METHOD IN TIME SERIES ERRORS-IN-VARIABLES MODELS

Norihisa Tsuga*, Masanobu Taniguchi* and Madan L. Puri**

Suppose that we observe (x_t, y_t) from the errors-in-variables model: $x_t = \xi_t + \delta_t$, $y_t = \beta \xi_t + \varepsilon_t$, where $\{\delta_t\}$ and $\{\varepsilon_t\}$ are i.i.d. measurement errors. Here we assume that $\{\xi_t\}$ is a non-Gaussian stationary process with zero mean and spectral density $f_\xi(\lambda)$. For this model, some estimators for β have been proposed in the literature. However, they are constructed under the assumption that the data are independent normal variates. Thus they do not contain the dependent structure of the data (e.g., time-lagged sample covariances, etc.). In this paper we propose a new class $\mathcal{A}$ of estimators of β , which is defined under consideration for dependent structures of (x_t, y_t, ξ_t) . Then the asymptotic distribution of $\hat{\beta} \in \mathcal{A}$ is derived. We give an asymptotically optimal estimator in this class. Comparison with the existing estimators is also discussed. Since the asymptotic variance of $\hat{\beta}$ is complicated we have illuminated some aspect of the asymptotics numerically using Mathematica.

Key Words and Phrases: errors-in-variables model, stationary process, spectral density, nonparametric spectral estimator, asymptotic distribution.

1. Introduction

In many applications a relationship is conjectured between ξ_t of “cause” random variables and η_t of “effect” random variables. Instead of observing ξ_t and η_t directly we may observe instead

$$(1.1) \quad \begin{aligned} x_t &= \xi_t + \delta_t, \\ y_t &= \eta_t + \varepsilon_t, \end{aligned}$$

where $\{\delta_t\}$ is a sequence of i.i.d. measurement errors with mean zero and variance δ (we denote this by $\{\delta_t\} \sim i.i.d.(0, \delta)$), and $\{\varepsilon_t\} \sim i.i.d.(0, \varepsilon)$.

Setting $\eta_t = \beta_1 \xi_t$ in (1.1), Robinson (1977) considered the model

$$(1.2) \quad \begin{aligned} x_t &= \xi_t + \delta_t, \\ y_t &= \beta_1 \xi_t + \varepsilon_t, \end{aligned}$$

and relaxed the i.i.d. and Gaussian assumption on $(\xi_t, \delta_t, \varepsilon_t)$. In his paper, the variates ξ_t , δ_t and ε_t were assumed to be independent up to first moments (i.e., a kind of martingale differences), hence ξ_t may be a sequence of fixed constants or (possibly serially dependent) random variables.

Then he proposed estimators for unknown parameters based on the i.i.d. and Gaussian likelihood. Under a suitable identification assumption, the asymptotics of the estimators were established.

Received June, 1998. Revised March, 1999. Accepted October, 1999.

* Department of Mathematical Science, Faculty of Engineering Science, Osaka University, Japan.

** Department of Mathematics, Indiana University, Bloomington, IN 47405, U.S.A.

In the case of serially correlated observations Eltinge (1989) considered a multivariate version of the model (1.1) with $\eta_t = \beta_1 \xi_t + q_t$, where q_t is mean zero “error in equation” variable associated with the deviation of the term η_t from the linear function $\beta_1 \xi_t$. Assuming that $(\xi_t, \delta_t, \varepsilon_t, q_t)$ constitutes a vector non-Gaussian linear process, he studied the asymptotic properties of an estimator of the coefficient.

Although the results by Eltinge (1989) and Robinson (1977) are fairly general, their estimators are constructed under the assumption that the data are independent normal variates. So, their estimators do not contain dependent structures of the models (i.e., time-lagged sample covariance functions, etc.). In this paper we propose a new class Λ of estimators of β in (1.1) with $\eta_t = \beta \xi_t$, where $\{\xi_t\}$ is assumed to be a stationary process. Here the estimators $\hat{\beta}$ in Λ are defined under consideration for “dependent structures” of (x_t, y_t, ξ_t) .

In Section 3, we will derive the asymptotic distribution of $\hat{\beta}$ under appropriate conditions. An asymptotically optimal estimator in this class is given. It is shown that $\hat{\beta}$ has a merit in comparison with the existing estimators. That is, we can make a version $\hat{\beta}^*$ of $\hat{\beta}$ so that $\hat{\beta}^*$ improves Eltinge (1989)’s estimator in our model. Since the asymptotics of $\hat{\beta}$ are complicated, some numerical studies for the asymptotic variance are given. The results illuminate some interesting points.

2. Model and estimator

Suppose that the observed stretch $\{(x_t, y_t); t=1, \dots, n\}$ is generated by the model

$$(2.1) \quad \begin{aligned} x_t &= \xi_t + \delta_t, \\ y_t &= \eta_t + \varepsilon_t, \end{aligned}$$

with the relationship

$$\eta_t = \beta \xi_t,$$

where $\{\xi_t\}$, $\{\delta_t\}$ and $\{\varepsilon_t\}$ satisfy the following assumption.

ASSUMPTION 1. (i) $\{\xi_t\}$ is a stationary process with zero mean and spectral density $f_\xi(\lambda)$.

(ii) $\{\delta_t\}$ is a sequence of i.i.d. $(0, \delta)$ random variables with finite fourth-order moment.

(iii) $\{\varepsilon_t\}$ is a sequence of i.i.d. $(0, \varepsilon)$ random variables with finite fourth-order moment.

(iv) $\{\xi_t\}$, $\{\delta_t\}$ and $\{\varepsilon_t\}$ are mutually independent.

Then it is easily shown that $\{(x_t, y_t)'\}$ is a two-dimensional stationary process with zero mean vector and spectral density matrix

$$(2.2) \quad \mathbf{f}(\lambda) = \begin{pmatrix} f_x(\lambda) & f_{xy}(\lambda) \\ f_{yx}(\lambda) & f_y(\lambda) \end{pmatrix} = \begin{pmatrix} f_\xi(\lambda) + \frac{\delta}{2\pi} & \beta f_\xi(\lambda) \\ \beta f_\xi(\lambda) & \beta^2 f_\xi(\lambda) + \frac{\varepsilon}{2\pi} \end{pmatrix}.$$

We estimate $\mathbf{f}(\lambda)$ by the following nonparametric estimator

$$(2.3) \quad \hat{\mathbf{f}}_n(\lambda) \equiv \begin{pmatrix} \hat{f}_x(\lambda) & \hat{f}_{xy}(\lambda) \\ \hat{f}_{yx}(\lambda) & \hat{f}_y(\lambda) \end{pmatrix} \equiv \int_{-\pi}^{\pi} W_n(\lambda - \mu) \mathbf{I}_n(\mu) d\mu,$$

where $\mathbf{I}_n(\lambda) = 1/(2\pi n) \{ \sum_{t=1}^n (x_t, y_t)' e^{i\lambda t} \} \{ \sum_{t=1}^n (x_t, y_t) e^{-i\lambda t} \}$, and $W_n(\lambda)$ satisfies the following.

ASSUMPTION 2. (i) $W_n(\lambda)$ can be expanded as

$$W_n(\lambda) = \frac{1}{2\pi} \sum_{l=-M}^M w\left(\frac{l}{M}\right) e^{-il\lambda},$$

where $w(x)$ is a continuous, even function with $w(0) = 1$, and satisfies

$$|w(x)| \leq 1, \quad \int_{-\infty}^{\infty} w^2(x) dx < \infty, \quad \lim_{x \rightarrow \infty} \frac{1 - w(x)}{|x|^2} = \kappa_2 < \infty.$$

(ii) $M = M(n)$ satisfies

$$n^{\frac{1}{4}}/M + M/n^{\frac{1}{2}} \longrightarrow 0, \quad \text{as } n \rightarrow \infty.$$

We will show in Section 3 that $\hat{\mathbf{f}}_n(\lambda)$ becomes a consistent estimator of $\mathbf{f}(\lambda)$ under appropriate conditions. Since the parameters β, δ and ϵ are not identifiable we set down the following assumption.

ASSUMPTION 3. The parameter δ is known.

Recall $\mathbf{f}(\lambda)$ given in (2.2). Substituting $\hat{f}_x(\lambda)$ into $f_x(\lambda)$ in the (1, 1) component of $\mathbf{f}(\lambda)$, we obtain a consistent estimator $\hat{f}_{\epsilon}(\lambda) = \hat{f}_x(\lambda) - \delta/(2\pi)$ of $f_{\epsilon}(\lambda)$. Next, comparing both of the (1, 2) components of (2.2), we propose an estimator $\hat{\beta}_a$ of β by the value which minimizes

$$(2.4) \quad \int_{-\pi}^{\pi} a(\lambda) \{ \hat{f}_{xy}(\lambda) - \beta \hat{f}_{\epsilon}(\lambda) \}^2 d\lambda,$$

with respect to β , where $a(\lambda) \geq 0$ is a known L_2 -integrable function. Since (2.4) is a quadratic form of β , this estimator has the explicit form

$$(2.5) \quad \hat{\beta}_a = \frac{\int_{-\pi}^{\pi} a(\lambda) \hat{f}_{\epsilon}(\lambda) \hat{f}_{xy}(\lambda) d\lambda}{\int_{-\pi}^{\pi} a(\lambda) \hat{f}_{\epsilon}^2(\lambda) d\lambda}.$$

REMARK 2.1. In general $\hat{\beta}_a$ contains the sample autocovariances of $\{x_t\}$ and the sample cross-covariances of $\{(x_t, y_t)\}$ from zero lag to $\pm M$ lags. Thus we may say that $\hat{\beta}_a$ provides some suitable information about the dependent structure of (2.1).

3. The main results

In this section we elucidate the asymptotics of $\hat{\beta}_a$. For this we need further assumptions.

ASSUMPTION 4. (i) $\{\xi_t\}$ is a linear process generated as

$$(3.1) \quad \xi_t = \sum_{j=0}^{\infty} g(j) e_{t-j}, \quad \text{with } g(0)=1,$$

where $\{e_t\}$ satisfies $E[e_t]=0$, $E[e_t^2]=\sigma_e$ and $E[e_t e_s]=0$ for $t \neq s$.

(ii) The coefficients $g(j)$ satisfy

$$\sum_{j=0}^{\infty} j^2 |g(j)| < \infty.$$

(iii) The fourth-order cumulant of e_t satisfies

$$\text{cum}[e_t, e_{t+t_1}, e_{t+t_2}, e_{t+t_3}] = \begin{cases} \kappa_4^e, & \text{for } t_1=t_2=t_3=0, \\ 0, & \text{otherwise.} \end{cases}$$

(iv) There exists $c > 0$ such that

$$\left| \sum_{j=0}^{\infty} g(j) e^{ij\lambda} \right| \geq c, \quad \text{for all } \lambda \in [-\pi, \pi].$$

Regarding the nonparametric spectral estimator $\hat{f}_n(\lambda)$ we get the following fundamental result.

PROPOSITION 3.1. (e.g., Hannan (1970, p. 280)). Suppose that Assumptions 1-4 hold. Then

$$(3.2) \quad \hat{f}_n(\lambda) = f(\lambda) + O_p\left(\sqrt{\frac{M}{n}}\right),$$

uniformly in $\lambda \in [-\pi, \pi]$.

This proposition implies that $\hat{f}_n(\lambda)$ is a $\sqrt{n/M}$ -consistent estimator of $f(\lambda)$. To derive the asymptotic distribution of $\hat{\beta}_a$ we need a further assumption.

Let $\mathcal{F}_t$ be the σ -field generated by $\{e_s; s \leq t\}$. Henceforth we denote Kronecker's delta by $\delta(t, s)$.

ASSUMPTION 5. Let $\tau \in \mathcal{N}$, the set of all positive integers.

(i) For each s ,

$$\text{var}[E\{e_t e_{t+s} | \mathcal{F}_{t-\tau}\} - \delta(s, 0)\sigma_e] = O(\tau^{-2-\epsilon}), \quad \epsilon > 0,$$

uniformly in t .

(ii)

$$E|E\{e_{t_1} e_{t_2} e_{t_3} e_{t_4} | \mathcal{F}_{t_1-\tau}\} - E\{e_{t_1} e_{t_2} e_{t_3} e_{t_4}\}| = O(\tau^{-1-\gamma}),$$

uniformly in t_1 , where $t_1 \leq t_2 \leq t_3 \leq t_4$ and $\gamma > 0$.

(iii) For any $\rho > 0$ and any integer $L \geq 0$, there exists $\zeta_\rho > 0$ such that

$$E[T(n, s)^2 \chi\{T(n, s) > \zeta_\rho\}] < \rho,$$

uniformly in n and s , where $\chi\{\cdot\}$ is the indicator function and

$$T(n, s) = \left[\sum_{r=0}^L \left\{ \sum_{t=1}^n (e_{t+s} e_{t+s+r} - \sigma_e \delta(r, 0)) / \sqrt{n} \right\}^2 \right]^{\frac{1}{2}}.$$

Assumption 5 is a kind of “asymptotically” higher order martingale difference condition. It may be noted that the ordinary linear process generated by i.i.d. e_t with finite fourth order moment satisfies Assumption 5. Now we get the main theorem.

THEOREM 3.1 *Suppose that Assumptions 1-5 hold. Then $\sqrt{n}(\hat{\beta}_a - \beta)$ converges in distribution to the normal distribution with zero mean and variance*

$$(3.3) \quad v_a = \left[\int_{-\pi}^{\pi} a^2(\lambda) \left\{ (\beta^2 \delta + \varepsilon) f_{\varepsilon}^3(\lambda) + \frac{1}{2\pi} (2\beta^2 \delta^2 + \delta \varepsilon) f_{\varepsilon}^2(\lambda) \right\} d\lambda \right. \\ \left. + \frac{\beta^2 \kappa_4^{\delta}}{(2\pi)^2} \left\{ \int_{-\pi}^{\pi} a(\lambda) f_{\varepsilon}(\lambda) d\lambda \right\}^2 \right] \left[\int_{-\pi}^{\pi} a(\lambda) f_{\varepsilon}^2(\lambda) d\lambda \right]^{-2},$$

where κ_4^{δ} is the fourth order cumulant of δ_t .

We put the proof of this theorem together with those of Theorem 3.2 and 3.3 in Section 4.

COROLLARY 3.1. *Define*

$$\hat{\beta}_1 = \frac{\int_{-\pi}^{\pi} \hat{f}_{\varepsilon}(\lambda) \hat{f}_{xy}(\lambda) d\lambda}{\int_{-\pi}^{\pi} \hat{f}_{\varepsilon}^2(\lambda) d\lambda}.$$

Then, under Assumptions 1-5, $\sqrt{n}(\hat{\beta}_1 - \beta)$ converges in distribution to the normal distribution with zero mean and variance

$$v_1 = \left[\int_{-\pi}^{\pi} \left\{ (\beta^2 \delta + \varepsilon) f_{\varepsilon}^3(\lambda) + \frac{1}{2\pi} (2\beta^2 \delta^2 + \delta \varepsilon) f_{\varepsilon}^2(\lambda) \right\} d\lambda \right. \\ \left. + \frac{\beta^2 \kappa_4^{\delta}}{(2\pi)^2} \left\{ \int_{-\pi}^{\pi} f_{\varepsilon}(\lambda) d\lambda \right\}^2 \right] \left[\int_{-\pi}^{\pi} f_{\varepsilon}^2(\lambda) d\lambda \right]^{-2}.$$

REMARK 3.1. Here it should be noted that $\sqrt{n}$ -consistency holds in Theorem 3.1 and Corollary 3.1 despite Proposition 3.1. This is due to the fact that integration of $\hat{f}_n(\lambda)$ recovers $\sqrt{n}$ -consistency.

Let us introduce the following class of estimators

$$\Lambda = \left\{ \hat{\beta}_a: a = a(\lambda) \geq 0 \quad \text{and} \quad \int_{-\pi}^{\pi} |a(\lambda)|^2 d\lambda < \infty \right\}.$$

The next theorem gives the lower bound of the asymptotic variance of estimators in Λ .

THEOREM 3.2. *Suppose that $\kappa_4^{\delta} = 0$. The lower bound of v_a in (3.3) is given by*

$$(3.4) \quad \left[\int_{-\pi}^{\pi} \frac{f_{\epsilon}(\lambda)^2}{(\beta^2 \delta + \epsilon) f_{\epsilon}(\lambda) + \frac{1}{2\pi} (2\beta^2 \delta^2 + \delta \epsilon)} d\lambda \right]^{-1} \stackrel{\text{say}}{=} v_{\text{opt}}.$$

Here the lower bound is attained if and only if

$$(3.5) \quad a(\lambda) = \frac{C}{(\beta^2 \delta + \epsilon) f_{\epsilon}(\lambda) + \frac{1}{2\pi} (2\beta^2 \delta^2 + \delta \epsilon)}, \quad a.s.(\lambda),$$

where $C(\neq 0)$ is a constant.

The optimal weight function (3.5) depends on the unknown values β , ϵ and $f_{\epsilon}(\lambda)$. Next we consider to estimate it. Recalling the equation (2.2) we estimate ϵ by the value $\hat{\epsilon}$ which minimizes

$$\int_{-\pi}^{\pi} \left\{ \hat{f}_y(\lambda) - \hat{\beta}_1^2 \hat{f}_{\epsilon}(\lambda) - \frac{\epsilon}{2\pi} \right\}^2 d\lambda,$$

with respect to ϵ . Evidently the estimator $\hat{\epsilon}$ is given by

$$\hat{\epsilon} = \int_{-\pi}^{\pi} \{ \hat{f}_y(\lambda) - \hat{\beta}_1^2 \hat{f}_{\epsilon}(\lambda) \} d\lambda.$$

Hence we use

$$\hat{a}(\lambda) = \left[(\hat{\beta}_1^2 \delta + \hat{\epsilon}) \hat{f}_{\epsilon}(\lambda) + \frac{1}{2\pi} (2\hat{\beta}_1^2 \delta^2 + \delta \hat{\epsilon}) \right]^{-1},$$

as an estimator of the optimal weight function. Define

$$\hat{\beta} = \frac{\int_{-\pi}^{\pi} \hat{a}(\lambda) \hat{f}_{\epsilon}(\lambda) \hat{f}_{xy}(\lambda) d\lambda}{\int_{-\pi}^{\pi} \hat{a}(\lambda) \hat{f}_{\epsilon}^2(\lambda) d\lambda}.$$

THEOREM 3.3. Suppose that $\kappa_1^2 = 0$. Then, under Assumptions 1-5, $\sqrt{n}(\hat{\beta} - \beta)$ converges in distribution to the normal distribution with zero mean and variance v_{opt} . Therefore, $\hat{\beta}$ becomes an optimal estimator of β .

REMARK 3.2. Let $\hat{\beta}_{\text{ELT}}$ be the estimator of β proposed by Theorem 3.3 of Eltinge (1990). Then it is not difficult to show that $\hat{\beta}_{\text{ELT}}$ is $\hat{\beta}_a$ with $a(\lambda) = \hat{f}_{\epsilon}(\lambda)^{-1}$. Strictly speaking this $\hat{\beta}_a$ does not belong to Λ . We write $\hat{f}_{xy}(\lambda)$ and $\hat{f}_{\epsilon}(\lambda)$ in the form

$$(3.6) \quad \hat{f}_{xy}(\lambda) = \sum_{l=-M}^M w_l^{(1)} \hat{R}_{xy}(l) e^{-il\lambda},$$

$$(3.7) \quad \hat{f}_{\epsilon}(\lambda) = \sum_{l=-M}^M w_l^{(2)} \hat{R}_{xx}(l) e^{-il\lambda} - \frac{\delta}{2\pi},$$

where $\{w_l^{(1)}\}$ and $\{w_l^{(2)}\}$ are appropriate weights, and

$$\hat{R}_{xy}(l) = \frac{1}{n} \sum_{t=1}^{n-|l|} x_t y_{t+|l|},$$

$$\hat{R}_{xx}(l) = \frac{1}{n} \sum_{t=1}^{n-|l|} x_t x_{t+|l|}.$$

Since

$$\hat{\beta}_{\text{ELT}} = \int_{-\pi}^{\pi} \hat{f}_{xy}(\lambda) d\lambda / \int_{-\pi}^{\pi} \hat{f}_x(\lambda) d\lambda = \frac{\hat{R}_{xy}(0)}{\{\hat{R}_{xx}(0) - \delta/2\pi\}},$$

$\hat{\beta}_{\text{ELT}}$ does not contain the lagged sample covariances $\hat{R}_{xy}(l)$, $l \neq 0$ and $\hat{R}_{xx}(l)$, $l \neq 0$. In view of Theorems 3.2 and 3.3, we can see that our estimator $\hat{\beta}$ always improves $\hat{\beta}_{\text{ELT}}$.

4. Numerical studies

We next investigate the numerical behavior of the lower limit (3.4) and the optimal weight function $a(\lambda)$ given by (3.5). Assume $\kappa_4^\delta = 0$, and

$$f_\varepsilon(\lambda) = \frac{1}{2\pi} \frac{1}{|1 - ue^{i\lambda}|^2}, \quad (\text{AR}(1)).$$

Consider the lower limit given in Theorem 3.2. Then we get

Figure 1. v_a in (4.1) for $\delta = \varepsilon = 0.1(0.3)1.0$

Figure 2. $a(\lambda)$ in (4.2) for $\delta = \varepsilon = 0.1(0.3)1.0$

$$(4.1) \quad v_a = (\beta^2 \delta + \varepsilon) \left[\frac{1}{(1-u^2)} - \frac{(2\beta^2 \delta^2 + \delta \varepsilon)}{\sqrt{\{(\beta^2 \delta + \varepsilon) + (2\beta^2 \delta^2 + \delta \varepsilon) + (2\beta^2 \delta^2 + \delta \varepsilon)u^2\}^2 - 4(2\beta^2 \delta^2 + \delta)^2 u^2}} \right]^{-1}.$$

In Figure 1, we plotted v_a in (4.1) for $\delta = \varepsilon = 0.1(0.3)1.0$ and for $(\beta, u) \in [0, 1] \times [-0.999, 0.999]$.

Also, in Figure 2, we plotted the optimal weight functions

$$(4.2) \quad a(\lambda) = \frac{1}{(\beta^2 \delta + \varepsilon) f_{\varepsilon}(\lambda) + \frac{1}{2\pi} (2\beta^2 \delta^2 + \delta \varepsilon)},$$

for $\beta = 0.4$, $\delta = \varepsilon = 0.1(0.3)1.0$ and for $(\lambda, u) \in [-\pi, \pi] \times [-0.999, 0.999]$.

In Figure 1, if measurement errors become large, the variance v_a becomes large. If $|u| \rightarrow 1$, then $v_a \rightarrow 0$, which means that if $\{\xi_t\}$ has a root near ± 1 then the estimator becomes good. In Figure 2, if measurement errors become large, $a(\lambda)$ becomes flat, which implies that if measurement errors are small, we need to adjust the weight $a(\lambda)$.

5. Proofs of theorems

In this section we give the proofs of the theorems in Section 3.

PROOF OF THEOREM 3.1. Let

$$\frac{C}{D} \equiv \frac{\int_{-\pi}^{\pi} a(\lambda) \hat{f}_{\epsilon}(\lambda) \hat{f}_{xy}(\lambda) d\lambda}{\int_{-\pi}^{\pi} a(\lambda) \hat{f}_{\epsilon}^2(\lambda) d\lambda} = \hat{\beta}_a,$$

and

$$\frac{c}{d} \equiv \frac{\int_{-\pi}^{\pi} a(\lambda) f_{\epsilon}(\lambda) f_{xy}(\lambda) d\lambda}{\int_{-\pi}^{\pi} a(\lambda) f_{\epsilon}^2(\lambda) d\lambda} = \beta.$$

Recalling Proposition 3.1, and using a generalization of Corollaries 5.1.5 and 5.1.6 of Fuller (1996) (see pp.225-226) we can see that

$$\begin{aligned} (5.1) \quad C - c &= \int_{-\pi}^{\pi} a(\lambda) \hat{f}_{\epsilon}(\lambda) \hat{f}_{xy}(\lambda) d\lambda - \int_{-\pi}^{\pi} a(\lambda) f_{\epsilon}(\lambda) f_{xy}(\lambda) d\lambda \\ &= \int_{-\pi}^{\pi} a(\lambda) f_{\epsilon}(\lambda) \{ \hat{f}_{xy}(\lambda) - f_{xy}(\lambda) \} d\lambda \\ &\quad + \int_{-\pi}^{\pi} a(\lambda) \{ \hat{f}_{\epsilon}(\lambda) - f_{\epsilon}(\lambda) \} f_{xy}(\lambda) d\lambda + o_p\left(\frac{1}{\sqrt{n}}\right) \\ &= \int_{-\pi}^{\pi} a(\lambda) f_{\epsilon}(\lambda) \{ \hat{f}_{xy}(\lambda) - f_{xy}(\lambda) \} d\lambda \\ &\quad + \int_{-\pi}^{\pi} a(\lambda) \{ \hat{f}_x(\lambda) - f_x(\lambda) \} f_{xy}(\lambda) d\lambda + o_p\left(\frac{1}{\sqrt{n}}\right) \\ (5.2) \quad &\stackrel{\text{say}}{=} l_c + o_p\left(\frac{1}{\sqrt{n}}\right), \end{aligned}$$

and

$$\begin{aligned} (5.3) \quad D - d &= \int_{-\pi}^{\pi} a(\lambda) \{ \hat{f}_{\epsilon}^2(\lambda) - f_{\epsilon}^2(\lambda) \} d\lambda \\ &= 2 \int_{-\pi}^{\pi} a(\lambda) f_{\epsilon}(\lambda) \{ \hat{f}_{\epsilon}(\lambda) - f_{\epsilon}(\lambda) \} d\lambda + o_p\left(\frac{1}{\sqrt{n}}\right) \\ &= 2 \int_{-\pi}^{\pi} a(\lambda) f_{\epsilon}(\lambda) \{ \hat{f}_x(\lambda) - f_x(\lambda) \} d\lambda + o_p\left(\frac{1}{\sqrt{n}}\right) \\ (5.4) \quad &\stackrel{\text{say}}{=} l_d + o_p\left(\frac{1}{\sqrt{n}}\right). \end{aligned}$$

Taniguchi, Puri and Kondo (1996) established that $l_c = O_p(1/\sqrt{n})$ and $l_d = O_p(1/\sqrt{n})$. Thus we have

$$\begin{aligned} (5.5) \quad \hat{\beta}_a - \beta &= \frac{C}{D} - \frac{c}{d} \\ &= (C - c) \frac{\partial}{\partial C} \left(\frac{C}{D} \right) \Big|_{C=c, D=d} + (D - d) \frac{\partial}{\partial D} \left(\frac{C}{D} \right) \Big|_{C=c, D=d} + O_p\left(\frac{1}{n}\right) \\ &= (C - c) \frac{1}{d} - (D - d) \left(\frac{c}{d^2} \right) + O_p\left(\frac{1}{n}\right) \end{aligned}$$

from (5.2) and (5.4). It follows from (5.1), (5.3) and (5.5) that

$$\begin{aligned}
 (5.6) \quad \sqrt{n}(\hat{\beta}_a - \beta) &= \frac{1}{d} \left[\sqrt{n} \int_{-\pi}^{\pi} a(\lambda) f_{\epsilon}(\lambda) \{ \hat{f}_{xy}(\lambda) - f_{xy}(\lambda) \} d\lambda \right. \\
 &\quad \left. + \sqrt{n} \int_{-\pi}^{\pi} a(\lambda) \{ \hat{f}_x(\lambda) - f_x(\lambda) \} f_{xy}(\lambda) d\lambda \right] \\
 &\quad - \frac{c}{d^2} \left[2\sqrt{n} \int_{-\pi}^{\pi} a(\lambda) f_{\epsilon}(\lambda) \{ \hat{f}_x(\lambda) - f_x(\lambda) \} d\lambda \right] + o_p(1) \\
 &= \frac{1}{d} \left[\sqrt{n} \int_{-\pi}^{\pi} a(\lambda) f_{\epsilon}(\lambda) \{ \hat{f}_{xy}(\lambda) - f_{xy}(\lambda) \} d\lambda \right] \\
 &\quad - \frac{1}{d} \left[\sqrt{n} \int_{-\pi}^{\pi} a(\lambda) \{ \hat{f}_x(\lambda) - f_x(\lambda) \} d\lambda \right] + o_p(1).
 \end{aligned}$$

The asymptotic normality of $\sqrt{n}(\hat{\beta}_a - \beta)$ follows from Theorem 1 of Taniguchi, Puri and Kondo (1996). Thus we next evaluate the asymptotic variance of $\sqrt{n}(\hat{\beta}_a - \beta)$. The derivation is based on Hosoya and Taniguchi (1982), Taniguchi (1987) and Taniguchi, Puri and Kondo (1996).

First we evaluate the Gaussian part of asymptotic variance of $\sqrt{n}(\hat{\beta}_a - \beta)$. The Gaussian part of asymptotic variances of the first and second terms in (5.6) are

$$\begin{aligned}
 &\frac{2\pi}{d^2} \int_{-\pi}^{\pi} a^2(\lambda) f_{\epsilon}^2(\lambda) \{ f_x(\lambda) f_y(\lambda) + f_{xy}^2(\lambda) \} d\lambda, \\
 &\frac{4\pi}{d^2} \int_{-\pi}^{\pi} a^2(\lambda) f_x^2(\lambda) f_{xy}^2(\lambda) d\lambda,
 \end{aligned}$$

respectively, and the Gaussian part of asymptotic covariance between the first and second terms in (5.6) is

$$-2 \cdot \frac{4\pi}{d^2} \int_{-\pi}^{\pi} a^2(\lambda) f_{\epsilon}(\lambda) f_x(\lambda) f_{xy}^2(\lambda) d\lambda.$$

Therefore, the Gaussian part of asymptotic variance of $\sqrt{n}(\hat{\beta}_a - \beta)$ is

$$\begin{aligned}
 &\frac{2\pi}{d^2} \int_{-\pi}^{\pi} a^2(\lambda) \{ f_{\epsilon}^2(\lambda) f_x(\lambda) f_y(\lambda) + f_{\epsilon}^2(\lambda) f_{xy}^2(\lambda) + 2f_x^2(\lambda) f_{xy}^2(\lambda) \\
 &\quad - 4f_{\epsilon}(\lambda) f_x(\lambda) f_{xy}^2(\lambda) \} d\lambda \\
 &= \frac{1}{d^2} \int_{-\pi}^{\pi} a^2(\lambda) \left\{ (\beta^2 \delta + \epsilon) f_{\epsilon}^3(\lambda) + \frac{1}{2\pi} (2\beta^2 \delta^2 + \delta \epsilon) f_{\epsilon}^2(\lambda) \right\} d\lambda.
 \end{aligned}$$

Second, we evaluate the non-Gaussian part of the asymptotic variance of $\sqrt{n}(\hat{\beta}_a - \beta)$. We can rewrite the model as

$$\begin{aligned}
 x_t &= \xi_t + \delta_t = L(B)e_t + \delta_t, \\
 y_t &= \beta_t \xi_t + \epsilon_t = \beta L(B)e_t + \epsilon_t,
 \end{aligned}$$

where

$$L(B) \equiv \sum_{j=0}^{\infty} g(j) B^j, \quad (Be_t = e_{t-1}).$$

Also, we can write

$$z_t \equiv \begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} L(B) & 1 & 0 \\ \beta L(B) & 0 & 1 \end{pmatrix} \begin{pmatrix} e_t \\ \delta_t \\ \varepsilon_t \end{pmatrix}.$$

Denote the frequency response matrix of $\begin{pmatrix} L(B) & 1 & 0 \\ \beta L(B) & 0 & 1 \end{pmatrix}$ by

$$\begin{pmatrix} k_{11}(\lambda) & k_{12}(\lambda) & k_{13}(\lambda) \\ k_{21}(\lambda) & k_{22}(\lambda) & k_{23}(\lambda) \end{pmatrix}.$$

Evidently, $\{e_t\}$, $\{\delta_t\}$ and $\{\varepsilon_t\}$ have, respectively, the fourth-order cumulant spectral densities

$$\tilde{Q}_{1111}^e(\lambda_1, \lambda_2, \lambda_3) = \frac{\kappa_4^e}{(2\pi)^3},$$

$$\tilde{Q}_{2222}^\delta(\lambda_1, \lambda_2, \lambda_3) = \frac{\kappa_4^\delta}{(2\pi)^3},$$

$$\tilde{Q}_{3333}^\varepsilon(\lambda_1, \lambda_2, \lambda_3) = \frac{\kappa_4^\varepsilon}{(2\pi)^3}.$$

By Lemma A 2.1 of Hosoya and Taniguchi (1982), the components of the fourth-order cumulant spectral density $\{z_t\}$ are given by

$$\tilde{Q}_{1111}^z(\lambda_1, \lambda_2, \lambda_3) = k_{11}(\lambda_1 + \lambda_2 + \lambda_3)k_{11}(-\lambda_1)k_{11}(-\lambda_2)k_{11}(-\lambda_3)\frac{\kappa_4^e}{(2\pi)^3} + \frac{\kappa_4^\delta}{(2\pi)^3},$$

$$\tilde{Q}_{1112}^z(\lambda_1, \lambda_2, \lambda_3) = \beta k_{11}(\lambda_1 + \lambda_2 + \lambda_3)k_{11}(-\lambda_1)k_{11}(-\lambda_2)k_{11}(-\lambda_3)\frac{\kappa_4^e}{(2\pi)^3},$$

$$\tilde{Q}_{1211}^z(\lambda_1, \lambda_2, \lambda_3) = \beta k_{11}(\lambda_1 + \lambda_2 + \lambda_3)k_{11}(-\lambda_1)k_{11}(-\lambda_2)k_{11}(-\lambda_3)\frac{\kappa_4^e}{(2\pi)^3},$$

$$\tilde{Q}_{1212}^z(\lambda_1, \lambda_2, \lambda_3) = \beta^2 k_{11}(\lambda_1 + \lambda_2 + \lambda_3)k_{11}(-\lambda_1)k_{11}(-\lambda_2)k_{11}(-\lambda_3)\frac{\kappa_4^e}{(2\pi)^3}.$$

Noting that $k_{11}(\lambda)k_{11}(-\lambda) = (2\pi/\sigma_e)f_\varepsilon(\lambda)$, we can see that

$$\tilde{Q}_{1111}^z(\lambda_1, \lambda_2, -\lambda_2) = \frac{\kappa_4^\delta}{2\pi\sigma_e^2}f_\varepsilon(\lambda_1)f_\varepsilon(\lambda_2) + \frac{\kappa_4^e}{(2\pi)^3},$$

$$\tilde{Q}_{1112}^z(\lambda_1, \lambda_2, -\lambda_2) = \tilde{Q}_{1211}^z(\lambda_1, \lambda_2, -\lambda_2) = \frac{\beta\kappa_4^e}{2\pi\sigma_e^2}f_\varepsilon(\lambda_1)f_\varepsilon(\lambda_2),$$

$$\tilde{Q}_{1212}^z(\lambda_1, \lambda_2, -\lambda_2) = \frac{\beta^2\kappa_4^e}{2\pi\sigma_e^2}f_\varepsilon(\lambda_1)f_\varepsilon(\lambda_2).$$

Then, using Lemma A 2.2 of Hosoya and Taniguchi (1982) we can show that the non-Gaussian part of the asymptotic variances of the first and second terms in (5.6) are

$$\begin{aligned} & 2\pi \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{a(\lambda_1)f_\varepsilon(\lambda_1)}{d} \frac{a(\lambda_2)f_\varepsilon(\lambda_2)}{d} \frac{\beta^2\kappa_4^e}{2\pi\sigma_e^2} f_\varepsilon(\lambda_1)f_\varepsilon(\lambda_2) d\lambda_1 d\lambda_2 \\ &= \left\{ \int_{-\pi}^{\pi} \frac{a(\lambda)f_\varepsilon(\lambda)}{d} \frac{\beta\sqrt{\kappa_4^e}}{\sigma_e} f_\varepsilon(\lambda) d\lambda \right\}^2, \end{aligned}$$

and

$$2\pi \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{a(\lambda_1)f_{xy}(\lambda_1)}{d} \frac{a(\lambda_2)f_{xy}(\lambda_2)}{d} \left\{ \frac{\kappa_4^e}{2\pi\sigma_e^2} f_{\epsilon}(\lambda_1)f_{\epsilon}(\lambda_2) + \frac{\kappa_4^{\delta}}{(2\pi)^3} \right\} d\lambda_1 d\lambda_2 \\ = \left\{ \int_{-\pi}^{\pi} \frac{a(\lambda)f_{\epsilon}(\lambda)}{d} \frac{\beta\sqrt{\kappa_4^e}}{\sigma_e} f_{\epsilon}(\lambda) d\lambda \right\}^2 + \left\{ \int_{-\pi}^{\pi} \frac{a(\lambda)f_{\epsilon}(\lambda)}{d} \frac{\beta\sqrt{\kappa_4^{\delta}}}{2\pi} d\lambda \right\}^2,$$

respectively, and that the non-Gaussian part of the asymptotic covariance between the first term and the second term in (5.6) is

$$-2 \cdot 2\pi \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{a(\lambda_1)f_{\epsilon}(\lambda_1)}{d} \frac{a(\lambda_2)f_{xy}(\lambda_2)}{d} \frac{\beta\kappa_4^e}{2\pi\sigma_e^2} f_{\epsilon}(\lambda_1)f_{\epsilon}(\lambda_2) d\lambda_1 d\lambda_2 \\ = -2 \left\{ \int_{-\pi}^{\pi} \frac{a(\lambda)f_{\epsilon}(\lambda)}{d} \frac{\beta\sqrt{\kappa_4^e}}{\sigma_e} d_{\epsilon}(\lambda) d\lambda \right\}^2.$$

Therefore, the non-Gaussian part of the asymptotic variance of $\sqrt{n}(\hat{\beta}_a - \beta)$ is given by

$$\frac{\beta^2 \kappa_4^{\delta}}{(2\pi d)^2} \left\{ \int_{-\pi}^{\pi} a(\lambda)f_{\epsilon}(\lambda) d\lambda \right\}^2.$$

Summarizing the above the asymptotic variance of $\sqrt{n}(\hat{\beta}_a - \beta)$ is

$$\frac{1}{d^2} \left[\int_{-\pi}^{\pi} a^2(\lambda) \left\{ (\beta^2 \delta + \epsilon) f_{\epsilon}^3(\lambda) + \frac{1}{2\pi} (2\beta^2 \delta^2 + \delta \epsilon) f_{\epsilon}^2(\lambda) \right\} d\lambda \right. \\ \left. + \frac{\beta^2 \kappa_4^{\delta}}{(2\pi)^2} \left\{ \int_{-\pi}^{\pi} a(\lambda)f_{\epsilon}(\lambda) d\lambda \right\}^2 \right],$$

which completes the proof. □

PROOF OF THEOREM 3.2. Put

$$A(\lambda) = a(\lambda)f_{\epsilon}(\lambda), \\ B(\lambda) = f_{\epsilon}(\lambda), \\ g(\lambda) = (\beta^2 \delta + \epsilon)f_{\epsilon}(\lambda) + \frac{1}{2\pi} (2\beta^2 \delta^2 + \delta \epsilon).$$

Then, by Schwarz's inequality we have

$$(5.7) \quad \frac{\int_{-\pi}^{\pi} A(\lambda)^2 g(\lambda) d\lambda}{\left\{ \int_{-\pi}^{\pi} A(\lambda) B(\lambda) d\lambda \right\}^2} \geq \left\{ \int_{-\pi}^{\pi} \frac{B(\lambda)^2}{g(\lambda)} d\lambda \right\}^{-1},$$

where the equality holds if and only if there exists a constant $C \neq 0$ such that

$$g(\lambda)A(\lambda) + CB(\lambda) = 0, \quad \text{a.s. } (\lambda).$$

Therefore, the proof follows from (5.7). □

PROOF OF THEOREM 3.3. First, $\sqrt{n}$ -consistency of $\hat{\epsilon}$ follows from the same arguments as in Theorem 3.1. Hence $\hat{a}(\lambda)$ becomes a $\sqrt{n/M}$ -consistent estimator of $a(\lambda)$. From the definition of $\hat{\beta}$ we can see

$$\begin{aligned}\sqrt{n}(\hat{\beta} - \beta) &= \sqrt{n}(\hat{\beta}_a - \beta) \frac{\int_{-\pi}^{\pi} a(\lambda) \hat{f}_{\epsilon}^2(\lambda) d\lambda}{\int_{-\pi}^{\pi} \hat{a}(\lambda) \hat{f}_{\epsilon}^2(\lambda) d\lambda} \\ &\quad + \sqrt{n} \frac{\int_{-\pi}^{\pi} \{\hat{a}(\lambda) - a(\lambda)\} \hat{f}_{\epsilon}(\lambda) \{\hat{f}_{xy}(\lambda) - \beta \hat{f}_{\epsilon}(\lambda)\} d\lambda}{\int_{-\pi}^{\pi} \hat{a}(\lambda) \hat{f}_{\epsilon}^2(\lambda) d\lambda} \\ &\stackrel{\text{say}}{=} (A) + (B).\end{aligned}$$

Since $\hat{a}(\lambda) - a(\lambda) = O_p(\sqrt{M/n})$ and $\hat{f}_{xy}(\lambda) - \beta \hat{f}_{\epsilon}(\lambda) = O_p(\sqrt{M/n})$, we obtain $(B) = o_p(1)$. Also we can see $(A) = \sqrt{n}(\hat{\beta}_a - \beta) + o_p(1)$. The assertion follows from Theorems 3.1 and 3.2. $\square$

REFERENCES

- [1] Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.
- [2] Anderson, T. W. (1984). The 1982 Wald memorial lectures. Estimating linear statistical relationships. *Ann. Statist.*, **12**, 1-45.
- [3] Eltinge, J. L. (1990). Errors in variables estimation in the presence of serially correlated observations. *Contemporary Mathematics*, **112**, 73-95.
- [4] Fuller, W. A. (1987). *Measurement Error Models*. Wiley, New York.
- [5] Fuller, W. A. (1996). *Introduction to Statistical Time Series*. Second Edition. Wiley, New York.
- [6] Hannan, E. J. (1970). *Multiple Time Series*. Wiley, New York.
- [7] Hausman, J. A. and Watson, M. W. (1985). Errors in variables and seasonal adjustment procedures. *J. Amer. Statist. Assoc.*, **80**, 531-540.
- [8] Hosoya, Y. and Taniguchi, M. (1982). A central limit theorem for stationary processes and the parameter estimation of linear processes. *Ann. Statist.*, **10**, 132-153. Correction: *Ann. Statist.*, **21**, 1115-1117.
- [9] Ioannides, D. A. and Alevizos, P. D. (1997). Nonparametric regression with errors in variables and applications. *Stat. and Prob. Letters*, **32**, 35-43.
- [10] Magnus, J. R. and Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley, New York.
- [11] Robinson, P. M. (1977). The estimation of a multivariate linear relation. *J. Multivariate Anal.*, **7**, 409-423.
- [12] Taniguchi, M. (1987). Minimum contrast estimation for spectral densities of stationary processes. *J. Roy. Statist. Ser. B*, **49**, 315-325.
- [13] Taniguchi, M., Puri, M. L. and Kondo, M. (1996). Nonparametric approach for non-Gaussian vector stationary process. *J. Multivariate Anal.*, **56**, 259-283.