J. Indones. Math. Soc.
Vol. 19, No. 1 (2013), pp. 99–110.

SECOND ORDER LEAST SQUARE ESTIMATION ON ARCH(1) MODEL WITH BOX-COX TRANSFORMED DEPENDENT VARIABLE

Herni Utami[1] and Subanar[2]

[1]**Department of Mathematics Gadjah Mada University, Indonesia**
herni.utami@gmail.com
[2]**Department of Mathematics Gadjah Mada University, Indonesia**
subanar@ugm.ac.id

Abstract. Box-Cox transformation is often used to reduce heterogeneity and to achieve a symmetric distribution of response variable. In this paper, we estimate the parameters of Box-Cox transformed ARCH(1) model using second-order least square method and then we study the consistency and asymptotic normality for second-order least square (SLS) estimators. The SLS estimation was introduced by Wang (2003, 2004) to estimate the parameters of nonlinear regression models with independent and identically distributed errors.

Key words: Box-Cox transformation, second-order least square, ARCH model.

Abstrak. Transformasi Box-Cox sering digunakan untuk mengurangi heterogenitas dan mencapai distribusi simetris dari variabel respon. Pada paper ini dibahas estimasi parameter dari model ARCH(1) di mana variabel responnya ditransformasi Box-Cox dengan menggunakan metode estimasi second-order least square dan selanjutnya diteliti konsistensi dan normalitas asimtotik dari estimator second-order least square. Metode ini pertama kali diperkenalkan oleh Wang (2003, 2004) untuk mengestimasi parameter model regresi nonlinier yang variabel errornya berdistribusi identik dan independen.

Kata kunci: Transformasi Box-Cox, second-order least square, model ARCH

1. INTRODUCTION

Time series related to finance usually have three typical characteristics (Chan (2002)):

2000 Mathematics Subject Classification: 37M10.
Received: 10-08-2012, revised: 07-06-2013, accepted: 20-06-2013.

(1) the unconditional distribution of financial time series such as stock price returns, has heavier tails than the normal distribution,
(2) the value of time series $\{X_t\}$ is not correlated with each other, but $\{X_t^2\}$ is strongly correlated with each other,
(3) the volatility clustering.

One of the models that can be used to model the above conditions is Autoregressive Conditional Heteroscedastic (ARCH) model proposed by Engle (1982).

Two popular estimation methods for ARCH model are maximum likelihood and least square methods. Weiss (1986) discussed properties of maximum likelihood estimation and least square estimation of the parameters of both regression and ARCH equation. Basawa (1976) studied consistency and asymptotic normality for maximum likelihood estimators in the case where the observed random variables may be dependent and not identically distributed. The least square estimation procedure for ARCH model is constructed in two stages. The first is to estimate the regression equation of the mean and the second is to estimate the regression equation of variance. Therefore, using least square method for estimating the ARCH model will not obtain estimator for the mean and the variance regression simultaneously. Wang and Leblanc (2008) estimated the parameters of nonlinear regression models with independent and identically distributed errors. We will propose second order least square (SLS) method to estimate parameters of ARCH model. The method does not require assumptions on the specific distribution of the errors and the estimators for mean and variance regression will be obtained simultaneously.

Box-Cox transformation can be used to reduce heterogeneity and achieve a symmetrical distribution of the response variable. Draper and Cox (1969) and Poiriers (1978) have shown that linearity, homoscedasticity, and normality cannot be done simultaneously with a certain Box-Cox transformation. Sarkar (2000) defined Box-Cox transformed ARCH model (BCARCH) and he considered maximum likelihood method to estimate parameters of BCARCH. Testing and estimation of BCARCH model are investigated and a Lagrange multiplier test is also developed to test Engle's linear ARCH model against this wider class of models. In this paper, we propose second-order least square method to estimate parameters of BCARCH model.

The paper is organized as follows. In section 2, we describe Box-Cox transformed ARCH model. Estimation method is discussed in section 3. We developed method for testing power Box-Cox transformation in section 4. Finally, in section 5, Monte Carlo simulations of finite sample performance of the estimator is provided.

2. BOX-COX TRANSFORMED ARCH MODEL

The family of ARCH model, which was introduced by Engle (1982) have proven useful in financial applications and have attracted great attention in economics and statistical literature (Alberola (2006), Gao, Yu, and Chen (2009), Hardle and Hafner (2000), Lamoureux, *et al.*(1990)). Let $(\mathbf{X}_t, Y_t)$ denote vector of predictor variables and response variable at the time t respectively. The ARCH(R) models proposed in this paper is defined by

$$Y_t|\Im_t \sim N(\mathbf{X}_t'\beta, h_t), \tag{1}$$

where $\Im_t$ is the information set containing information about the process up to and including time $t-1$ and

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \cdots + \alpha_R\varepsilon_{t-R}^2. \tag{2}$$

The error term, ε_t, has mean zero and variance h_t which is split into a stochastic piece u_t and time-dependent variation h_t characterizing the typical size of the term so that $\varepsilon_t = u_t\sqrt{h_t}$. Coefficients $\alpha_0 \geq 0, \alpha_i > 0$, so that conditional variance is strictly positive, $\mathbf{X}_t$ is a $k \times 1$ vector of fixed observation at the time t on p independent variables which may include this lagged value of the dependent variable, $\beta' = (\beta_1, \beta_2, ..., \beta_p)$ is a vector of associated regression coefficients.

Sarkar (2000) stated that the Box-Cox transformed ARCH(1) model is generalization of the ARCH model and can be represented by

$$Y_t^{(\lambda)} = \mathbf{X}_t'\beta + \varepsilon_t, \tag{3}$$

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2, \tag{4}$$

with

$$\varepsilon_t = u_t\sqrt{h_t} \tag{5}$$

where $0 \leq \alpha_0, 0 < \alpha_1 < 1$, (u_t) is a sequence of iid random variables with $E(u_t) = 0$ and $E(u_t^2) = 1$. The Box-Cox transformed value of the (original) dependent variable y_t i.e.

$$y_t^{(\lambda)} = \begin{cases} \left(Y_t^{\lambda} - 1\right)/\lambda & ,\lambda \neq 0 \\ \log Y_t & ,\lambda = 0 \end{cases} \tag{6}$$

The transformation in equation (6) is valid only for $y_t > 0$ and, therefore, modifications have to be made for negative observation. Box and Cox proposed the shifted power transformation with the form

$$y_t^{(\lambda)} = \begin{cases} \frac{(y_t+c)^{\lambda}-1}{\lambda}, & \lambda \neq 0 \\ \log(y_t + c), & \lambda = 0 \end{cases} \tag{7}$$

where λ is the power transformation and c is chosen such that $y_t + c > 0$ for $t = 1, 2, ..., T$. The λ is a parameter in this model, and the parameter indicates degree of nonlinearity in the data. The model reduces to the linear model when $\lambda = 1$. Hence, we develop test for the linear model by hypothesis $H_0 : \lambda = 1$ vs $H_1 : \lambda \neq 1$.

The ARCH(1) model assume that $E\left[\log(\alpha_1\varepsilon_t^2)\right] < 0$. The assumption is known to be necessary for stationarity, see Nelson (1990) for coefficient of conditional variance of ε_t on the GARCH (1,1) model is zero.

Conditional mean of ε_t is given by

$$E(\varepsilon_t|\Im_{t-1}) = 0, \tag{8}$$

and conditional variance of ε_t is

$$\begin{aligned} h_t &= E(\varepsilon_t^2|\Im_{t-1}) \\ &= \alpha_0 + \alpha_1\varepsilon_{t-1}^2. \end{aligned}$$

where $\alpha' = (\alpha_0, \alpha_1)$ is a vector of parameters in the ARCH or variance equation. The complete parameter vector for the model is $\theta' = (\lambda, \beta', \alpha')$. The parameter space as $\Theta \subset R^{p+3}$ is compact set that has at least one interior point.

3. ESTIMATION

In this section, we briefly outline the estimation procedure for model (3) and (4) with the second-order least square estimation method proposed by Wang and Leblanc (2008), Abarin and Wang (2006). If $\hat{\theta}_{\mathbf{SLS}}$ is second-order least square estimator for θ, then it is determined by minimizing the squared distance of the response variable to its first conditional moment and the square response variable to its second conditional moment of response variable:

$$Q_T(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T}\rho_t'(\boldsymbol{\theta})\mathbf{W}_t\rho_t(\boldsymbol{\theta}) \tag{9}$$

where $\rho_t(\boldsymbol{\theta}) = (Y_t^{(\lambda)} - E(Y_t^{\lambda}|\Im_t), (Y_t^{(\lambda)})^2 - E((Y_t^{(\lambda)})^2|\Im_t))'$ $\mathbf{W}_t = \mathbf{W}(\mathbf{X}_t)$ is weight that is a 2x2 nonnegative definite matrix which may depend on $\mathbf{X}_i$.

The SLSE for θ can be represented

$$\hat{\boldsymbol{\theta}}_{SLS} = \arg\min_{\boldsymbol{\theta}} Q_T(\boldsymbol{\theta}), \tag{10}$$

where $\boldsymbol{\theta} \in \Theta$. In order to find $\boldsymbol{\theta}$ which minimizes $Q_T(\boldsymbol{\theta})$ in equation (9), we recommend using the algorithm proposed by Berndt et al (1974).

Lemma 3.1. *Let ε_t be a ARCH(1) process,*

$$\varepsilon_t = \sqrt{h_t}z_t,\ z_t \sim IID(0, \sigma^2), \tag{11}$$

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2,\ 0 \text{ ¡ } \alpha_1 < 1. \tag{12}$$

Then $\{\varepsilon_t^2\}$ is an ergodic process.

PROOF. Sequence (z_t) is iid, so (z_t) is stationary and ergodic. Repeatedly substituting for ε_{t-1}^2 in equation (12), we have, for $t \geq 1$,

$$h_t = \alpha_0\left(\sum_{j=0}^{\infty}\alpha_1^j\prod_{i=0}^{j}z_{t-i}^2\right). \tag{13}$$

Suppose

$$g(z_0, z_1, z_2, ...) = \alpha_0\left(\sum_{j=0}^{\infty}\alpha_1^j\prod_{i=0}^{j}z_{t-i}^2\right). \tag{14}$$

Let a sequence space $S = \{\mathbf{z} = (z_k) : z_k \in R,\ k = 0, 1, 2, ...\}$. For $\mathbf{z}, \mathbf{y} \in S$ and $j \leq t$, we define $a_j = \prod_{i=0}^{j} z_{t-i}^2$ and $b_j = \prod_{i=0}^{j} y_{t-i}^2$ and a function $\rho : S \times S \to R$ such that for any $\mathbf{z}, \mathbf{y} \in S$,

$$\begin{aligned}\rho(\mathbf{z}, \mathbf{y}) &= \max_j \{|a_j - b_j|\} \\ &= \max_j \left\{ \left| \prod_{i=0}^{j} z_{t-i}^2 - \prod_{i=0}^{j} y_{t-i}^2 \right| \right\}.\end{aligned}$$

It is easy to show that ρ is a metric on S. For $\rho(\mathbf{z}, \mathbf{y}) = \|\mathbf{z} - \mathbf{y}\|$ then ρ is a norm. Given $\varepsilon > 0$, there exists a $\delta = \frac{1-\alpha_1}{\alpha_0}\varepsilon > 0$ such that for all $\mathbf{z}, \mathbf{y} \in S$ with $\|\mathbf{z} - \mathbf{y}\| = \max_j \left\{ \left| \prod_{i=0}^{j} z_{t-i}^2 - \prod_{i=0}^{j} y_{t-i}^2 \right| \right\} < \delta$, then

$$\begin{aligned}|g(\mathbf{z}) - g(\mathbf{y})| &= \left| \alpha_0 \sum_{j=0}^{\infty} \alpha_1^j \prod_{i=0}^{j} z_{t-i}^2 - \alpha_0 \sum_{j=0}^{\infty} \alpha_1^j \prod_{i=0}^{j} y_{t-i}^2 \right| \\ &\leq \alpha_0 \sum_{j=0}^{\infty} \alpha_1^j \left| \prod_{i=0}^{j} z_{t-i}^2 - \prod_{i=0}^{j} y_{t-i}^2 \right| \\ &< \alpha_0 \sum_{j=0}^{\infty} \alpha_1^j \frac{1-\alpha_1}{\alpha_0} \varepsilon \\ &= (1-\alpha_1)\varepsilon \sum_{j=0}^{\infty} \alpha_1^j \\ &= \varepsilon\end{aligned}$$

Therefore, function g is continuous. By using ergodic theory, $\{\varepsilon_t^2\}$ is an ergodic process. ■

Theorem 3.2. (Meyn and Tweedie (1993)) *Function $f_n : R^d \to R, n \in N$ are continuous and they have partial derivative due to each variable. If there exists constant M such that $\|f_n\| \leqslant M$ for $n \in N$ and $\mathbf{x} \in R^d$ then the family $\{f_n, n \in N\}$ is equicontinuous.*

Assumption 1 Parameter space $\Theta \subset \mathbb{R}^{p+R+2}$ is compact.
Assumption 2 $(\mathbf{W}_t)_{t \in \mathbb{N}} \xrightarrow{a.s} \mathbf{W}_0$.
Assumption 3 $E(\varepsilon_t^4) < \infty$.

Theorem 3.3. *Under assumption 1-3, the estimator SLS $\hat{\boldsymbol{\theta}}_{SLS} \xrightarrow{a.s} \boldsymbol{\theta}_0$ as $T \to \infty$.*

Proof. By using ergodic theory, $\{Q_T(\boldsymbol{\theta})\}$ is a ergodic process and we have

$$Q_T(\boldsymbol{\theta}) \xrightarrow{a.s} E\left(\rho_t'(\boldsymbol{\theta})\mathbf{W}_0\rho_t(\boldsymbol{\theta})\right) = Q(\boldsymbol{\theta}).$$

The expected value of $\rho_t'(\boldsymbol{\theta})W_0\rho_t(\boldsymbol{\theta})$ can be described by

$$\begin{aligned} Q(\boldsymbol{\theta}) &= E\left(\rho_t'(\boldsymbol{\theta})\right)\mathbf{W}_0\rho_t(\boldsymbol{\theta}) \\ &= E\left\{\left[\rho_t(\boldsymbol{\theta}_0)+(\rho_t(\boldsymbol{\theta})-\rho_t(\boldsymbol{\theta}_0))\right]'\mathbf{W}_0\left[\rho_t(\boldsymbol{\theta}_0)+(\rho_t(\boldsymbol{\theta})-\rho_t(\boldsymbol{\theta}_0))\right]\right\} \\ &= E\left\{\rho_t'(\boldsymbol{\theta}_0)\mathbf{W}_0\rho_t(\boldsymbol{\theta}_0)+2E\left\{\rho_t'(\boldsymbol{\theta}_0)\mathbf{W}_0(\rho_t(\boldsymbol{\theta})-\rho_t(\boldsymbol{\theta}_0))\right\}\right\} \\ &\quad + E\left\{(\rho_t(\boldsymbol{\theta})-\rho_t(\boldsymbol{\theta}))'\mathbf{W}_0(\rho_t(\boldsymbol{\theta})-\rho_t(\boldsymbol{\theta}_0))\right\}. \end{aligned}$$

Since $\rho_t'(\boldsymbol{\theta}_0)\mathbf{W}_0(\rho_t(\boldsymbol{\theta})-\rho_t(\boldsymbol{\theta}_0))$ does not depend on Y_t, it is a function of $\Im_{t-1}$, and $E\left\{(\rho_t(\boldsymbol{\theta})-\rho_t(\boldsymbol{\theta}))'\mathbf{W}_0(\rho_t(\boldsymbol{\theta})-\rho_t(\boldsymbol{\theta}_0))\right\} \geqslant 0$ we have

$$\begin{aligned} Q(\boldsymbol{\theta}) &= Q(\boldsymbol{\theta}_0)+E\left((\rho_t(\boldsymbol{\theta})-\rho_t(\boldsymbol{\theta}_0))'\mathbf{W}_0(\rho_t(\boldsymbol{\theta})-\rho_t(\boldsymbol{\theta}_0))\right) \\ &\geq Q(\boldsymbol{\theta}_0). \end{aligned}$$

It is clear that $Q(\boldsymbol{\theta}) = Q(\boldsymbol{\theta}_0)$ if only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. It means that $Q(\boldsymbol{\theta})$ has a unique minimum.

Note that $\hat{\boldsymbol{\theta}}_{SLS} = \arg\min_{\boldsymbol{\theta}\in\Theta} Q_T(\boldsymbol{\theta})$ and $\boldsymbol{\theta}_0 = \arg\min_{\boldsymbol{\theta}\in\Theta} Q(\boldsymbol{\theta})$ which imply

$$Q_T(\hat{\boldsymbol{\theta}}_{SLS}) \leqslant Q_T(\boldsymbol{\theta}), \text{ for every } \boldsymbol{\theta}\in\Theta \tag{15}$$

and

$$Q(\boldsymbol{\theta}_0) \leqslant Q_T(\boldsymbol{\theta}), \text{ for every } \boldsymbol{\theta}\in\Theta. \tag{16}$$

By using inequality (15) and (16) we observe that

$$Q_T(\hat{\boldsymbol{\theta}}_{SLS}) - Q(\hat{\boldsymbol{\theta}}_{SLS}) \leqslant Q_T(\hat{\boldsymbol{\theta}}_{SLS}) - Q(\boldsymbol{\theta}_0) \leqslant Q_T(\boldsymbol{\theta}_0) - Q(\boldsymbol{\theta}_0). \tag{17}$$

Therefore from the above we have

$$\begin{aligned} \left|Q_T(\hat{\boldsymbol{\theta}}_{SLS}) - Q(\boldsymbol{\theta}_0)\right| &\leqslant \max\left\{|Q_T(\boldsymbol{\theta}_0)-Q(\boldsymbol{\theta}_0)|, |Q_T(\hat{\boldsymbol{\theta}}_{SLS})-Q(\hat{\boldsymbol{\theta}}_{SLS})|\right\} \\ &\leq \sup_{\boldsymbol{\theta}\in\Theta}|Q_T(\boldsymbol{\theta})-Q(\boldsymbol{\theta})|. \end{aligned} \tag{18}$$

$Q_T(\boldsymbol{\theta})$ is a continuous function on a compact set Θ and differentiable for every element i^{th} of $\boldsymbol{\theta}$, θ_i for $i = 1, 2, ..., p+3$.

The derivative of $Q_T(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ is denoted by

$$\begin{aligned} \nabla Q_T(\boldsymbol{\theta}) &= \frac{1}{T}\sum \frac{\partial}{\partial\boldsymbol{\theta}}\left(\rho_t'(\boldsymbol{\theta})\mathbf{W}_t\rho_t(\boldsymbol{\theta})\right) \\ &= \frac{2}{T}\sum \frac{\partial\rho_t'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{W}_t\rho_t(\boldsymbol{\theta}), \end{aligned}$$

where

$$\begin{aligned} \frac{\partial\rho_t'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} &= \begin{pmatrix} \frac{\partial}{\partial\boldsymbol{\theta}}(Y_t^{(\lambda)} - \mathbf{X}_t'\boldsymbol{\beta}) & \frac{\partial}{\partial\boldsymbol{\theta}}((Y_t^{(\lambda)})^2 - (\mathbf{X}_t'\boldsymbol{\beta})^2 - h_t) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial Y_t^{(\lambda)}}{\partial\boldsymbol{\theta}} & 2\frac{\partial Y_t^{(\lambda)}}{\partial\boldsymbol{\theta}} \\ -\mathbf{X}_t & -2\mathbf{X}_t'\boldsymbol{\beta}\mathbf{X}_t \\ 0 & -\varepsilon_{t-1}^2 \end{pmatrix}. \end{aligned}$$

Since $\boldsymbol{\theta}\in\Theta\subset R^{p+3}$ ia a compact set and $\nabla Q_T(\boldsymbol{\theta})$ is a continuous for every $\boldsymbol{\theta}\in\Theta$, then $\{\nabla Q_T(\boldsymbol{\theta})\}$ is a compact set. In other words, there exists a $M < \infty$

such that $\|\nabla Q_T()\| < M$ for every $T \in N$. By using theorem 3.2 we get $\{Q_T(\boldsymbol{\theta})\}$ equicontinuous which implies

$$\sup_{\boldsymbol{\theta}\in\Theta} |Q_T(\boldsymbol{\theta}) - Q(\boldsymbol{\theta})| \xrightarrow{a.s} 0 \text{ for } T \to \infty.$$

By inequality (18) we observe that

$$\left| Q_T(\hat{\boldsymbol{\theta}}_{SLS}) - Q(\boldsymbol{\theta}_0) \right| \leqslant \sup_{\boldsymbol{\theta}\in\Theta} |Q_T(\boldsymbol{\theta}) - Q(\boldsymbol{\theta})| \xrightarrow{a.s} 0.$$

This implies that $Q_T(\hat{\boldsymbol{\theta}}_{SLS}) \xrightarrow{a.s} Q(\boldsymbol{\theta}_0)$. Since $Q(\boldsymbol{\theta})$ has a unique minimum we have $\hat{\theta}_{SLS} \xrightarrow{a.s} \theta_0$. ■

Assumption 4 $\mathbf{A}_0 = E\left(\frac{\partial \rho_t'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathbf{W}_t \rho_t(\boldsymbol{\theta}) \rho_t'(\boldsymbol{\theta}) \mathbf{W}_t \frac{\partial \rho_t'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)$ is a nonsingular matrix.

Assumption 5 $E\left\|q_t(\boldsymbol{\theta})|\Im_{t-1}\right\|^4 < \infty$ where $q_t(\boldsymbol{\theta}) = \frac{\partial \rho_t'(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \mathbf{W}_t \rho_t(\boldsymbol{\theta})$

Theorem 3.4. *Under Assumptions 4 and 5, as $T \to \infty$,*

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_{SLS} - \boldsymbol{\theta}_0) \xrightarrow{d} N(\mathbf{0}, \mathbf{B}_0^{-1}\mathbf{A}_0\mathbf{B}_0^{-1}),$$

where

$$\mathbf{A}_0 = E\left[\frac{\partial \rho_t'(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \mathbf{W}_t \rho_t(\boldsymbol{\theta}_0) \rho_t'(\boldsymbol{\theta}_0) \mathbf{W}_t \frac{\partial \rho_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right]$$

and

$$\mathbf{B}_0 = E\left(\frac{\partial \rho_t'(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \mathbf{W}_t \frac{\partial \rho_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right).$$

Proof. Since $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} Q_T(\boldsymbol{\theta})$, we have $\frac{\partial Q_T(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}} = 0$. By equation Taylor expansion in Θ

$$\sqrt{T}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) = -\left(\frac{1}{T}\frac{\partial^2 Q_T(\tilde{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}\right)^{-1} \frac{1}{\sqrt{T}} \frac{\partial Q_T(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}}. \tag{19}$$

Using the equation (19), the asymptotic distribution of $\sqrt{T}(\hat{\boldsymbol{\theta}}_{SLS} - \boldsymbol{\theta}_0)$ will be normal if:

(1) $\frac{1}{\sqrt{T}} \frac{\partial Q_T(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} = \frac{1}{\sqrt{T}} \sum \frac{\partial q_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \xrightarrow{d} N(\mathbf{0}, 4\mathbf{A}_0)$ for nonrandom $\mathbf{A}_0 > 0$

(2) $\frac{1}{T} \frac{\partial^2 Q_T()}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \xrightarrow{p} 2\mathbf{B}_0$ for nonrandom $\mathbf{B}_0 > 0$.

The method of the proof is to show that two conditions are satisfied.

(1) Since $E(\varepsilon_t|\Im_t) = 0$ and $E\left(Y_t^{(\lambda)2} - E(Y_t^{(\lambda)2}|\Im_t)\right) = 0$ then $E\left[\frac{\partial q_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}}\right] = 0$ and we have

$$\begin{aligned} E\left(\frac{\partial q_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \frac{\partial q_t(\boldsymbol{\theta}_0)'}{\partial \boldsymbol{\theta}}\right) &= 4E\left(\frac{\partial \rho_t'(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}} \mathbf{W}_t \rho_t(\boldsymbol{\theta}_0) \rho_t'(\boldsymbol{\theta}_0) W_t \frac{\partial \rho_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}\right) \\ &= 4A_0. \end{aligned}$$

Furthermore we can apply a Martingale central limit theorem (Billingsley, 1961 and 1965), so we obtain:

$$\frac{1}{\sqrt{T}}\frac{\partial Q_T(\theta_0)}{\partial\theta} = \frac{1}{\sqrt{T}}\sum\frac{\partial q_t(\theta_0)}{\partial\theta} \xrightarrow{d} N\left(\mathbf{0}, 4A_0\right).$$

(2) The second derivative of $Q_T(\boldsymbol{\theta})$ is given by

$$\frac{\partial^2 Q_T(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} = \frac{2}{T}\sum\left[\frac{\partial\rho_t'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}W_t\frac{\partial\rho_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} + (\rho_t(\boldsymbol{\theta})\mathbf{W}_t\otimes I_{p+4})\frac{\partial vec\left(\partial\rho_t'(\boldsymbol{\theta})/\partial\boldsymbol{\theta}\right)}{\partial\boldsymbol{\theta}'}\right] \tag{20}$$

where

$$\frac{\partial vec\partial\rho_t'(\theta)/\partial\theta}{\partial\theta'} = \begin{bmatrix} \partial^2 y_t^{(\lambda)}/\partial\lambda^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \partial^2 y_t^{(\lambda)2}/\partial\lambda^2 & 0 & 0 \\ 0 & -2\partial(\mathbf{X}_t'\beta\mathbf{X}_t)/\partial\beta' & 0 \\ 0 & 0 & -\partial^2 h_t/\partial\alpha\partial\alpha' \end{bmatrix}.$$

By the ergodic theory, we get $\frac{1}{T}\sum\frac{\partial\rho_t'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{W}_t\frac{\partial\rho_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'} \xrightarrow{p} E\left(\frac{\partial\rho_t'(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\mathbf{W}_t\frac{\partial\rho_t(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}'}\right)$. Therefore, based on the equation (20), we obtain

$$\frac{1}{T}\frac{\partial^2 Q_T(\tilde{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'} \xrightarrow{p} 2\mathbf{B}_0$$

for nonrandom $\mathbf{B}_0 > 0$.

■

4. TESTING

Since Box-Cox transformed ARCH model is a generalization of the original ARCH model in which dependent variable has been transformed by the Box-Cox transformation, we need to test whether linear ARCH model provides an adequate description of the data or not. From section (3) we obtain that SLS estimators of Box-Cox transformed ARCH model are asymptotically normal in probability, so we can use z-test in the linearity testing. Consider testing a hypothesis about the first of coefficient θ. Theorem 3.3 implies that under the $H_0 : \lambda = 1$ (i.e., the linear ARCH regression model),

$$\sqrt{T}(\hat{\lambda}_{SLS} - \lambda) \xrightarrow{d} N(0, \text{var}(\hat{\lambda}_{SLS}))$$

and

$$\widehat{\text{var}}(\hat{\lambda}_{SLS}) \xrightarrow{p} \text{var}(\hat{\lambda}_{SLS}),$$

where $\widehat{\text{var}}(\hat{\lambda}_{SLS})$ is the (1,1) element of the $(P+3)\times(P+3)$ matrix $\hat{\mathbf{B}}_0^{-1}\hat{\mathbf{A}}_0\hat{\mathbf{B}}_0^{-1}$, where

$$\hat{\mathbf{B}}_0 = \frac{1}{T}\sum \frac{\partial \rho_t'(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}}\mathbf{W}_t\frac{\partial \rho_t(\hat{\boldsymbol{\theta}})}{\partial \boldsymbol{\theta}'},$$

and

$$\hat{\mathbf{A}}_0 = \frac{1}{T}\sum \frac{\partial \rho_t'(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}}\mathbf{W}_t\rho_t(\boldsymbol{\theta}_0)\rho_t'(\boldsymbol{\theta}_0)\mathbf{W}_t\frac{\partial \rho_t(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}'}.$$

The test statistics of the hypothesis is

$$t = \frac{\sqrt{T-(p+3)}\hat{\lambda}_{SLS}-1}{\sqrt{\widehat{\text{var}}(\hat{\lambda}_{SLS})}} \to t_{T-(P+3)}.$$

5. SIMULATION

In order to study the performance of the SLS estimators of $\boldsymbol{\theta}$ in finite samples, we simulated 100 series that is generated from ARCH(1) process with samples size $T = 50, 100, 200, 350$:

$$Y_t^{(\lambda)} = \beta Y_{t-1}^{(\lambda)} + \varepsilon_t, \tag{21}$$

and

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2. \tag{22}$$

where ε_t has mean zero and variance h_t. We use values of the parameters of the model in which $\lambda = 0.25, \beta = 0.2, \alpha_0 = 0.60$, and $\alpha_1 = 0.15$. Figure 1-4 show the results, where SLS estimators go to the true value when $T \to \infty$. We report the Monte Carlo means and their mean squared errors (MSE) on the Table 1. The results show that SLS estimators performance are well, where the MSE decreases. The p-value in the table is less than 0.05, it means that λ value is significantly different from one. As the result, nonlinear ARCH model is adequate.

Table 1. SLS estimators of model (21) and (22)

	$\lambda = 0.25$			$\beta = 0.20$		$\alpha_0 = 0.60$		$\alpha_1 = 0.15$	
T	$\hat{\lambda}_{SLS}$	MSE	p-value	$\hat{\beta}_{SLS}$	MSE	$\hat{\alpha}_{0.SLS}$	MSE	$\hat{\alpha}_{1.SLS}$	MSE
50	0.255	0.046	0.0005	0.163	0.024	0.618	0.032	0.063	0.058
100	0.253	0.018	0.0003	0.209	0.013	0.603	0.012	0.078	0.030
200	0.253	0.010	0.0003	0.194	0.007	0.602	0.008	0.109	0.029
350	0.252	0.006	0.0002	0.197	0.003	0.602	0.004	0.120	0.025

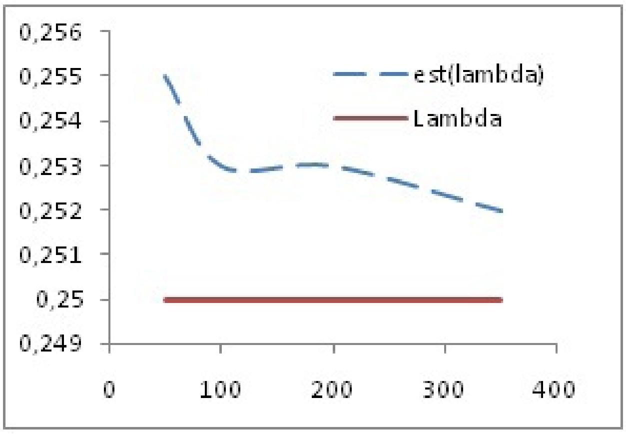

Figure 1. SLS estimation of λ

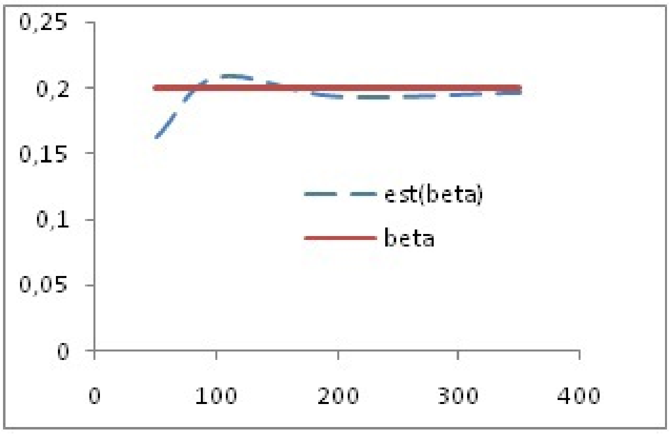

Figure 2. SLS estimation of β

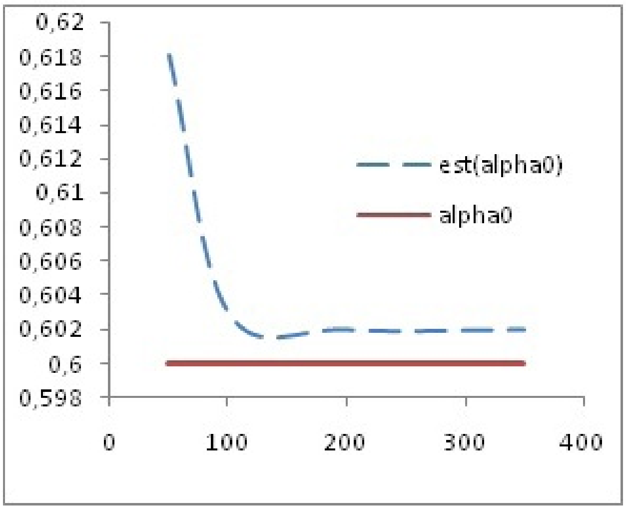

Figure 3. SLS estimation of α_0

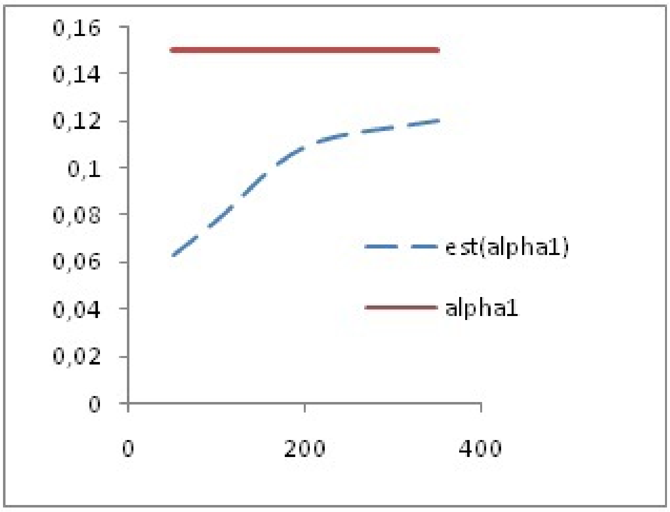

Figure 4. SLS estimation of α_1

REFERENCES

[1] Abarin, T., and Wang, L., "Comparison GMM with Second-Order Least Square Estimation in Nonlinear Models", *Far East Journal of Theoritical Statistics*, **20(2)** (2006), 179-196.

[2] Alberola, R., "Estimating Volatility Returns Using ARCH Models. An Empirical Case: The Spanish Energy Market", *Lect. Econ*, **66** (2006), 251-275.

[3] Basawa, I.V., Feigin, P.D., and Heyde, C.C., "Asymptotic Properties of Maximum Likelihood Estimators for Stochastic Processes", *The Indian Journal of Statistics*, **28(3)** (1976), 259-270.

[4] Berndt, E.K., Hall, R.E., Hausman, J.A., "Estimation and Inference in Nonlinear Structural Models", *Ann. Econom. Social Measurement*, **4** (1974), 653-665.

[5] Billingsley, P., "The Lindeberg-Levy Theorem for Martingales", *Proceedings of The American Mathematical Society*,**12(5)** (1961), 788-792.

[6] Billingsley, P., *Ergodic Theory and Information*, Wiley, New York, 1965.

[7] Brockwell, P.J. and Davis, R.A., *Time Series: Theory and Methods*, Springer-Verlag, Corolado, 1990.

[8] Capinski, M., Kopp, E., *Measure, Integral and Probability*, Springer, New York, 2003.

[9] Chan, N., H., *Time Series: Applications to Finance*, John Wiley and Sons, Canada, 2002.

[10] Draper, N.R., Cox, D.R., "On Distributions and Their Trnsformation to Normality", *Royal statistical Society- Series B* **38** (1969), 472-476.

[11] Engle, R.F., "Autoregressive Conditional Heteroscedasticity with Estimates of The Variance of U.K. Inflation", *Econometrica*, **50** (1982), 987-1008.

[12] Gao, S., He Yu, Chen, H., "Wind Speed Forecast for Wind Farms Based on ARMA-ARCH Model", *IEEE* (2009), 1-4.

[13] Hannan, E.J., *Multiple Time Series*, Wiley, New York, 1970.

[14] Hayashi, F., *Econometric*, Princeton University Press, United Kingdom, 2000.

[15] Hardle, W and Hafner, C.M., "Discrete Time Option Pricing with Flexible Volatility Estimation", *Journal Finance and Stochaties*, **4(2)** (2000), 189-207.

[16] Lamoureux, Christoper G., Lastrapes, William D., "Heteroscedasticity in Stock Return Data: Volume Vesus GARCH Effect", *Journal of Finance*, **45(1)** (1990), 221-229.

[17] Meyn, S.P., Tweedie, R.L., *Markov Chains and Stochastic Stability*, Springer-Verlag, 1993.

[18] Poiries, D.J., "The Use of the Box-Cox Transformation in Limited Dependent Variable Models", *Journal of the American Statistical Association*, **73** (1978), 284-287.

[19] Rao, C.R., *Linear Statistical Inference and Its Applications*, Wiley, Canada, 2002.

[20] Sarkar, N., "ARCH model with Box-Cox Transformed Dependent Variable", *Statistics and Probability Letters*, **50** (2000), 365-374.

[21] Wang, L., "Estimation of Nonlinear Berkson-Type Measurement Error Models", *Statistica Sinica*, **13** (2003), 1201-1210.

[22] Wang, L., "Estimation of Nonlinear Models with Berkson Measurement Error", *Annals of Statistics*, **32** (2004), 2559-2579.

[23] Wang, L., Leblanc, A., "Second-Order Nonlinear Least Square Estimation", *Ann Inst Math* **60** (2008), 883-900.

[24] Weiss, A.A., "Asymptotic Theory for ARCH Model: Estimating and Testing", *Econometric Theory* **2(1)** (1986), 107-131.