JOURNAL OF TIME SERIES ANALYSIS
J. Time. Ser. Anal. **36:** 61–66 (2015)
Published online 19 September 2014 in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1111/jtsa.12092

ORIGINAL ARTICLE

INFERENCE FOR A SPECIAL BILINEAR TIME-SERIES MODEL

SHIQING LING[a] LIANG PENG[b] AND FUKANG ZHU[c,*]

[a] *Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong, China*
[b] *Department of Risk Management and Insurance, Georgia State University, Atlanta, GA, USA*
[c] *School of Mathematics, Jilin University, Changchun 130012, China*

It is well known that estimating bilinear models is quite challenging. Many different ideas have been proposed to solve this problem. However, there is not a simple way to do inference even for its simple cases. This article proposes a generalized autoregressive conditional heteroskedasticity-type maximum likelihood estimator for estimating the unknown parameters for a special bilinear model. It is shown that the proposed estimator is consistent and asymptotically normal under only finite fourth moment of errors.

Received 8 February 2014; Revised 24 August 2014; Accepted 25 August 2014

Keywords: Asymptotic distribution; bilinear model; LSE; MLE.
JEL. Primary C12; C13; C22.

1. INTRODUCTION

The general bilinear time-series model is defined by the equation

$$Y_t = \mu + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \psi_j \varepsilon_{t-j} + \sum_{l=1}^{m} \sum_{l'=0}^{k} b_{ll'} Y_{t-l} \varepsilon_{t-l'} + \varepsilon_t, \tag{1}$$

where $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with mean zero and variance σ^2. It was proposed by Granger and Andersen (1978a) and has been widely applied in many areas such as control theory, economics and finance. The structure of model (1) has been studied in the literature especially for some special cases. For example, Subba Rao (1981) considered model (1) with $\psi_1 = \cdots = \psi_q = 0$; Davis and Resnick (1996) studied the asymptotic behaviour of the correlation function for the simple bilinear model $Y_t = bY_{t-1}\varepsilon_{t-1} + \varepsilon_t$; Pham and Tran (1981), Turkman and Turkman (1997) and Basrak *et al.* (1999) studied the model $Y_t = \phi_1 Y_{t-1} + bY_{t-1}\varepsilon_{t-1} + \varepsilon_t$; Zhang and Tong (2001) considered the model $Y_t = bY_{t-1}\varepsilon_t + \varepsilon_t$. A sufficient condition for stationarity of the general model was obtained by Liu and Brockwell (1988), which is far away from the necessary one as pointed out by Liu (1989). A simplified sufficient condition is given by Liu (1990a).

It is known that estimating the general bilinear model is quite challenging. Many different ideas have been proposed to solve this problem for some special cases of (1); see Pham and Tran (1981), Guegan and Pham (1989), Wittwer (1989), Liu (1990b), Kim and Billard (1990), Kim *et al.* (1990), Sesay and Subba Rao (1992), Gabr (1998) and Hili (2008). However, the asymptotic theory is either rarely established or only derived by assuming that ε_t follows a normal distribution in these papers. The Hellinger distance estimation in Hili (2008) even assumes that the density of ε_t is known. To understand this difficulty, let us look at the least squares estimator (LSE) considered by Pham and Tran (1981). The LSE is equivalent to the quasi-maximum likelihood estimator (quasi-MLE), which

* Correspondence to: Fukang Zhu, School of Mathematics, Jilin University, Changchun 130012, China. E-mail: zfk8010@163.com

is the minimizer of $L_n(\theta) = \sum_{t=1}^n \varepsilon_t^2(\theta)$, where θ is the vector consisting of all parameters in the model, and its true value is $\theta_0, \varepsilon_t(\theta_0) = \varepsilon_t$ and

$$\varepsilon_t(\theta) = Y_t - \mu - \sum_{i=1}^{p} \phi_i Y_{t-i} - \sum_{j=1}^{q} \psi_j \varepsilon_{t-j}(\theta) - \sum_{l=1}^{m} \sum_{l'=0}^{k} b_{ll'} Y_{t-l} \varepsilon_{t-l'}(\theta).$$

Given a sample $\{Y_1, \ldots, Y_n\}$, one needs an efficient way to calculate the residual $\varepsilon_t(\theta)$ such that the effect from the initial values $\{Y_0, Y_{-1}, \ldots\}$ is ignorable. This is the so-called invertibility of the model. Although Liu (1990a) gave a sufficient condition for invertibility, it still remains unknown on how to use it to derive the asymptotic limit of the aforementioned LSE. Another type of invertibility was proposed by Granger and Andersen (1978b). That is, model (1) is said to be invertible if $\lim_{t\to\infty} E(\varepsilon_t - \hat{\varepsilon}_t)^2 = 0$, where $\hat{\varepsilon}_t$ is an estimator of ε_t. Along this direction, the invertibility of a special bilinear model was studied by Subba Rao (1981), Pham and Tran (1981) and Wittwer (1989). This type of invertibility may be useful for forecasting, but it is not useful for proving asymptotic normality of estimators of parameters. This is because we need the property of $\varepsilon_t(\theta)$ at a neighbourhood of the true parameter θ_0 for deriving the asymptotic limit of the estimator. For example, to obtain the asymptotic normality of the LSE, we need the score function $\partial\varepsilon_t(\theta)/\partial\theta$ to have a finite second moment, which in general results in some very restrictive requirements for model (1). Let us further illustrate this issue as follows.

For the following simple bilinear model

$$Y_t = bY_{t-2}\varepsilon_{t-1} + \varepsilon_t, \tag{2}$$

one needs $\prod_{i=1}^m Y_{t-i}$ has a finite moment for any m so as to have $E\{\partial\varepsilon_t(\theta)/\partial\theta\}^2 < \infty$. Grahn (1995) showed that $EY_t^{2m} < \infty$ if and only if $b^{2m}E\varepsilon_t^{2m} < 1$. Note that $E|Y_t|^m < \infty$ for any m is equivalent to $b = 0$ when $\varepsilon_t \sim \mathrm{N}(0, \sigma^2)$. Thus, it is almost impossible to establish the asymptotic normality of the LSE for model (2) unless some special conditions are imposed. Instead, Grahn (1995) proposed a non-standard conditional LSE procedure for model (2) by using the facts that $E(Y_t^2|Y_s, s \le t-2) = \sigma^2 + b^2\sigma^2 Y_{t-2}^2$ and $E(Y_tY_{t-1}|Y_s, s \le t-2) = b\sigma^2 Y_{t-2}$. Although Grahn (1995) derived the asymptotic normality for the conditional LSE, the asymptotic variance and its estimator are not given, so some *ad hoc* method such as bootstrap method is needed to construct confidence intervals for b. Furthermore, the moment condition required is $EY_t^8 < \infty$, which reduces to $b^8\sigma^8 < 1/105$ when $\varepsilon_t \sim \mathrm{N}(0, \sigma^2)$. This is quite restrictive on the parameter space of (b, σ). When $\varepsilon_t \sim \mathrm{N}(0, \sigma^2)$, Giordano (2000) and Giordano and Vitale (2003) obtained the formula of the asymptotic variance for the conditional LSE of b, which can be estimated too. Liu (1990b) considered the LSE for the model $Y_t = \phi Y_{t-p} + bY_{t-p}\varepsilon_{t-q} + \varepsilon_t$ with $p \ge 1$ and obtained its asymptotic normality by assuming that $\partial\varepsilon_t(\theta)/\partial\theta$ has a finite second moment. As in model (2), this condition may only hold when $b = 0$ if $\varepsilon_t \sim \mathrm{N}(0, \sigma^2)$. When $|\varepsilon_t| \le c$ (a constant) holds almost surely and $\phi = 0$, Liu (1990b) showed that this condition holds when $|b| \le 1/(2c)$, which is a small parameter space when c is large. In general, one cannot check whether this condition holds when ε_t is not bounded. That is, a general asymptotic theory for LSE or MLE has not been established for the model in Liu (1990b) up to now.

In this article, we propose a generalized autoregressive conditional heteroskedasticity-type MLE (GMLE) for estimating the unknown parameters. It is shown that the GMLE is consistent and asymptotically normal under only finite fourth moment of errors. We organize this article as follows. Section 2 presents our main results. Section 3 reports some simulation results. For saving space, all proofs, additional simulation results and some remarks are kept in the arXiv version (arXiv:1405.3029).

2. ESTIMATION AND ASYMPTOTIC RESULTS

Throughout, we consider the following special bilinear model:

$$Y_t = \mu + \phi Y_{t-2} + bY_{t-2}\varepsilon_{t-1} + \varepsilon_t, \tag{3}$$

DOI: 10.1111/jtsa.12092

where $\{\varepsilon_t\}$ is a sequence of i.i.d. random variables with mean zero and variance $\sigma^2 > 0$. Then the following statement is given as Theorem 1 in the arXiv version of this article, which also follows from Theorem 1 in Kristensen (2009).

Assume $E \ln|\phi + b\varepsilon_1| < 0$, and then there exists a unique strictly stationary solution to model (3), and the solution is ergodic and has representation $Y_t = \mu + \varepsilon_t + \sum_{i=1}^{\infty} \prod_{r=0}^{i-1} (\phi + b\varepsilon_{t-2r-1})(\mu + \varepsilon_{t-2i})$.

Next, we estimate the unknown parameters. Let $\mathcal{F}_t$ be the σ fields generated by $\{\varepsilon_s : s \le t\}$. Assume that $\{Y_1, Y_2, \ldots, Y_n\}$ are generated by model (3). By noting that $E[Y_t|\mathcal{F}_{t-2}] = \mu + \phi Y_{t-2}, \mathrm{Var}[Y_t|\mathcal{F}_{t-2}] = E[(Y_t - \mu - \phi Y_{t-2})^2 |\mathcal{F}_{t-2}] = \sigma^2 \left(1 + b^2 Y_{t-2}^2\right)$, we propose to estimate parameters by maximizing the following quasi-log-likelihood function:

$$L_n(\theta) = \sum_{t=1}^{n} \ell_t(\theta) \quad \text{and} \quad \ell_t(\theta) = -\frac{1}{2}\left[\ln\left[\sigma^2\left(1 + b^2 Y_{t-2}^2\right)\right] + \frac{(Y_t - \mu - \phi Y_{t-2})^2}{\sigma^2\left(1 + b^2 Y_{t-2}^2\right)}\right],$$

where $\theta = (\mu, \phi, \sigma^2, b^2)^\top$ is the unknown parameter and its true value is denoted by θ_0. The maximizer $\hat{\theta}_n$ of $L_n(\theta)$ is called the GMLE of θ_0. Although the estimation idea has appeared in Francq and Zakoïn (2004), Ling (2004) and Truquet and Yao (2012), the challenge here is that $\{\partial \ell_t(\theta)/\partial\theta\}$ is no longer a martingale difference, which complicates the derivation of the asymptotic limit. A straightforward calculation shows that

$$\frac{\partial \ell_t(\theta)}{\partial \mu} = \frac{Y_t - \mu - \phi Y_{t-2}}{\sigma^2\left(1 + b^2 Y_{t-2}^2\right)}, \quad \frac{\partial \ell_t(\theta)}{\partial \phi} = \frac{Y_{t-2}\left(Y_t - \mu - \phi Y_{t-2}\right)}{\sigma^2\left(1 + b^2 Y_{t-2}^2\right)},$$

$$\frac{\partial \ell_t(\theta)}{\partial \sigma^2} = -\frac{1}{2\sigma^2}\left[1 - \frac{(Y_t - \mu - \phi Y_{t-2})^2}{\sigma^2\left(1 + b^2 Y_{t-2}^2\right)}\right], \quad \frac{\partial \ell_t(\theta)}{\partial b^2} = -\frac{Y_{t-2}^2}{2\left(1 + b^2 Y_{t-2}^2\right)}\left[1 - \frac{(Y_t - \mu - \phi Y_{t-2})^2}{\sigma^2\left(1 + b^2 Y_{t-2}^2\right)}\right].$$

By solving $\sum_{t=1}^{n} \partial \ell_t(\theta)/\partial\mu = \sum_{t=1}^{n} \partial \ell_t(\theta)/\partial\phi = \sum_{t=1}^{n} \partial \ell_t(\theta)/\partial\sigma^2 = 0$, we can write the GMLE for μ, ϕ, σ^2 explicitly in terms of b^2. Hence, using these explicit expressions and the equation $\sum_{t=1}^{n} \partial \ell_t(\theta)/\partial b^2 = 0$, we can first obtain the GMLE for b^2 and then obtain the GMLE for μ, ϕ, σ^2.

It is easy to check that $E[\partial \ell_t(\theta_0)/\partial\theta|\mathcal{F}_{t-2}] = 0$, but $\{\partial \ell_t(\theta_0)/\partial\theta\}_{t=1}^{\infty}$ cannot be a martingale difference. Therefore, we cannot use the central limit theory for martingale difference to derive the asymptotic limit. Instead, we will show that $\{\partial \ell_t(\theta_0)/\partial\theta\}_{t=1}^{\infty}$ is a near-epoch dependent sequence so that the asymptotic limit of the proposed GMLE can be derived. Denote

$$\Omega = E\left[\frac{\partial \ell_t(\theta_0)}{\partial\theta} + \frac{\partial \ell_{t-1}(\theta_0)}{\partial\theta}\right]\left[\frac{\partial \ell_t(\theta_0)}{\partial\theta} + \frac{\partial \ell_{t-1}(\theta_0)}{\partial\theta}\right]^\top - E\left[\frac{\partial \ell_t(\theta_0)}{\partial\theta}\frac{\partial \ell_t(\theta_0)}{\partial\theta^\top}\right],$$

$$\Sigma = \mathrm{diag}\left\{E\left[\frac{1}{\sigma_0^2\left(1 + b_0^2 Y_{t-2}^2\right)}\begin{pmatrix} 1 & Y_{t-2} \\ Y_{t-2} & Y_{t-2}^2 \end{pmatrix}\right], \; E\begin{pmatrix} \frac{1}{2\sigma_0^4} & \frac{Y_{t-2}^2}{2\sigma_0^2(1+b_0^2 Y_{t-2}^2)} \\ \frac{Y_{t-2}^2}{2\sigma_0^2(1+b_0^2 Y_{t-2}^2)} & \frac{Y_{t-2}^4}{2\left(1+b_0^2 Y_{t-2}^2\right)^2} \end{pmatrix}\right\}.$$

The following theorem gives the asymptotic properties of the GMLE, which is Theorem 2 in the arXiv version.

Theorem 1. Suppose the parameter space Θ is a compact subset of $\{\theta : E \ln|\phi + b\varepsilon_1| < 0, |\mu| \le \bar{\mu}, |\phi| \le \bar{\phi}, \underline{\omega} \le \sigma^2 \le \overline{\omega}, \underline{\alpha} \le b^2 \le \overline{\alpha}\}$, where $\overline{\mu}, \overline{\phi}, \underline{\omega}, \overline{\omega}, \underline{\alpha}$ and $\overline{\alpha}$ are some finite positive constants, and the true parameter value θ_0 is an interior point in Θ. Further assume $E\varepsilon_1^4 < \infty$. Then as $n \to \infty$, we have (a) $\hat{\theta}_n \to \theta_0$ almost surely; (b) $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} \mathrm{N}(0, \Sigma^{-1}\Omega\Sigma^{-1})$.

DOI: 10.1111/jtsa.12092

Since $\sum_{t=1}^{n} \partial \ell_t(\theta)/\partial b^2 = 0$ is equivalent to $\sum_{t=1}^{n} \partial \ell_t(\theta)/\partial b = 0$, one cannot estimate b by the above GMLE. So as to estimate b, we need a consistent estimator for the sign of b. Write $(Y_t - \mu - \phi Y_{t-2})(Y_{t-1} - \mu - \phi Y_{t-3}) = \varepsilon_t(bY_{t-3}\varepsilon_{t-2} + \varepsilon_{t-1}) + b^2 Y_{t-3}Y_{t-2}\varepsilon_{t-2}\varepsilon_{t-1} + bY_{t-2}\varepsilon_{t-1}^2$. It is easy to see that $E\{(Y_t - \mu - \phi Y_{t-2})(Y_{t-1} - \mu - \phi Y_{t-3})|\mathcal{F}_{t-2}\} = b\sigma^2 Y_{t-2}$, which motivates to estimate b by minimizing the following least squares $\sum_{t=2}^{n}\left\{(Y_t - \mu - \phi Y_{t-2})(Y_{t-1} - \mu - \phi Y_{t-3}) - b\sigma^2 Y_{t-2}\right\}^2$ with μ, ϕ and σ^2 being replaced by the corresponding GMLE. However, so as to avoid requiring some moment conditions on Y_t, we propose to minimize the weighted least squares $\sum_{t=2}^{n}\left\{(Y_t - \mu - \phi Y_{t-2})(Y_{t-1} - \mu - \phi Y_{t-3}) - b\sigma^2 Y_{t-2}\right\}^2 / \{(1 + Y_{t-2}^2)\sqrt{1 + Y_{t-3}^2}\}$ with μ, ϕ, σ^2 being replaced by the corresponding GMLE. This results in

$$\tilde{b}_n = \left(\hat{\theta}_{n3}\sum_{t=2}^{n}\frac{Y_{t-2}^2}{\left(1+Y_{t-2}^2\right)\sqrt{1+Y_{t-3}^2}}\right)^{-1}\sum_{t=2}^{n}\frac{(Y_t - \hat{\theta}_{n1} - \hat{\theta}_{n2}Y_{t-2})(Y_{t-1} - \hat{\theta}_{n1} - \hat{\theta}_{n2}Y_{t-3})Y_{t-2}}{\left(1+Y_{t-2}^2\right)\sqrt{1+Y_{t-3}^2}}.$$

Like Theorem 1(a), it is easy to show that $\tilde{b}_n = b + o_p(1)$. Using $\tilde{b}_n$ to estimate the sign of b, we obtain an estimator for b as $\hat{b}_n = \text{sgn}(\tilde{b}_n)\sqrt{\hat{\theta}_{n4}}$. It easily follows from Theorem 1 that $\hat{b}_n = b + o_p(1)$ and the asymptotic limit of $2b\sqrt{n}(\hat{b}_n - b)$ is the same as that of $\sqrt{n}(\hat{\theta}_{n4} - b^2)$ given in Theorem 1. As stated in the simulation study, we propose to use $2b\sqrt{n}(\hat{b}_n - b)$ rather than $2\hat{b}_n\sqrt{n}(\hat{b}_n - b)$ to construct a confidence interval for b, although both share the same asymptotic limit. Moreover, we do not propose to estimate b directly by $\tilde{b}_n$. The reason is that like Grahn (1995), we cannot derive the formula and a consistent estimator for the asymptotic variance of $\sqrt{n}(\tilde{b}_n - b)$. Moreover, $\tilde{b}_n$ is a less efficient estimator than $\hat{b}_n$ in general.

Theorem 1 excludes the case of $b = 0$, which reduces the bilinear model to a linear model. Hence, testing $H_0 : b = 0$ is of interest. Write $\Theta = [-\bar{\mu}, \bar{\mu}] \times [-\bar{\phi}, \bar{\phi}] \times [\underline{\omega}, \overline{\omega}] \times [0, \overline{\alpha}]$, where $\overline{\mu}, \overline{\phi}, \underline{\omega}, \overline{\omega}$ and $\overline{\alpha}$ are some finite positive constants. Then the case of $b = 0$ means that $\theta = (\mu, \phi, \sigma^2, b^2)^\top$ lies at the boundary of the compact set Θ, which implies that the case of $b = 0$ is the well-known non-standard situation of maximum likelihood estimation. The following theorem is Theorem 3 in the arXiv version.

Theorem 2. Suppose the parameter space Θ satisfies $E\ln|\phi + b\varepsilon_1| < 0$, and the true parameter value $\theta_0 = (\mu_0, \phi_0, \sigma_0^2, 0)^\top$ satisfies that $(\mu_0, \phi_0, \sigma_0^2)^\top$ is an interior point of $[-\bar{\mu}, \bar{\mu}] \times [-\bar{\phi}, \bar{\phi}] \times [\underline{\omega}, \overline{\omega}]$. Further assume $E\varepsilon_1^4 < \infty$. Then as $n \to \infty$, we have (a) $\hat{\theta}_n \to \theta_0$ almost surely; (b) $\sqrt{n}(\hat{\theta}_n - \theta_0) \xrightarrow{d} (Z_1, Z_2, Z_3, Z_4)^\top I(Z_4 > 0) + (Z_1 - \sigma_{14}\sigma_{44}^{-1}Z_4, Z_2 - \sigma_{24}\sigma_{44}^{-1}Z_4, Z_3 - \sigma_{34}\sigma_{44}^{-1}Z_4, 0)^\top I(Z_4 < 0)$, where $(Z_1, Z_2, Z_3, Z_4)^\top \sim \text{N}(0, \Sigma^{-1}\Omega\Sigma^{-1})$, $\Sigma^{-1}\Omega\Sigma^{-1} = (\sigma_{ij})$ and Σ and Ω are given in Theorem 1.

3. SIMULATION

We investigate the finite sample performance of the GMLE by drawing 1000 random samples of size $n = 200$ and 1000 from model (3) with $\mu = 0, b = \pm 0.1$ or $\pm 1, \phi = 0$ or 0.9, and $\varepsilon_t \sim \text{N}(0, 1)$. We compute the GMLE $\hat{\theta}_n = (\hat{\theta}_{n1}, \ldots, \hat{\theta}_{n4})^\top$ for $\theta = (\mu, \phi, \sigma^2, b^2)^\top$ and $\hat{b}_n$. For an estimator $\hat{\beta}$, we use $E(\hat{\beta})$, $\text{SD}(\hat{\beta})$ and $\widehat{\text{SD}}(\hat{\beta})$ to denote the sample mean of $\hat{\beta}$, sample standard deviation of $\hat{\beta}$ and sample mean of the standard deviation estimator of $\hat{\beta}$ based on the 1000 samples.

Table I reports these quantities for the case of $n = 200$, which show that the proposed GMLE has a small bias (i.e. $E(\cdot)$ close to the true value) and the proposed variance estimator is accurate too (i.e. $\widehat{\text{SD}}(\cdot)$ close to $\text{SD}(\cdot)$). Results are $n = 1000$ can be found in the arXiv version, which show that $\text{SD}(\hat{\beta})$ and $\widehat{\text{SD}}(\hat{\beta})$ are much smaller when $n = 1000$ than those when $n = 200$. Although the proposed estimator for b has a small bias, the proposed variance estimator performs badly when b is small. This is due to some very small values of $\hat{\theta}_{n4}$. However, the variance estimator for $2b\hat{b}_n$ is reasonably well and much accurate than that for $\hat{b}_n$. Hence, we suggest to use $2b\sqrt{n}(\hat{b}_n - b)$ instead of $\sqrt{n}(\hat{b}_n - b)$ to construct a confidence interval for b in practice.

DOI: 10.1111/jtsa.12092

Table I. Sample mean and sample standard deviation are reported for the proposed GMLE for $(\mu, \phi, \sigma^2, b^2)^\top$ and b with $n = 200$

(b, ϕ)	(0.1, 0)	(0.1, 0.9)	(−0.1, 0)	(−0.1, 0.9)	(1, 0)	(1, 0.9)	(−1, 0)	(−1, 0.9)
$E(\hat{\theta}_{n1})$	0.0008	−0.0020	0.0002	0.0021	−0.0014	0.0054	0.0039	−0.0009
$\mathrm{SD}(\hat{\theta}_{n1})$	0.0707	0.0890	0.0720	0.0882	0.1058	0.1315	0.1015	0.1369
$\widehat{\mathrm{SD}}(\hat{\theta}_{n1})$	0.0701	0.0740	0.0703	0.0740	0.0999	0.1244	0.0997	0.1241
$E(\hat{\theta}_{n2})$	−0.0026	0.8793	−0.0092	0.8797	−0.0113	0.8887	−0.0052	0.8892
$\mathrm{SD}(\hat{\theta}_{n2})$	0.0714	0.0399	0.0706	0.0400	0.1000	0.0878	0.0979	0.0891
$\widehat{\mathrm{SD}}(\hat{\theta}_{n2})$	0.0698	0.0352	0.0693	0.0350	0.0977	0.0868	0.0974	0.0869
$E(\hat{\theta}_{n3})$	0.9648	0.9896	0.9744	0.9919	1.0209	1.0254	1.0263	1.0280
$\mathrm{SD}(\hat{\theta}_{n3})$	0.1115	0.1161	0.1047	0.1179	0.1976	0.2535	0.1993	0.2524
$\widehat{\mathrm{SD}}(\hat{\theta}_{n3})$	0.1196	0.1253	0.1209	0.1259	0.1855	0.2245	0.1880	0.2280
$E(\hat{\theta}_{n4})$	0.0355	0.0124	0.0344	0.0117	0.9987	1.0256	0.9882	1.0201
$\mathrm{SD}(\hat{\theta}_{n4})$	0.0615	0.0166	0.0570	0.0168	0.3183	0.3205	0.3246	0.3041
$\widehat{\mathrm{SD}}(\hat{\theta}_{n4})$	0.0738	0.0174	0.0728	0.0172	0.2786	0.2667	0.2797	0.2671
$E(\hat{b}_n)$	0.0879	0.0780	−0.0986	−0.0736	0.9872	1.0015	−0.9812	−0.9996
$\mathrm{SD}(\hat{b}_n)$	0.1666	0.0793	0.1572	0.0793	0.1555	0.1505	0.1593	0.1446
$\widehat{\mathrm{SD}}(\hat{b}_n)$	4.7444	0.8028	4.5939	0.8511	0.1379	0.1294	0.1393	0.1303
$E(2b\hat{b}_n)$	0.0176	0.0156	0.0197	0.0147	1.9744	2.0030	1.9624	1.9992
$\mathrm{SD}(2b\hat{b}_n)$	0.0333	0.0159	0.0314	0.0159	0.3111	0.3010	0.3187	0.2892
$\widehat{\mathrm{SD}}(2b\hat{b}_n)$	0.0738	0.0174	0.0728	0.0172	0.2786	0.2667	0.2797	0.2671

Simulation results for testing $H_0 : b = 0$ against $H_a : b \neq 0$ can be found in the arXiv version, which shows that the proposed test has a reasonably accurate size and non-trivial power.

ACKNOWLEDGEMENTS

We thank the editor, an associate editor and two reviewers for suggestions on writing this article as a short note. Research was supported by Hong Kong RGC grants HKUST641912, 603413 and FSGRF12SC12; NSF grant DMS-1005336; Simons Foundation; NSF of China grants 11371168 and 11271155; SRFDP grant 20110061110003; Science and Technology Developing Plan of Jilin Province under grant 20130522102JH; and SRF for ROCS, SEM.

REFERENCES

Basrak B, Davis RA, Mikosch T. 1999. The sample ACF of a simple bilinear process. *Stochastic Processes and their Applications* **83**: 1–14.

Davis RA, Resnick SI. 1996. Limit theory for bilinear processes with heavy-tailed noise. *Annals of Applied Probability* **6**: 1191–1210.

Francq C, Zakoïn JM. 2004. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli* **10**: 605–637.

Gabr MM. 1998. Robust estimation of bilinear time series models. *Communications in Statistics-Theory and Methods* **27**: 41–53.

Giordano F. 2000. The variance of CLS estimators for a simple bilinear model. *Quaderni di Statistica* **2**: 147–155.

DOI: 10.1111/jtsa.12092

Giordano F, Vitale C. 2003. CLS asymptotic variance for a particular relevant bilinear time series model. *Statistical Methods and Applications* **12**: 169–185.
Grahn T. 1995. A conditional least squares approach to bilinear time series estimation. *Journal of Time Series Analysis* **16**: 509–529.
Granger CWJ, Andersen AP. 1978a. *An Introduction to Bilinear Time Series Models*. Vandenhoeck & Ruprecht: Göttingen.
Granger CWJ, Andersen AP. 1978b. On the invertibility of time series models. *Stochastic Processes and their Applications* **8**: 87–92.
Guegan D, Pham DT. 1989. A note on the estimation of the parameters of the diagonal bilinear model by the method of least squares. *Scandinavian Journal of Statistics* **16**: 129–136.
Hili O. 2008. Hellinger distance estimation of general bilinear time series models. *Statistical Methodology* **5**: 119–128.
Kim WY, Billard L. 1990. Asymptotic properties for the first-order bilinear time series model. *Communications in Statistics-Theory and Methods* **19**: 1171–1183.
Kim WY, Billard L, Basawa IV. 1990. Estimation for the first-order diagonal bilinear time series model. *Journal of Time Series Analysis* **11**: 215–229.
Kristensen D. 2009. On stationarity and ergodicity of the bilinear model with applications to GARCH models. *Journal of Time Series Analysis* **30**: 125–144.
Ling S. 2004. Estimation and testing stationarity for double-autoregressive models. *Journal of the Royal Statistical Society Series B* **66**: 63–78.
Liu J. 1989. A simple condition for the existence of some stationary bilinear time series. *Journal of Time Series Analysis* **10**: 33–39.
Liu J. 1990a. A note on causality and invertibility of a general bilinear time series model. *Advances in Applied Probability* **22**: 247–250.
Liu J. 1990b. Estimation for some bilinear time series. *Stochastic Models* **6**: 649–665.
Liu J, Brockwell PJ. 1988. On the general bilinear time series model. *Journal of Applied Probability* **25**: 553–564.
Pham DT, Tran LT. 1981. On the first-order bilinear time series model. *Journal of Applied Probability* **18**: 617–627.
Sesay SAO, Subba Rao T. 1992. Frequency-domain estimation of bilinear time series models. *Journal of Time Series Analysis* **13**: 521–545.
Subba Rao T. 1981. On the theory of bilinear time series models. *Journal of the Royal Statistical Society Series B* **43**: 244–255.
Truquet L, Yao J. 2012. On the quasi-likelihood estimation for random coefficient autoregressions. *Statistics* **46**: 505–521.
Turkman KF, Turkman MAA. 1997. Extremes of bilinear time series models. *Journal of Time Series Analysis* **18**: 305–319.
Wittwer G. 1989. Some remarks on bilinear time series models. *Statistics* **20**: 521–529.
Zhang Z, Tong H. 2001. On some distributional properties of a first-order nonnegative bilinear time series model. *Journal of Applied Probability* **38**: 659–671.

DOI: 10.1111/jtsa.12092