Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 72(1) 2018 39–48

Unit Root and Cointegration with Logistic Errors

Nimitha John and N. Balakrishna
Cochin University of Science and Technology, Kochi

Received 20 January 2016; Revised 31 March 2017; Accepted 23 February 2018

SUMMARY

Cointegration analysis and the existence of unit root often suggest an economic relationship in the long run for more than one non stationary time series. In this paper, a unit root process and cointegration model of first order for $I(1)$ processes which allows for logistic innovation is defined. We propose the maximum likelihood estimator of the cointegrating vector from a first order vector autoregressive process. Then we develop a likelihood ratio test for unit root and cointegration associated with two time series. Monte Carlo simulations are performed to verify the finite sample properties of the estimator and the test statistic. To account for the distortions caused by the specific sample, a bootstrap test based on MLE is performed. Rubber consumption and export data are analysed to illustrate the applications of the proposed model.

AMS classifications : 91G70, 91B84, 37M10.

Keywords: Bootstrapping, Cointegration, Maximum likelihood estimation, Unit root.

1. INTRODUCTION

The methods for analysing time series are developed by assuming that the observed series is a realization of certain discrete parameter stationary stochastic process. However, a time series representing a real situation need not be stationary. Box and Jenkins (1970) argued that a non-stationary time series can be converted in to stationary one by successive differences. According to them, if an observed non-stationary series $\{X_t\}$ becomes stationary after successively differencing d times, then the original series $\{X_t\}$ is referred to as a series of integration of order d and is denoted by $I(d)$. If the series is stationary, then it is denoted by $I(0)$. The study on basic linear time series reveals that, if the series becomes stationary ARMA, after $d-$ differences then there are d unit roots in the characteristic polynomial of the underlying autoregressive model. That is, we are assuming that the non stationarity was only due to the presence of unit roots. Granger (1981) pointed out that set of all time series which achieve stationarity after differencing may have linear combinations which are stationary without differencing. Engle and Granger (1987) formalized this idea and introduced the concept of co-integration. That is, if a linear combination of several $I(d)$ time series provides a stationary series then the constituent series are said to be co-integrated. Some examples for cointegrating series are income and expenditures series, short and long term interest rates and prices of same commodity of different markets etc.

During the last decade, several estimation methods and test procedures for cointegration among nonstationary time series have appeared in literature. Systematic analysis of a set of cointegrating time series can be performed by representing them in the form of a vector autoregression or error correction models, see Engle and Granger (1987). One of the efficient methods for cointegration analysis is the maximum likelihood approach, suggested by Johansen (1988). This method starts from a vector autoregressive (VAR) model representation for a set of variables with Gaussian errors. Since the problem of cointegration and the unit root are closely related, test for cointegration can be carried out by testing for unit root from the residuals of cointegrating regression

Corresponding author: Nimitha John
E-mail address: nimitha.tj@gmail.com

series. Some of the test procedures for cointegration in the literature includes the Dickey-Fuller unit root test, Engle and Granger two step estimator, Johansen likelihood ratio test etc. Engle and Granger (1987) suggest an efficient estimation technique of the error correction model with the assumption of Gaussianity of errors. All the above mentioned theories and studies are based on the assumption that the possibly cointegrated VAR or error correction model (ECM) has normally distributed errors and, hence, they have the same likelihood function as the classical Johansen method. But in practical situations most of the series we come across are far from Gaussian and hence a study of cointegrating models with non Gaussian errors is necessary. Kim and Schmidt (1993) considered the finite sample accuracy(size) of the Dickey fuller unit root test when the errors were conditionally hetroskedastic. Lee and Tse (1996) examined the performance of Johansen likelihood ratio tests for cointegration in the presence of GARCH errors.

To the best of our knowledge, apart from the above cointegration tests based on conditionally hetroskedastic errors, to date, there is no studies on cointegration and error correction model when the innovations are non normal in their distributions. There are several standard non normal distributions in literature and each distributions may need independent attention. In this paper, we study the properties of two cointegrating time series and then model them with independent and identically distributed logistic error variables. We propose an estimation procedure for cointegrating parameters using the method of conditional maximum likelihood estimation and then develop a test procedure for unit root and cointegration when the innovation processes are generated by iid logistic distribution. Since the underlying distribution of the test statistic is well-known (asymptotic Chi-square), a bootstrap method provides a way to account for the distortions caused by the finite sample. To account that, we perform a bootstrap test based on MLE for the likelihood ratio test for cointegration.

Rest of the paper is organized as follows. In Section 2, we define a cointegration model with logistic innovations and study the likelihood based estimation in Section 3. In Sections 4 and 5 we study the problems of testing of hypothesis on unit root and cointegration. A simulation study is conducted in Section 6 followed by bootstrap method in Section 7. To illustrate the applications of our model, a data analysis is presented in Section 8. And finally in Section 9, we summarises the conclusions of the study.

2. COINTEGRATING MODEL WITH LOGISTIC INNOVATIONS

Let $\{x_{1t}\}$ and $\{x_{2t}\}$ be two cointegrating time series and both are $I(1)$. It is customary to represent the series in the form of an ECM for further analysis. Towards that end, let $\{\varepsilon_t\} = (\varepsilon_{1t}, \varepsilon_{2t})'$, t =1,2,3,... be iid logistic random variables with independent marginals. Following Engle and Granger (1987), let us write

$$x_{1t} + \beta x_{2t} = u_{1t}, u_{1t} = u_{1t-1} + \varepsilon_{1t} \quad (1)$$

$$x_{1t} + \alpha x_{2t} = u_{2t}, u_{2t} = \phi u_{2t-1} + \varepsilon_{2t}, |\phi| < 1, \quad (2)$$

where $\{(\varepsilon_{1t}, \varepsilon_{2t})'\}$, t =1,2,3,.. is a sequence of iid bivariate random variables with independent marginals following symmetric logistic distribution with probability density function of the form

$$f(\varepsilon_{it}) = \frac{e^{-\varepsilon_{it}}}{(1+e^{-\varepsilon_{it}})^2}, i = 1,2,$$

with $E(\varepsilon_{it}) = 0$ and $V(\varepsilon_{it}) = \frac{\pi^2}{3}$.

Reason for the above two series become cointegrated is as follows: The reduced form for the process in (1) and (2) will make the variables x_{1t} and x_{2t} as a linear combination of u_{1t} and u_{2t} and therefore both the series will be nonstationary (Integrated of order 1). From (1) and (2) we will get,

$$x_{1t} = \left(\frac{\alpha}{\alpha-\beta}\right) u_{1t} - \left(\frac{\beta}{\alpha-\beta}\right) u_{2t} \quad (3)$$

and

$$x_{2t} = \left(\frac{1}{\alpha-\beta}\right) u_{2t} - \left(\frac{1}{\alpha-\beta}\right) u_{1t}. \quad (4)$$

Hence from the above two equations, it is clear that x_{1t} and x_{2t} are non stationary as they are linear combinations of a stationary and a non stationary series. Since x_{1t} and x_{2t} are integrated series, equation (2) describes a stationary linear combination of the nonstationary variables. Thus the variables x_{1t} and x_{2t} are cointegrated and hence we can say that they have a long run relationship in equilibrium. But if $\phi \to 1$, then the series are uncorrelated random walks and hence they are no longer cointegrated. The model given above has been studied by Engle and Granger (1987) in detail with possibly correlated white noise and the model can be transformed in to the error correction form by subtracting the lagged values from both the sides.

Before studying the properties of the model it is convenient to reparameterise the model in (1) and (2) by subtracting the lagged values from both the sides. Let Δ be a difference operator, on applying Δ operator on x_{1t} and x_{2t} of both sides of equations (1) and (2) and after some algebra we will get,

$$\Delta x_{1t} = \delta\beta(x_{1,t-1} + \alpha x_{2,t-1}) + \xi_{1t} \text{ and}$$

$$\Delta x_{2t} = -\delta(x_{1,t-1} + \alpha x_{2,t-1}) + \xi_{2t} ,$$

where ξ_{1t} and ξ_{2t} are linear combinations of the $\{\varepsilon_{it}\}$, i=1,2 and $\delta = \frac{1-\phi}{\alpha-\beta}$. Hence the Error Correction representation becomes:

$$\Delta x_{1t} = \delta\beta z_{t-1} + \xi_{1t} \tag{5}$$

$$\Delta x_{2t} = -\delta z_{t-1} + \xi_{2t} \tag{6}$$

where $z_{t-1} = x_{1,t-1} + \alpha x_{2,t-1}$.

Here α is the cointegrating parameter. The above error correction representation has three unknown parameters and we estimate them by the method of conditional maximum likelihood.

3. MAXIMUM LIKELIHOOD ESTIMATION FOR ERROR CORRECTION MODEL

Estimation of model parameters is one of the important problems involved in modelling of Gaussian and non Gaussian time series. Tiku et.al (1999) developed estimation method for a regression model with autocorrelated errors following a shift scaled Student's t distribution. Wong and Bian (2005) extended the work of Tiku et.al to the case, where the underlying distribution is a generalised logistic distribution using the modified maximum likelihood estimators since they found their maximum likelihood estimates are intractable. However, in our model we do not encounter such a problem while estimating the cointegration parameters using logistic innovations and hence we can proceed with the estimation technique using the conditional maximum likelihood estimates. If an explicit form for the innovation density function is available, then the conditional likelihood based inference is possible for error correction model given in (5) and (6). To obtain the maximum likelihood estimation of parameters in the error correction model, the innovation random variables are assumed to follow iid logistic distribution with the joint density function:

$$f(\xi_{1t}, \xi_{2t}) = \frac{e^{-[(\Delta x_{1t} - \delta\beta z_{t-1}) + (\Delta x_{2t} + \delta z_{t-1})]}}{(1+e^{-(\Delta x_{1t} - \delta\beta z_{t-1})})^2 (1+e^{-(\Delta x_{2t} + \delta z_{t-1})})^2} \tag{7}$$

The parameter vector to be estimated is $\theta = (\alpha, \beta, \delta)'$ and the conditional log-likelihood function for the ECM is given by

$$L_T(\theta) = \sum_{t=1}^{n} \{-[(\Delta x_{1t} - \delta\beta z_{t-1}) + (\Delta x_{2t} + \delta z_{t-1})]$$

$$-2\log(1 + e^{-(\Delta x_{1t} - \delta\beta z_{t-1})}) - 2\log(1 + e^{-(\Delta x_{2t} + \delta z_{t-1})})\}.$$

The form of the above log likelihood function suggest that we have to maximize it by some numerical methods. Hence on differentiating the log-likelihood function with respect to the parameter vector θ, we will get three equations given by,

$$\frac{\partial L_T(\theta)}{\partial \delta} = \sum_{t=1}^{n} z_{t-1}(\beta - 1) - \frac{2\beta z_{t-1} e^{-(\Delta x_{1t} - \delta\beta z_{t-1})}}{(1+e^{-(\Delta x_{1t} - \delta\beta z_{t-1})})} + \frac{2 z_{t-1} e^{-(\Delta x_{2t} + \delta z_{t-1})}}{(1+e^{-(\Delta x_{2t} + \delta z_{t-1})})}\} = 0. \tag{8}$$

$$\frac{\partial L_T(\theta)}{\partial \alpha} = \sum_{t=1}^{n} \{\delta x_{2t-1}(\beta - 1 - \frac{2\beta e^{-(\Delta x_{1t} - \delta\beta z_{t-1})}}{(1+e^{-(\Delta x_{1t} - \delta\beta z_{t-1})})} + \frac{2 e^{-(\Delta x_{2t} + \delta z_{t-1})}}{(1+e^{-(\Delta x_{2t} + \delta z_{t-1})})})\} = 0. \tag{9}$$

$$\frac{\partial L_T(\theta)}{\partial \beta} = \sum_{t=1}^{n} \{\delta z_{t-1}(1 - \frac{2 e^{-(\Delta x_{1t} - \delta\beta z_{t-1})}}{(1+e^{-(\Delta x_{1t} - \delta\beta z_{t-1})})})\} = 0. \tag{10}$$

These equations are solved numerically and are illustrated using simulated samples in Table 3.

The study of cointegrating models with logistic errors uses the properties of first order autoregressive models with logistic innovations, which we discuss in the next Section.

4. UNIT ROOT TEST FOR AR(1) MODEL WITH LOGISTIC ERRORS

Dickey and Fuller (1979) developed a unit root test for cointegration among nonstationary time series when the innovations were assumed to follow Gaussian series. Some other authors have examined the size distortions of this test when the errors were conditionally hetroskedastic. In particular, our interest is to analyse the time series in the presence of non-normal innovations, specifically logistic errors. If the time series are integrated of same order and are non stationary, then test for cointegration can be carried out by developing a unit root test for the residual series of either cointegrating regression equation or of the ECM. If the residuals obtained from the error correction model are stationary, then the variables could explain a long run behaviour in the equilibrium and hence they are cointegrated.

Let us consider the first order autoregressive process $\{Y_t\}$ defined by,

$$Y_t = \phi Y_{t-1} + e_t, \tag{11}$$

where $Y_0 = 0$ and $\{e_t\}$ is a sequence of independent logistic random variables with mean zero. Note that $Y_t = e_t + \phi e_{t-1} + ... + \phi^{t-1} e_1$ and if $|\phi| < 1$, Y_t converges to a stationary process as $t \to \infty$ with $E(Y_t) = 0$ and $V(Y_t) = \pi^2/3(1-\phi^2$. If a realisation ($Y_1,Y_2,Y_3,.....,Y_n$) of a first order autoregressive time series are given, we are interested in finding an estimator of ϕ and in tests of the null hypothesis that $H_0: \phi = 1$. Mostly, the alternative hypothesis of interest, $H_1: \phi < 1$ is that the time series Y_t was generated by $Y_t = \alpha + \phi Y_{t-1} + e_t$, where $|\phi| < 1$. One can also consider the alternative of interest that the time series is generated by $Y_t = \alpha + \beta t + \phi Y_{t-1}$, where $|\phi| < 1$. Our interest is to find $\hat{\phi}$, the maximum likelihood estimator for ϕ in model (11) and hence shall obtain the test procedure of unit root under the null hypothesis. An ordinary least square estimate of the regression coefficient in the autoregressive equation is obtained to be as

$$\hat{\phi} = \left(\sum_{t=1}^{n} Y_{t-1}^2\right)^{-1}\left(\sum_{t=1}^{n} Y_{t-1}Y_t\right).$$

Let us suppose n observations, say $Y_1,Y_2,Y_3,.....,Y_n$ are available for the analysis and we shall obtain the likelihood function based on n observations generated by the model (11). The joint density function of $(e_1, e_2, e_3, ...e_n)$ is

$$\prod_{t=1}^{n} \frac{e^{-e_t}}{(1+e^{-e_t})^2},$$

and the log-likelihood function of ϕ conditioned on Y_0 is

$$L(\phi|Y_0) = -\{\sum_{t=1}^{n} (Y_t - \phi Y_{t-1}) + 2\log(1 + e^{-(Y_t - \phi Y_{t-1})})\}. \tag{12}$$

The critical points of the above log likelihood function can be obtained by setting the first derivative with respect to ϕ equal to zero.

$$\partial L(\phi|Y_1) = \sum_{t=1}^{n} \left(Y_{t-1} - 2Y_{t-1}\frac{e^{-(Y_t-\phi Y_{t-1})}}{(1+e^{-(Y_t-\phi Y_{t-1})})}\right)$$

$$= \sum_{t=1}^{n} [Y_{t-1} - 2Y_{t-1}\Gamma],$$

where $\Gamma = \frac{e^{-(Y_t-\phi Y_{t-1})}}{(1+e^{-(Y_t-\phi Y_{t-1})})}$. The first order partial derivative equation suggest that the value of ϕ that maximises (12) must satisfy $\sum_{t=1}^{n} [Y_{t-1} - 2Y_{t-1}\Gamma] = 0$. This equation can be solved by some numerical technique and if any such solution exists specifies a critical point, which is either a maximum or a minimum. It should be noted that if the partial derivative of second order is negative, then the critical point will be a maximum.

$$\frac{\partial^2 L(\phi|Y_1)}{\partial\phi\,\partial\phi'} = -\sum_{t=1}^{n} Y_{t-1}\left(2Y_{t-1}\frac{e^{-(Y_t-\phi Y_{t-1})}}{(1+e^{-(Y_t-\phi Y_{t-1})})^2}\right)$$

$$= \sum_{t=1}^{n} - 2{Y_{t-1}}^2\Gamma(1-\Gamma).$$

Now let us consider the hypothesis $H_0: \phi = 1$ against the alternative $H_1: \phi < 1$. Under H_0, the maximum value of the likelihood function is

$$= e^{-\sum_{t=1}^{n}(Y_t - Y_{t-1})} \prod_{t=1}^{n} \left(1 + e^{-\sum_{t=1}^{n}(Y_t-Y_{t-1})}\right)$$

and under the alternative, the maximum value of likelihood function is, $L_1 = e^{-\sum_{t=1}^{n}(Y_t - \hat{\phi}Y_{t-1})} \prod_{t=1}^{n} \left(1 + e^{-\sum_{t=1}^{n}(Y_t-\hat{\phi}Y_{t-1})}\right)$. For $\hat{\phi} \in H_1$, the likelihood ratio test rejects H_0 when $e^{-\sum_{t=1}^{n}(Y_t(1-\hat{\phi})} \prod_{t=1}^{n}\left(\frac{1+e^{-(Y_t-\hat{\phi}Y_{t-1})}}{1+e^{-(Y_t-Y_{t-1})}}\right)$ is small. Wilks (1938) established that under suitable regularity conditions, the distribution of $-2\log\lambda$ is asymptotically Chi-square distribution. The regularity conditions are all verified and hence the decision of unit root in the model can be made by comparing the likelihood ratio test statistic

$$-2\log\lambda = -2\sum_{t=1}^{n}\left[Y_{t-1}(1-\hat{\phi}) + 2\log\left(\frac{1+e^{-(Y_t-\hat{\phi}Y_{t-1})}}{1+e^{-(Y_t-Y_{t-1})}}\right)\right]. \tag{13}$$

with the corresponding Chi-squared table value at a given level of significance.

5. TEST FOR COINTEGRATION IN AN ECM

Modern economic theory often suggests that certain pairs of financial or economic variables should be linked by some long run economic relationship. One of the primary interest concerned with such variables is that to test whether the set of variables are cointegrated. There are several test procedures available for cointegration when the disturbances in vector error correction model are i.i.d Gaussian and some authors have examined the performance of these tests by comparing the sizes and powers of the tests in which the model assumptions are violated. If $\phi \to 1$ in the cointegrating equation, then the series will be a random walk and therefore model cannot explain any long run behaviour in the observed series. Hence it is necessary to assure that the variables are all integrated with same order and are non stationary before we test for the presence of cointegration. Then the idea of testing the presence of unit root in the auto regression equation (2) can be extended to testing the presence of cointegration using a similar approach that considered in Engle and Granger (1987). That is, once the series are identified to be unit root nonstationary with same order of integration, we can extend the test procedure of unit root for testing the presence of cointegration to the residuals of the fitted error correction model. The null hypothesis of unit root $\phi = 1$ can then be identically equal to testing $\delta = 0$ in the error correction model. Note that, unlike the usual cointegration test that applied to the residuals of the cointegrating regression, here we apply the test for the residuals from error correction model. So to test for cointegration, the null hypothesis that has to be taken is no cointegration , $H_0: \delta = 0$ against the alternative hypothesis of $\delta \neq 0$. Once the model parameters are estimated from the data, we tests the residuals from the error correction model using the test procedure described below. If the residuals are stationary, (ie; the null hypothesis of no cointegration is rejected) then we can conclude that the variables will be cointegrated.

Denote the residuals from the error correction model by,

$$\hat{\xi}_{1t} = \Delta x_{1t} - \hat{\beta}\hat{\delta}_1 \hat{z}_{t-1}\ , \quad \hat{\xi}_{2t} = \Delta x_{2t} + \hat{\delta}_1 \hat{z}_{t-1}$$

and the hypothesis of interest is $H_0: \delta = 0$ against $H_1: \delta \neq 0$.

For model (5), under the null hypothesis, the maximum value of the likelihood function is

$$L_0 = e^{-\sum_{t=1}^{n} \Delta x_{1t}} \prod_{t=1}^{n} \left(1 + e^{-\Delta x_{1t}}\right)^{-2}$$

and under the alternative hypothesis, the maximum value of the likelihood function is,

$$L_1 = e^{-\sum_{t=1}^{n} (\Delta x_{1t} - \hat{\beta}\hat{\delta}_1 z_{t-1})} \prod_{t=1}^{n} \left(1 + e^{-(\Delta x_{1t} - \hat{\beta}\hat{\delta}_1 z_{t-1})}\right)^{-2}.$$

For the model (6), under the null, the maximum of the likelihood function is

$L_0 = e^{-\sum_{t=1}^{n} \Delta x_{2t}} \prod_{t=1}^{n} \left(1 + e^{-\Delta x_{2t}}\right)^{-2}$ and under the alternative, maximum of the likelihood function is $L_1 = e^{-\sum_{t=1}^{n} (\Delta x_{2t} + \hat{\delta}_1 \hat{z}_{t-1})} \prod_{t=1}^{n} \left(1 + e^{-(\Delta x_{2t} + \hat{\delta}_1 \hat{z}_{t-1})}\right)^{-2}$.

Hence we reject the null hypothesis of no cointegration if the likelihood ratio tests statistic

$$-2\left[\sum_{t=1}^{n} \hat{\delta}_1 z_{t-1} + 2\log\frac{(1+e^{-(\Delta x_{2t} + \hat{\delta}_1 \hat{z}_{t-1})})}{(1+e^{-(\Delta x_{2t})})}\right] \quad (14)$$

or

$$-2\left[\sum_{t=1}^{n} -\hat{\beta}\hat{\delta}_1 z_{t-1} + 2\log\frac{(1+e^{-(\Delta x_{1t} - \hat{\delta}_1 \hat{\beta} \hat{z}_{t-1})})}{(1+e^{-\Delta x_{1t}})}\right] \quad (15)$$

is too large or too small based on Chi-square critical value. As our study deals with two time series, we have two error correction models that represent the cointegrating relationship. Though both the ECM has a unique representation for the long run cointegrating relationship, which is represented by the term z_t, the null hypothesis of no cointegration will be rejected if at-least one of the above test statistic exceeds the Chi-square critical value.

6. SIMULATION STUDY

As the estimating equations do not admit explicit solutions, we analyse the performance of the above methods by simulation. Hence we carry out a simulation study to understand the performance of the estimator and test statistic described in Sections 4 and 5 for various sample sizes and for different specified values of the model parameters. For the simulation purpose, we first generate the innovation random variable from a logistic distribution. Then for specified values of the model parameter, we simulated the sequence Y_t, t=1,2,...,n using the relation described in (11). Based on this sample, we obtain the maximum likelihood estimates of ϕ using the procedure described in Section 4. We used sample autocorrelation as the initial estimate while solving the log likelihood equations by iterative methods. For the given values of the model parameter, we repeated the experiment 100 times for computing the estimates and then averaged them over the repetitions. Next we compute the likelihood ratio test statistic given in equation (13) for various sample sizes and for different parameter values. Finally we compute the number of rejections in 500 trials for testing the null hypothesis of interest. The numerical computations are carried out for various value of the model parameter and are summarised in Tables 1 and 2.

Table 1. The average estimates and the corresponding root mean squares errors of the MLE

Sample size	True value ϕ	MLE $\hat{\phi}$	RMSE
100	-0.8	-0.7766	0.0706
	-0.5	-0.5006	0.0883
	-0.3	-0.2951	0.0874
	0.2	0.1797	0.0977
	0.3	0.2781	0.0917
	0.6	0.5699	0.0829
	0.8	0.8832	0.0708
300	0.8	0.7921	0.0354
	0.5	-0.5007	0.0456
	-0.3	-0.2874	0.0547
	0.2	0.2006	0.0492
	0.3	0.3008	0.0571
	0.6	0.6007	0.0425
	0.8	0.7944	0.0341
500	-0.8	-0.7949	0.0477
	-0.5	-0.4995	0.0375
	-0.3	-0.2989	0.0446
	0.2	0.1967	0.0394
	0.3	0.3031	0.0404
	0.6	0.5977	0.0357
	0.8	0.7929	0.0271

Table 2. No of rejections in 500 trials of the hypotheis $H_0: \phi = 1$ against $H_1: \phi < 1$ using the test statistic given in (13) for different values of ϕ

Sample size	n=50	n=100	n=250	n= 350
	.01 .05 .1 .2	.01 .05 .1 .2	.01 .05 .1 .2	.01 .05 .1 .2
$\phi = .50$	500 500 500 500	500 500 500 500	500 500 500 500	500 500 500 500
$\phi = .20$	500 500 500 500	500 500 500 500	500 500 500 500	500 500 500 500
$\phi = .80$	218 327 411 460	466 493 499 500	500 500 500 500	500 500 500 500
$\phi = .85$	122 216 312 389	376 448 482 496	500 500 500 500	500 500 500 500
$\phi = .90$	59 133 197 283	197 311 399 464	500 500 500 500	500 500 500 500
$\phi = .95$	56 103 167 273	53 108 167 277	285 388 439 480	482 496 500 500
$\phi = 1.0$	7 14 33 35	8 15 38 45	1 8 24 50	5 12 33 45

Next to evaluate the accuracy of the estimation and testing procedure of the error correction model, a simulation study is carried out for different sample sizes and for different values of the model parameters. For the study, we first generate a sample of size, say T, from the autoregressive equations in (1) and (2) with innovation random variables generated from logistic distribution. For different values of the model parameters, we simulated the time series $\{x_{1t}\}$ and $\{x_{2t}\}$ using the equations (3) and (4). Based on this sample , we generate the error correction model using (5) and (6). Finally we obtained the MLE of the parameters by solving the likelihood equations in (8), (9) and (10). We then repeated the experiment 50 times for computing the estimates and then averaged them over the repetitions. After the parameters of ECM being estimated, we test for cointegration using the residuals from the error correction model. We use the test statistic given in (14) and (15) to compute the number of rejections of the null hypothesis under the various alternatives. In practical situations, we could reject the null hypothesis of no cointegration based on either of the two test statistics. The numerical computations for estimation and testing are carried out for various values of the model parameters and are summarised in Tables 3,4 and 5.

Table 3. The average estimates and the corresponding root mean squared errors of MLE

Sample	True values			MLE		
	δ	β	α	$\hat{\delta}$	$\hat{\beta}$	$\hat{\alpha}$
300	2.6	1.5	1.8	2.708(0.3125)	1.4965(0.0156)	1.8016(0.0049)
	0.5	2	3	0.5080(0.0851)	1.9600(0.2108)	3.0096(0.0258)
	0.3	2.5	3.5	0.3178(0.0683)	2.4966(0.1028)	3.529(0.0547)
	0.2	3	4	0.1863(0.0588)	2.9448(0.1667)	4.027(0.0837)
	0.1	3	4	0.1018(0.0429)	2.7962(0.1550)	4.029(0.1978)
500	2.6	1.5	1.8	2.6777(0.2319)	1.5009(0.0108)	1.8040(0.0275)
	0.5	2	3	0.5133(0.0624)	1.9915(0.0727)	3.035(0.1361)
	0.3	2.5	3.5	0.3107(0.0537)	2.489(0.0507)	3.5116(0.0269)
	0.2	3	4	0.2037(0.0488)	3.0027(0.0643)	4.0179(0.0451)
	0.1	3	4	0.1048(0.0390)	2.9740(0.1498)	4.0560(0.238)
700	2.6	1.5	1.8	2.6651(0.0182)	1.5003(0.0068)	1.8006(0.0019)
	0.5	2	3	0.5075(0.0522)	2.0208(0.0477)	3.0244(0.1113)
	0.3	2.5	3.5	0.3035(0.0391)	2.4976(0.0418)	3.5072(0.0194)
	0.2	3	4	0.1904(0.0379)	2.9789(0.0619)	4.0079(0.0421)
	0.1	3	4	0.1047(0.0260)	3.0078(0.0948)	4.0182(0.0653)

Table 4. No of rejections in 500 trials of the hypothesis $H_0: \phi = 1$ (δ=0) against the alternative of $\phi < 1$ ($\delta \neq 0$) using the first ECM

ECM-1								
Sample size	50				100			
	0.01	0.05	0.1	0.2	0.01	0.05	0.1	0.2
ϕ=0.5	500	500	500	500	500	500	500	500
ϕ=0.8	450	462	472	475	486	489	491	492
ϕ=0.9	413	429	441	452	470	475	480	492
ϕ=.95	220	250	270	285	295	300	325	347

Table 5. No of rejections in 500 trials of the hypothesis $H_0: \phi = 1$ (δ=0) against the alternative of $\phi < 1$ ($\delta \neq 0$) using the second ECM

ECM-2								
Sample size	50				100			
	0.01	0.05	0.1	0.2	0.01	0.05	0.1	0.2
ϕ=0.5	500	500	500	500	500	500	500	500
ϕ=0.8	477	488	489	491	484	490	491	496
ϕ=0.9	453	465	477	483	476	480	484	490
ϕ=.95	259	270	280	280	370	320	340	363

Note that from Table 1 and 3, for series of length 100 and 300, estimates are reasonably satisfactory and become more accurate with increasing sample size. From Table 2, it is evident that as ϕ becomes closer to 1, the number of rejections of the null hypothesis of unit root becomes smaller. For example, in a length of 50 series, the hypothesis $H_0: \phi = 1$ was rejected 7 times at the 0.01 significance level, while it was rejected 218 times when ϕ was 0.8. Hence we claim that the derived test statistic is powerful for testing the presence of unit root in an observed nonstationary time series. From Tables 4 and 5, it is seen that for large values of ϕ or as ϕ increases to 1, the number of rejections of the null hypothesis in 500 trial decreases.

7. BOOTSTRAP METHOD

In this section we address the accuracy of a bootstrap algorithm in small samples for testing the presence of cointegration in an ECM. In recent years, there has been an increasing interest in parametric and non parametric bootstrap inference for econometric and financial time series. The technique of parametric bootstrap suggest estimation of the sampling distribution of the statistic using random sampling methods and it may also be used for constructing tests of hypothesis. Here we provide a simulation based parametric bootstrap method that involves simulating data sets using the maximum likelihood estimates and hence computing the likelihood ratio test statistic for each available simulated data set.

In small sample situations, the asymptotic likelihood ratio test discussed in the earlier sections may not be suitable for determining the cointegrating relationship between two or more time series. The theoretical chi-square distribution for likelihood ratio test will provides much better results if the sample size is reasonably large. Hence for finite sample situations, we can use a parametric bootstrap approach in which we constructs the distribution of the likelihood ratio test statistic empirically. So we provide a Monte Carlo simulation to compare the performance of bootstrap testing with the usual method based on an asymptotic approximation of the distribution of the test statistic. The method involves 4 steps.

As a starting point, we estimate the parameters of the cointegration model using the conditional maximum likelihood estimation method and then obtain the asymptotic likelihood ratio test statistic for the real data. Secondly, we generate a bootstrap sample using the maximum likelihood estimates as the initial values and hence compute the likelihood ratio test statistic for the bootstrap sample. Thirdly, we repeat the above step 10000 times which yield an estimate of the distribution of the likelihood ratio test statistic. Finally, we compute the empirical quantiles of the test statistic and then take the decision on the null hypothesis of no cointegration by comparing the calculated critical values with the calculated likelihood ratio test value.

We analyse a real data set for testing the presence of cointegration by our proposed model in section 8.

8. DATA ANALYSIS

In this section, we illustrate the analysis of cointegration with non normal innovation using the real data set. The data set consists of monthly observations on consumption and export of natural rubber collected from "The Rubber Board", Ministry of Commerce and Industry, Govt. of India, Kottayam. The Figure 1(a) and Figure 1(b) provide the time series plot of the log transformed data and it indicates that the time series is nonstationary.

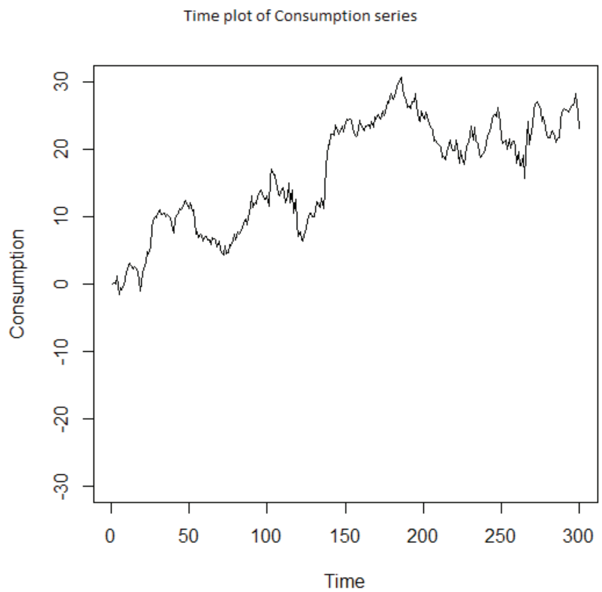

Fig. 1(a). Time plot of consumption series

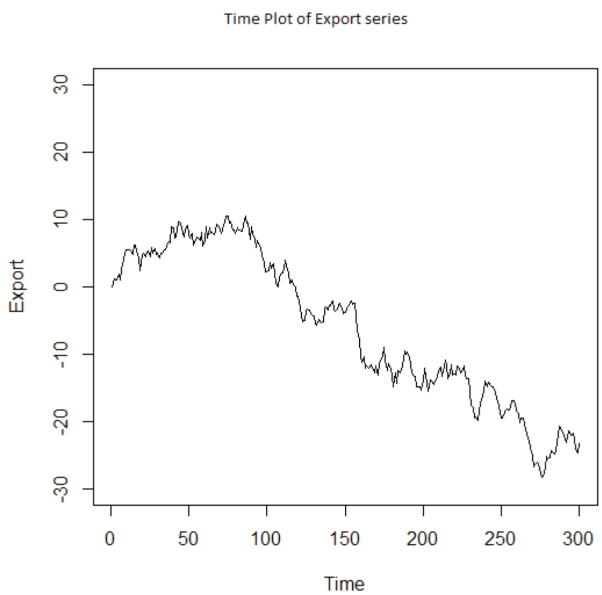

Fig. 1(b). Time plot of Export series

First we tested the data for cointegration with normally distributed errors using the Johansen test for cointegration. Table below shows the Johansen test of cointegration for normally distributed errors. Johansen's trace test tests the null hypothesis of r cointegrating vectors against the alternative hypothesis of n cointegrating vectors. If r=0, it means that there is no relationship among the variables that is stationary.

Table 6. Johansen Trace test

Cointegration rank	test	10%	5%	1%
r=1	4.9	7.52	9.24	12.97
r=0	32.29	17.85	19.96	24.6

The maximum eigen value test tests the null hypothesis of r cointegrating vectors against the alternative of $(r+1)$ cointegrating vectors. From the tables, it can be seen that in both cases the null hypothesis of one cointegrating vector is not rejected. This implies that, cointegration exist between the rubber consumption and export series.

Table 7. Johansen Eigen Value test

Cointegration rank	test	10%	5%	1%
r=1	4.9	7.52	9.24	12.97
r=0	27.39	13.75	15.67	20.2

The parameter estimates are obtained as $\hat{\alpha} = -1.33$, $\hat{\beta} = -0.0059$ and $\hat{\delta} = -0.203$. $X_{1t} - 1.33X_{2t}$ is the estimated cointegrating relationship using the Johansen test. Finally to evaluate the adequacy of the model using normal errors, we checked whether the residual series obtained from the fitted model follows normal distribution. But the assumption of normality is rejected for the residual series, hence we tested for cointegration with errors generated by logistic innovations. Although the plot seems to be nonstationary, it is important to test whether a series is stationary or not before we test for cointegration. Hence we performed a unit root test developed for logistic error variables to the data set in order to test whether the series is stationary or not.

The p values obtained for testing the unit root for consumption and export series are obtained as 0.9928 and 0.9723 respectively. Since both the p values obtained are very large, we do no reject the null hypothesis of unit root and hence the series meets cointegration test condition. Next we perform a maximum likelihood estimation described in the above section in order to find the parameter estimates of an error correction model of order 1. The parameter estimates are obtained as $\hat{\alpha} = -1.43$, $\hat{\beta} = -0.0158$ and $\hat{\delta} = -0.1542$. Thus the estimated cointegrating relationship, if any exist, is $X_{1t} - 1.43X_{2t}$, where $\{X_{1t}\}$ is the month-wise rubber consumption series and $\{X_{2t}\}$ is the month-wise rubber export series. The residuals from the error correction model is obtained as

$$\xi_{1t} = \Delta x_{1t} - \hat{\delta}\hat{\beta}\hat{z}_{t-1}$$

$$\xi_{2t} = \Delta x_{2t} + \hat{\delta}\hat{\beta}\hat{z}_{t-1}, t = 1,2,3, \ldots.$$

Using the parameter estimates of the ECM, we tested whether the residuals follow a logistic distribution using Kolmogorov-Smirnov test. The p values obtained for the series are 0.965 and 0.303 respectively, which indicates that logistic distribution is suitable for the residuals. The probability-probability plot of the residuals are shown in Figure 2(a) and Figure 2(b) also confirms the result.

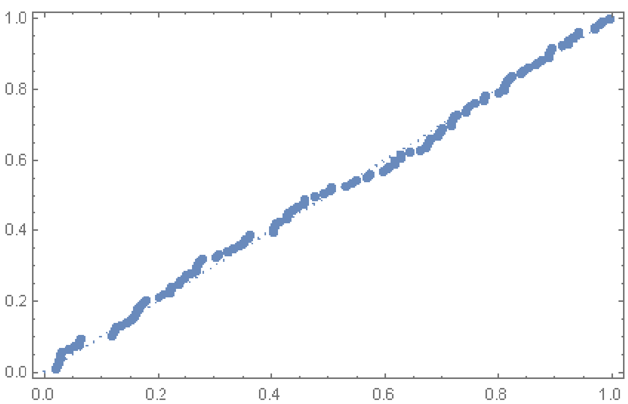

Fig. 2(a). PP- plot of residuals of consumption series

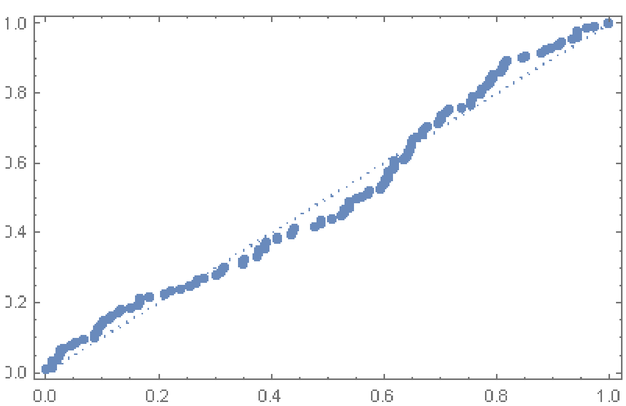

Fig. 2(b). PP- plot of residuals of Export series

Finally we performed a bootstrap algorithm for testing cointegration using the error correction model with logistic errors. The bootstrap values and asymptotic values are given in Table 8.

Table 8. Emperical levels for bootstrap and asymptotic tests

Nominal level	0.2		0.1		0.05	
bootstrap	0.024	2.7197	0.0005	3.849	-0.0957	5.146
Asymptotic	0.016	2.71	0.0039	3.84	0.0098	5.02

The value of the test statistic obtained for the error correction equation of export series is 0.000219 and from Table 8 we can conclude that we reject the null hypothesis of no cointegration at 10 percent level of significance using a bootstrap test and asymptotic test implying that the residuals from the ECM are stationary. If the null hypothesis of no cointegration is rejected, then the cointegrating vector parameter estimate provides an estimate of a long run relationship. That is, $X_{1t} - 1.43X_{2t}$ is the cointegrating relationship and the cointegrated vector is $[1, -1.43]'$.

Thus in both situations, that is with normal and non normal errors, the existence of cointegration relationship in the data is identified. But the residual series obtained from the cointegrating regression using normal errors rejects the assumption of normality. Hence we proceed with the vector autoregression model that allows for logistic innovations to arrive at a right conclusion.

9. CONCLUSIONS

This study provides the evidence of the long-run cointegrating relationship of Rubber consumption and Export series. In this paper, we developed maximum likelihood estimation of the cointegration vector in a first order vector autoregressive model that allows for logistic innovations. Then we developed LRT to detect the presence of cointegration by developing a unit root test for the residuals of the ECM. All the estimating equations are solved by using numerical techniques. From the simulation studies, it is observed that the proposed procedure is powerful for detecting the presence of unit root and cointegration. Along with the usual asymptotic test, a bootstrap test based on MLE is carried out to account for the size distortions caused by the finite samples. The data analysis confirms that the proposed model detects the presence of cointegration.

ACKNOWLEDGEMENT

The authors thank the referees for their insightful comments. Nimitha John wishes to acknowledge the Kerala State Council for Science Technology and Environment under Grant number 1413/2012/

KCSTE for the financial support. The research work of N.Balakrishna is partially supported by the DST SERB under the project No.SR/S4/MS:837/13.

REFERENCES

Box,G.E, & Jenkins, G.M. (1970) Time series analysis: forecasting and control, Holden day.

Dickey, D.A, & Fuller, W.A, (1979) Distribution of the estimators for autoregressive time series with a unit root, *J. Amer. Statist. Assoc.*, **74**, 427-431.

Engle, R.F. & Granger, C.W.J. (1987) Cointegration and error correction: Representation, estimation and testing, *Econometrica*, **306**, 251-276.

Granger, C.W.J. (1981) Some Properties of Time Series Data and their use in Econometric Model Specification, *J. Economet*., 121-130.

Johansen, S.(1988) Statistical analysis of cointegration vectors, *J. Eco. Dyn. Control*, **12**, 231-254.

Kim, K. & Schmidt, P. (1993) Unit root tests with conditional heteroskedasticity, *J. Economet.*, **59**, 287-300.

Lee, T.H. & Tse, Y.(1996) Cointegration tests with conditional heteroskedasticity, *J. Economet.*, **73**, 401-410.

Tiku, M.L., Wong, W.K. & Bian, G.(1999) Estimating parameters in autoregressive models in non-normal situations: Symmetric innovations, *Comm. Statist.-Theory Methods*, **28**, 315-341.

Wilks, S.S.(1938) The large-sample distribution of the likelihood ratio for testing composite hypotheses, *The Annal. Math. Statist.*, **9**, 60-62.

Wong, W.K. & Bian, G.(2005) Estimating parameters in autoregressive models with asymmetric innovations, *Statistics & probability letters*, **71**, 61-70.