The Maximum Likelihood Method of Random Coefficient Dynamic Regression Model
Autcha Araveeporn

Abstract—The Random Coefficient Dynamic Regression (RCDR) model is to developed from Random Coefficient Autoregressive (RCA) model and Autoregressive (AR) model. The RCDR model is considered by adding exogenous variables to RCA model. In this paper, the concept of the Maximum Likelihood (ML) method is used to estimate the parameter of RCDR(1,1) model. Simulation results have shown the AIC and BIC criterion to compare the performance of the the RCDR(1,1) model. The variables as the stationary and weakly stationary data are good estimates where the exogenous variables are weakly stationary. However, the model selection indicated that variables are nonstationarity data based on the stationary data of the exogenous variables.

Keywords—Autoregressive, Maximum Likelihood Method, Nonstationarity, Random Coefficient Dynamic Regression, Stationary.

I. INTRODUCTION

Most data are collected in the form of time series that often exhibits nonstationarity and stationary models. The nonstationarity models might be caused by several aspects including changes in trend volatility and random walk. The heteroscedasticity or volatility has been modeled in the literature by various authors, for instance, [1], [2] evaluated risk in finance, [3] monitored the reliability of nonlinear prediction. The stationary process does not change when shifted in time or space. The stationary models have been widely used in the time series data modeling such as the AutoRegressive (AR) model, Moving Average (MA) model and AutoRegressive Moving Average (ARMA) model.

There are several volatility models in time series, starting by [4] who introduced AutoRegressive Conditional Heteroscedastic model (ARCH) which was obtained the predictive variance for U.K. inflation rate. To obtain more flexibility, the ARCH model has been extended by [5] who produced the Generalized ARCH (GARCH) model. The GARCH model is allowed the past data time series and the past volatility. The stationary process does not change when shifted in time or space. The stationary models have been widely used in the time series data modeling such as the AutoRegressive (AR) model, Moving Average (MA) model and AutoRegressive Moving Average (ARMA) model.

The Maximum Likelihood Method of Random Coefficient Dynamic Regression (RCDR) model is to developed from RCA model and Autoregressive (AR) model. The RCDR model is considered by adding exogenous variables to RCA model. In this paper, the concept of the Maximum Likelihood (ML) method is used to estimate the parameter of RCDR(1,1) model. Simulation results have shown the AIC and BIC criterion to compare the performance of the the RCDR(1,1) model. The variables as the stationary and weakly stationary data are good estimates where the exogenous variables are weakly stationary. However, the model selection indicated that variables are nonstationarity data based on the stationary data of the exogenous variables.

Keywords—Autoregressive, Maximum Likelihood Method, Nonstationarity, Random Coefficient Dynamic Regression, Stationary.

I. INTRODUCTION

Most data are collected in the form of time series that often exhibits nonstationarity and stationary models. The nonstationarity models might be caused by several aspects including changes in trend volatility and random walk. The heteroscedasticity or volatility has been modeled in the literature by various authors, for instance, [1], [2] evaluated risk in finance, [3] monitored the reliability of nonlinear prediction. The stationary process does not change when shifted in time or space. The stationary models have been widely used in the time series data modeling such as the AutoRegressive (AR) model, Moving Average (MA) model and AutoRegressive Moving Average (ARMA) model.

There are several volatility models in time series, starting by [4] who introduced AutoRegressive Conditional Heteroscedastic model (ARCH) which was obtained the predictive variance for U.K. inflation rate. To obtain more flexibility, the ARCH model has been extended by [5] who produced the Generalized ARCH (GARCH) model. The GARCH model is allowed the past data time series and the past volatility. The stationary process does not change when shifted in time or space. The stationary models have been widely used in the time series data modeling such as the AutoRegressive (AR) model, Moving Average (MA) model and AutoRegressive Moving Average (ARMA) model.

There are several volatility models in time series, starting by [4] who introduced AutoRegressive Conditional Heteroscedastic model (ARCH) which was obtained the predictive variance for U.K. inflation rate. To obtain more flexibility, the ARCH model has been extended by [5] who produced the Generalized ARCH (GARCH) model. The GARCH model is allowed the past data time series and the past volatility. The stationary process does not change when shifted in time or space. The stationary models have been widely used in the time series data modeling such as the AutoRegressive (AR) model, Moving Average (MA) model and AutoRegressive Moving Average (ARMA) model.

The Maximum Likelihood Method of Random Coefficient Dynamic Regression (RCDR) model is to developed from RCA model and Autoregressive (AR) model. The RCDR model is considered by adding exogenous variables to RCA model. In this paper, the concept of the Maximum Likelihood (ML) method is used to estimate the parameter of RCDR(1,1) model. Simulation results have shown the AIC and BIC criterion to compare the performance of the the RCDR(1,1) model. The variables as the stationary and weakly stationary data are good estimates where the exogenous variables are weakly stationary. However, the model selection indicated that variables are nonstationarity data based on the stationary data of the exogenous variables.

Keywords—Autoregressive, Maximum Likelihood Method, Nonstationarity, Random Coefficient Dynamic Regression, Stationary.

I. INTRODUCTION

Most data are collected in the form of time series that often exhibits nonstationarity and stationary models. The nonstationarity models might be caused by several aspects including changes in trend volatility and random walk. The heteroscedasticity or volatility has been modeled in the literature by various authors, for instance, [1], [2] evaluated risk in finance, [3] monitored the reliability of nonlinear prediction. The stationary process does not change when shifted in time or space. The stationary models have been widely used in the time series data modeling such as the AutoRegressive (AR) model, Moving Average (MA) model and AutoRegressive Moving Average (ARMA) model.

The Maximum Likelihood Method of Random Coefficient Dynamic Regression (RCDR) model is to developed from RCA model and Autoregressive (AR) model. The RCDR model is considered by adding exogenous variables to RCA model. In this paper, the concept of the Maximum Likelihood (ML) method is used to estimate the parameter of RCDR(1,1) model. Simulation results have shown the AIC and BIC criterion to compare the performance of the the RCDR(1,1) model. The variables as the stationary and weakly stationary data are good estimates where the exogenous variables are weakly stationary. However, the model selection indicated that variables are nonstationarity data based on the stationary data of the exogenous variables.

Keywords—Autoregressive, Maximum Likelihood Method, Nonstationarity, Random Coefficient Dynamic Regression, Stationary.
time series dynamic modeling when the observations of time series data have correlated with exogenous variables, the dynamic modeling will help accurate the coefficient model.

Essentially, we will extend the RCA model by adding the exogenous variables of \( y_t \) denoted as

\[
\begin{align*}
  x_t &= \alpha_t + \sum_{i=1}^{p} \beta_i x_{t-i} + \sigma \varepsilon_t \\
  \beta_t &= \mu \beta + \sigma \varepsilon_t \\
  \alpha_t &= \sum_{j=1}^{q} \eta_j y_{t-j} + \varepsilon_t
\end{align*}
\]

(2)

The \( \varepsilon_t \) and \( \mu \beta \) are the sequences of independent of random vectors with mean zero and unit variance, so the model (2) is called Random Coefficient Dynamic Regression (RCDR) Model or RCDR(p,q) model.

We will consider the simplified case of RCDR model with \( p = q = 1 \) and \( \sigma = 1 \); denoted by the RCDR(1,1), and we can rewrite as

\[
\begin{align*}
  x_t &= \alpha_t + \beta_t x_{t-1} + \varepsilon_t \\
  \beta_t &= \mu \beta + \sigma \varepsilon_t \\
  \alpha_t &= \eta y_{t-1} + \varepsilon_t
\end{align*}
\]

(3)

where \( \beta_t \)'s are iid random variables with mean \( \mu \beta \) and variance \( \sigma \beta^2 \), \( \varepsilon_t \)'s are iid random variables with mean 0 and variance \( \sigma^2 \), and \( \beta_t \)'s and \( \varepsilon_t \)'s are independent.

The parameters of RCDR(1, 1) consist of the intercept term \( \eta \), the mean \( \mu \beta \) and variance \( \sigma \beta^2 \) of the coefficient \( \beta_t \) and the variance \( \sigma^2 \) of the \( \varepsilon_t \), or defined as \( \theta = (\eta, \mu \beta, \sigma \beta^2, \sigma^2) \). In the literature, there is the RCA(1) with the slight modifications to model setup that used the nature of problem at hand might motivate to assume the two random variables \( \beta_t \) and \( \varepsilon_t \) to be correlated, see [11].

### III. PARAMETER ESTIMATION FOR RCDR(1, 1)

The method of maximum likelihood has been widely used in estimation. For any set of observations, \( x_1, \ldots, x_n \), time series or not, the likelihood function \( L(\theta) \) is define to be the joint probability density of obtaining the data actually observed. However, it is considered as a function of the unknown parameters in the model with the observed data held fixed.

To estimate parameter of RCDR(1,1) model, we propose the maximum likelihood method to estimate parameter \( \theta = (\eta, \mu \beta, \sigma \beta^2, \sigma^2) \). The time series data \( \{x_t\} \) and \( \{y_t\} \) from (3) obtain following:

\[
\begin{align*}
  E(x_t | x_{t-1}) &= \alpha_t + \mu \beta x_{t-1} \\
  Var(x_t | x_{t-1}) &= (1 + \eta^2) \sigma^2 + \sigma^2 \beta^2 x_{t-1}^2
\end{align*}
\]

The maximum likelihood method consider the likelihood function from (3) to estimate \( \mu \beta \) as

\[
L(\theta) = L(\theta | x_t, x_{t-1}) = \prod_{t=2}^{n} f(x_t | x_{t-1})
\]

\[
= \left( \frac{1}{2\pi} \right)^{n/2} \prod_{t=2}^{n} \left[ \left( 1 + \eta^2 \right) \sigma^2 + \sigma^2 \beta^2 x_{t-1}^2 \right]^{-1/2} \exp \left\{ -\frac{1}{2} \sum_{t=2}^{n} \left( x_t - \alpha_t + \mu \beta x_{t-1} \right)^2 \right\}
\]

(4)

Constructing the new likelihood function by setting parameter, let \( \eta^2 = \omega, \tau = \frac{\sigma^2}{\omega} \), and substitute \( \alpha_t = \eta y_{t-1} \), so it show that

\[
L(\theta) = \left( \frac{1}{2\pi} \right)^{n/2} \prod_{t=2}^{n} \left[ \sigma^2 \right]^{-1/2} \left[ (1 + \omega) + \tau x_{t-1}^2 \right]^{-1/2} \exp \left\{ -\frac{1}{2} \sum_{t=2}^{n} \left( x_t - \eta y_{t-1} - \mu \beta x_{t-1} \right)^2 \right\}
\]

(5)

The ln likelihood function following;

\[
\begin{align*}
  \ln L(\theta) &= -\frac{n}{2} \ln (2\pi) - \frac{n}{2} \ln \sigma^2 \\
  &-\frac{1}{2} \sum_{t=2}^{n} \ln \left[ (1 + \omega) + \tau x_{t-1}^2 \right] \\
  &-\left\{ \frac{1}{2} \sum_{t=2}^{n} \left( x_t - \eta y_{t-1} - \mu \beta x_{t-1} \right)^2 \right\}
\end{align*}
\]

(6)

The next step is differentiable from (6) with respect to \( \mu \beta, \eta, \sigma^2 \)

\[
\begin{align*}
  \frac{\partial \ln L(\theta)}{\partial \mu \beta} &= \sum_{t=2}^{n} \frac{(x_t - \eta y_{t-1} - \mu \beta x_{t-1}) x_{t-1}}{\sigma^2 \left[ (1 + \omega) + \tau x_{t-1}^2 \right]} \\
  \frac{\partial \ln L(\theta)}{\partial \eta} &= \sum_{t=2}^{n} \frac{(x_t - \eta y_{t-1} - \mu \beta x_{t-1}) y_{t-1}}{\sigma^2 \left[ (1 + \omega) + \tau x_{t-1}^2 \right]} \\
  \frac{\partial \ln L(\theta)}{\partial \sigma^2} &= -\frac{n}{2\sigma^2} \sum_{t=2}^{n} \frac{(x_t - \eta y_{t-1} - \mu \beta x_{t-1})^2}{(1 + \omega) + \tau x_{t-1}^2} + \frac{1}{2\sigma^4} \sum_{t=2}^{n} \frac{(x_t - \eta y_{t-1} - \mu \beta x_{t-1})^2}{(1 + \omega) + \tau x_{t-1}^2}
\end{align*}
\]

(7)

(8)

(9)

Now we get

\[
\frac{\partial \ln L(\theta)}{\partial (\mu \beta, \eta, \sigma^2)} = 0
\]

We obtain the estimators;

\[
\begin{align*}
  \hat{\mu \beta} &= \frac{a_1 - \hat{\eta} a_2}{a_3} \\
  \hat{\eta} &= \frac{a_4 - \hat{\mu \beta} a_2}{a_5} \\
  \hat{\sigma^2} &= (n)^{-1} \sum_{t=2}^{n} \frac{(x_t - \hat{\eta} y_{t-1} - \hat{\mu \beta} x_{t-1})^2}{(1 + \omega) + \tau x_{t-1}^2}
\end{align*}
\]

(10)

(11)

(12)
That is, we have profiled the log-likelihood as a function of \( \eta \) only. If the sample sizes become infinite. In this section, we will consider properties as \( n \to \infty \).

The ML estimates \( \hat{\eta}, \hat{\mu}_s, \hat{\sigma}^2 \), and \( \hat{\sigma}_\beta^2 \) can be obtained by calculating \( \hat{\tau} \), where \( \hat{\tau} \) is the minimizer of the following function of \( \tau \),

\[
g(\tau) = \ln(\sigma^2) + \sum_{t=2}^{n} \ln((1 + \omega) + \tau x_{t-1}^2)
\]

That is, we have profiled the log-likelihood as a function of \( \tau \) only.

The ML estimates \( \hat{\eta}, \hat{\mu}_s, \hat{\sigma}^2 \), and \( \hat{\sigma}_\beta^2 \) are obtained by

\[
\ln L(\hat{\theta}) = -\inf_\theta \ln L(\theta)
\]

\[
\ell(\hat{\eta}, \hat{\mu}_s, \hat{\sigma}^2, \hat{\sigma}_\beta^2) = -\inf_{(\eta, \mu_s, \sigma^2, \sigma_\beta^2)} \ln L(\eta, \mu_s, \sigma^2, \sigma_\beta^2)
\]

\[
\hat{\sigma}^2 = (n-1) \sum_{t=2}^{n} \frac{(x_t - \hat{\eta} y_{t-1} - \hat{\mu}_s x_{t-1})^2}{(1 + \omega) + \hat{\tau} x_{t-1}^2}
\]

and

\[
\hat{\omega} = \hat{\eta}^2
\]

A. The Properties of ML Estimators

For the point estimation, we might consider properties as if the sample sizes becomes infinite. In this section, we will look at the properties of ML estimators: consistency and asymptotic efficiency.

1) Consistency

The consistency of

\[
\theta = (\eta_1, \ldots, \eta_p, \mu_{j1}, \ldots, \mu_{j_p}, \sigma_1^2, \ldots, \sigma_p^2, \sigma_\beta^2)^T
\]

will be shown by examining

\[
\lim_{n \to \infty} P_\theta(|W_n - \theta| \geq \epsilon) = 0
\]

where \( W_n = W_n(x_1, \ldots, x_n) \) is a consistent sequence of estimators of parameter \( \theta \). Recall that, for an estimator \( W_n \), Chebychev’s Inequality states

\[
P_\theta(|W_n - \theta| \geq \epsilon) \leq \frac{E_\theta(|W_n - \theta|)^2}{\epsilon^2}
\]

For the second term, we have

\[
E \left( \frac{(x_t - \sum_{j=1}^{q} \eta_j y_{t-j} - \sum_{i=1}^{p} \mu_{ji} x_{t-i})^2}{(1 + \sum_{j=1}^{q} \omega_j) + \sum_{i=1}^{p} \tau_i x_{t-i}^2} \right)
\]

\[
\leq \left| E \left( \frac{(x_t - \sum_{j=1}^{q} \eta_j y_{t-j} - \sum_{i=1}^{p} \mu_{ji} x_{t-i})^2}{(1 + \sum_{j=1}^{q} \omega_j) + \sum_{i=1}^{p} \tau_i x_{t-i}^2} \right) \right|
\]

\[
< \infty
\]

Therefore, we can conclude that

\[
\left| E \left( \frac{(x_t - \sum_{j=1}^{q} \eta_j y_{t-j} - \sum_{i=1}^{p} \mu_{ji} x_{t-i})^2}{(1 + \sum_{j=1}^{q} \omega_j) + \sum_{i=1}^{p} \tau_i x_{t-i}^2} \right) \right| = \frac{2}{n}
\]

Moreover, (13) says that the sample size becomes infinite, the estimators will be arbitrarily close to the parameter with zero in probability.

2) Asymptotic Efficiency [12]

Let \( x_1, \ldots, x_n \) be iid \( f(x|\theta) \), let \( \hat{\theta} \) denote the ML estimator of \( \theta \), and let \( W_n \) be a continuous function of \( \theta \).

\[
\sqrt{n} |W_n - \theta| \to n(0, \nu(\theta))
\]

where \( \nu(\theta) \) is the Cramér-Rao Lower Bound. That is, \( W_n \) is a consistent and asymptotically efficient estimator of \( \theta \).

Under the property of consistency, the variance of estimator is

\[
Var(\hat{\theta}) = E_\theta(W_n - \theta)^2 \approx \frac{1}{n} E_\theta(\frac{\partial \ln L(\theta)}{\partial \theta})^2
\]

(13)

Suppose that

\[
\sqrt{n} \left( \frac{W_n - \theta}{\sigma} \right) \to Z \text{ in distribution}
\]

where \( Z \sim \text{Normal}(0,1) \). By applying Slutsky’s Theorem we conclude

\[
W_n - \theta = \frac{\sigma}{\sqrt{n}} \left( \sqrt{n} \frac{W_n - \theta}{\sigma} \right) \to \lim_{n \to \infty} \left( \frac{\sigma}{\sqrt{n}} \right) Z = 0
\]

so \( W_n - \theta \to 0 \) in distribution. We know that convergence in distribution to a point is equivalent to convergence in probability, so \( W_n \) is consistent estimator of \( \theta \).

IV. A SIMULATION STUDY

The simulation study to estimate parameter \( \theta = (\eta_1, \mu_{j1}, \sigma_1^2, \sigma_\beta^2)^T \) for the performance of ML method. At the beginning, we generate data \( y_t, t = 1, 2, \ldots, n \) from the AR(1) model by taking \( \eta = 0.1, 0.5 \) and 0.9 following;

\[
\alpha_t = \eta y_{t-1} + \varepsilon_t : AR(1)
\]

(14)

To illustrate the implication of AR model, Figure 1 shows the 100 sample sizes for each 3 coefficients (\( \eta = 0.1, 0.5 \) and 0.9). It should be noted that \( \eta = 0.1 \) is stationary, \( \eta = 0.5 \) is weakly stationary, and \( \eta = 0.9 \) is the nonstationary case.
Next, we consider the RCDR(1, 1) model where \( y_t \) are generated from the AR(1). Therefore we obtain the \( x_t \) in terms of RCDR(1, 1) written as

\[
x_t = \eta y_{t-1} + \beta_1 x_{t-1} + \epsilon_t
\]  

In Figures 3-4, we present the data generating in 6 cases:

1) \( \sigma^2 = 1, \mu_\beta = 0.5 \) and \( \sigma^2_\beta = 0.25 \)
2) \( \sigma^2 = 1, \mu_\beta = 0.995 \) and \( \sigma^2_\beta = 0.01 \)
3) \( \sigma^2 = 1, \mu_\beta = 0.1 \) and \( \sigma^2_\beta = 0.99 \)
4) \( \sigma^2 = 1, \mu_\beta = -0.995 \) and \( \sigma^2_\beta = 0.01 \)
5) \( \sigma^2 = 1, \mu_\beta = -0.1 \) and \( \sigma^2_\beta = 0.99 \)
6) \( \sigma^2 = 1, \mu_\beta = 0 \) and \( \sigma^2_\beta = 1 \)

It should be noted that Case 2 is the nonstationary case and the Case 4 tends to be around its mean value of 0 as the stationary process. For Case 1, 3, 5, and 6, we can’t define the character of the time series plot when there are slightly different figures that depended on the \( \eta \) under \( y_t \). In order to assess the model performance, it is customary to use some type of model selection criteria such as Akaike Information Criterion (AIC) introduced by [13] and Bayesian Information Criterion (BIC) studied by [14]. In our simulation studies we explore the performance of the RCDR(1, 1) model in picking up the true models. AIC and BIC are defined as,

\[
AIC(\theta) = -2 \ln L(\theta) + 2m
\]

\[
BIC(\theta) = -2 \ln L(\theta) + m \ln(n)
\]

where \( \ln L(\cdot) \) is the log-likelihood function, \( m \) is the number of parameters in the model, and \( n \) is the number of the sample sizes.

The selection of the chosen model is then made by considering the smallest AIC and BIC in each case.

The results of the simulations were carried out using R program that was used to generate data and performed the parameter values from ML method. We simulated data with the sample sizes \( n = 100 \) and \( 500 \), and repeated the data generation for model fitting 500 times. Tables I-III show various Monte Carlo (MC) of the estimates obtained by taking \( \eta = 0.1, 0.5 \) and 0.9 in 6 cases.

The third and the fourth columns of these tables represent the AIC and BIC criterion to perform in picking up the chosen model when the estimators are fitted. From Tables I, it appears that both AIC and BIC are performing reason-
ably well when the data are generated from true parameter $\sigma^2 = 1$, $\mu = 0$, $\sigma_\beta^2 = 1$ at sample sizes $n = 100$, and $\sigma^2 = 1$, $\mu = 0.1$, $\sigma_\beta^2 = 0.99$ at sample sizes $n = 500$ when $\eta = 0.1$. However, the $\eta = 0.5$, model is good fit at parameter $\sigma^2 = 1$, $\mu = 0.1$, $\sigma_\beta^2 = 0.99$. On the other hand, the $\eta = 0.9$, parameters $\sigma^2 = 1$, $\mu = 0.1$, $\sigma_\beta^2 = 0.99$ prefer RCDE model at sample sizes $n = 100$, but the sample sizes $n = 500$ is fitted in parameter $\sigma^2 = 1$, $\mu = -0.995$, $\sigma_\beta^2 = 0.01$. For each $\eta$, the small sample sizes are fitted better than the large sample sizes.

### TABLE I

<table>
<thead>
<tr>
<th>Case</th>
<th>Parameters</th>
<th>$n$</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>157.983</td>
<td>168.383</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0.5$, $\sigma_\beta^2 = 0.25$</td>
<td>n=500</td>
<td>893.249</td>
<td>910.107</td>
</tr>
<tr>
<td>Case 2</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>159.326</td>
<td>169.745</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0.995$, $\sigma_\beta^2 = 0.01$</td>
<td>n=500</td>
<td>894.414</td>
<td>911.273</td>
</tr>
<tr>
<td>Case 3</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>158.258</td>
<td>168.679</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0.1$, $\sigma_\beta^2 = 0.99$</td>
<td>n=500</td>
<td>892.983</td>
<td>909.841</td>
</tr>
<tr>
<td>Case 4</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>158.266</td>
<td>168.687</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0.995$, $\sigma_\beta^2 = 0.01$</td>
<td>n=500</td>
<td>893.110</td>
<td>909.968</td>
</tr>
<tr>
<td>Case 5</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>158.181</td>
<td>168.602</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0.1$, $\sigma_\beta^2 = 0.99$</td>
<td>n=500</td>
<td>893.029</td>
<td>909.888</td>
</tr>
<tr>
<td>Case 6</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>157.920</td>
<td>168.341</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0$, $\sigma_\beta^2 = 1$</td>
<td>n=500</td>
<td>893.483</td>
<td>909.841</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case</th>
<th>Parameters</th>
<th>$n$</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>158.078</td>
<td>168.498</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0.5$, $\sigma_\beta^2 = 0.25$</td>
<td>n=500</td>
<td>893.443</td>
<td>910.302</td>
</tr>
<tr>
<td>Case 2</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>159.341</td>
<td>169.762</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0.995$, $\sigma_\beta^2 = 0.01$</td>
<td>n=500</td>
<td>893.975</td>
<td>910.833</td>
</tr>
<tr>
<td>Case 3</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>157.879</td>
<td>168.300</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0.1$, $\sigma_\beta^2 = 0.99$</td>
<td>n=500</td>
<td>893.264</td>
<td>910.123</td>
</tr>
<tr>
<td>Case 4</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>158.085</td>
<td>168.509</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0.995$, $\sigma_\beta^2 = 0.01$</td>
<td>n=500</td>
<td>893.275</td>
<td>910.133</td>
</tr>
<tr>
<td>Case 5</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>157.940</td>
<td>168.361</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0.1$, $\sigma_\beta^2 = 0.99$</td>
<td>n=500</td>
<td>893.302</td>
<td>910.161</td>
</tr>
<tr>
<td>Case 6</td>
<td>$\eta = 0.1$, $\sigma^2 = 1$</td>
<td>n=100</td>
<td>158.399</td>
<td>168.820</td>
</tr>
<tr>
<td></td>
<td>$\mu = 0$, $\sigma_\beta^2 = 1$</td>
<td>n=500</td>
<td>893.353</td>
<td>910.211</td>
</tr>
</tbody>
</table>

### APPENDIX A

#### THE CONDITION OF ML ESTIMATORS

To use the RCDE(1,1) model to verify that a function of parameters has a local maximum at estimators, it must be shown that the following three conditions hold [12].

1. The first-order partial derivatives are 0.

2) At least one second-order partial is negative.

This condition can be seen after the first-order partial derivatives from (7) and (8).

$$\frac{\partial^2 \ln L(\theta)}{\partial \beta^2} = - \sum_{t=2}^{n} \frac{y_t^2}{\sigma^2[(1 + \omega + \tau x_t^2)]}$$

$$\frac{\partial^2 \ln L(\theta)}{\partial \beta^2} = - \sum_{t=2}^{n} \frac{x_t^2}{\sigma^2[(1 + \omega + \tau x_t^2)]}$$

3) The Jacobian of the second-order partial derivatives is positive.

For the log likelihood function, the second-order partial derivatives can be written in the symmetric matrix and denoted

$$V = \begin{pmatrix}
V_{11} & V_{12} & V_{13} \\
V_{21} & V_{22} & V_{23} \\
V_{31} & V_{32} & V_{33}
\end{pmatrix}$$

where each element $V_{ij}$ of $V$ given by

$$V_{11} = \frac{\partial^2 \ln L(\theta)}{\partial \beta^2} = - \sum_{t=2}^{n} \frac{y_t^2}{\lambda_t}$$

$$V_{12} = \frac{\partial^2 \ln L(\theta)}{\partial \beta^2} = - \sum_{t=2}^{n} \frac{x_t \cdot y_{t-1}}{\lambda_t}$$

$$V_{13} = \frac{\partial^2 \ln L(\theta)}{\partial \beta^2} = - \sum_{t=2}^{n} \frac{x_t}{\lambda_t^2}$$

$$V_{22} = \frac{\partial^2 \ln L(\theta)}{\partial \beta^2} = - \sum_{t=2}^{n} \frac{y_{t-1}}{\lambda_t}$$

$$V_{23} = \frac{\partial^2 \ln L(\theta)}{\partial \beta^2} = - \sum_{t=2}^{n} \frac{x_t}{\lambda_t^2}$$
where $\lambda_t = \sigma^2[1 + \omega + \tau x_{t-1}^2]$ and $u_t = x_t - \eta y_{t-1} - \mu \beta x_{t-1}$.

$$
V_{33} = \frac{\partial^2 \ln L(\theta)}{\partial (\sigma^2)^2} = \sum_{t=2}^{n} \frac{1}{\lambda_t^2} - \frac{1}{n} \sum_{t=2}^{n} \frac{u_t^2}{\lambda_t^2}
$$

Thus, we can compute the Jacobian by

$$
V_{11}V_{22}V_{33} + V_{12}V_{23}V_{31} + V_{13}V_{21}V_{32}
$$

$$
- V_{31}V_{22}V_{13} - V_{32}V_{23}V_{11} - V_{12}V_{21}V_{33} > 0
$$

**References**


