ANADOLU ÜNİVERSİTESİ

Bilim ve Teknoloji Dergisi A-Uygulamalı Bilimler ve Mühendislik
Cilt: 14 Sayı: 3 2013
Sayfa: 263-276

ARAŞTIRMA MAKALESİ/**RESEARCH ARTICLE**

Esin FİRUZAN[1], Selim Orhun SUSAM[2]

PARAMETERS' PROPERTIES OF BIVARIATE COINTEGRATED VAR (1) PROCESS

ABSTRACT

Vector autoregressive (VAR) process is common tool for capturing the autocorrelation pattern among VAR models which are generalized form of the univariate autoregression (AR) models. In our study, bivariate cointegrated VAR (1) is considered. Monte Carlo simulation study is performed to examine the finite sample performance of estimators corresponding to the asymptotic distribution for different $\hat{\rho}$ and $\hat{\alpha}$ in MATLAB R2011A software package.

Keywords: Cointegration, Vector autoregressive process, Maximum likelihood estimator, Least square estimator

İKİ DEĞİŞKENLİ EŞBÜTÜNLEŞİK VAR(1) SÜRECİNİN PARAMETRELERİNİN ÖZELLİKLERİ

ÖZ

Tek değişkenli otoregresif sürecin genelleştirilmiş hali olan vektör otoregresif süreç değişkenler arasında otokorelasyon örneklerini modelleyerek yaygınca kullanılan bir süreçtir.Çalışmamızda iki değişkenli birinci dereceden vektör otoregresif süreç göz önüne alınmıştır. Monte Carlo simulasyonu yardımıyla $\hat{\rho}$ ve $\hat{\alpha}$ nın sonlu örneklem tahmin edicilerinin asismptotik özellikleri MATLAB R2011A programı kullanılarak incelenmiştir.

Anahtar Kelimeler: Eşbütünleşme, Vektör otoregresif süreç, En çok olabilirlik tahmin edicisi , En küçük kareler tahmin edicisi

[1], Faculty of Science University of , Dokuz Eylul, Buca, Izmir
Tel: 0232 411 85 57, E-mail: esin.firuzan@deu.edu.tr

[2], Faculty of Science University of ,Dokuz Eylul, Buca, Izmir
Tel: 0232 411 85 63, E-mail: orhun.susam@deu.edu.tr

Recieved: 28 November 2012 **Accepted:** 24 June 2013

1. INTRODUCTION

The aim of the study is to examine asymptotic properties of parameters depends on $\hat{\rho}$ and $\hat{\alpha}$ under the cointegration for bivariate VAR (1) process. The Monte Carlo simulation study is performed to examine the finite sample performance of $\hat{\rho}$ and $\hat{\alpha}$ in relation to the asymptotic distribution for different $\hat{\rho}$ and $\hat{\alpha}$.

2. COINTEGRATION IN VECTOR AUTOREGRESSIVE PROCESS

Vector autoregressive (VAR) process is common tool for capturing the autocorrelation pattern among VAR models which are generalized form of the univariate *autoregression* (AR) models. In this study, we consider nonstationary bivariate VAR (1) as follows:

$$X_t = AX_{t-1} + u_t \qquad t = 1,2,\ldots,n$$

where A is a nonsingular matrix which involves the coefficients for VAR(1) process, $X_t = (X_{1t}, X_{2t})'$, the error process u_t is *iid* $N(0,\Sigma_u)$ with $\Sigma_u > 0$, and the process is initialized at t=0 by $X_{i0} = 0$.

$$A = \begin{bmatrix} \rho & \theta \\ 0 & \alpha \end{bmatrix}, \qquad \theta \neq 0$$

So, the considered nonstationary VAR (1) process can be expressed in two simultaneous equations. It is clear that X_{1t} is related with both X_{1t-1} and X_{2t-1}, however X_{2t} is related with only X_{2t-1} in the model.

$$X_{1t} = \hat{\rho} X_{1t-1} + \hat{\theta} X_{2t-1} + \hat{u}_{1t}$$

$$X_{2t} = 0 X_{1t-1} + \hat{\rho} X_{2t-1} + \hat{u}_{2t}$$

In our study, we are interested in the specific case I (1) and I (0). I (1) represents that stationary process after first differencing. The two-dimensional VAR (1) process $X_t = AX_{t-1} + u_t$ is called cointegrated if $|\Pi| := |A - I_2|$ has no unit roots for ΔX_t or I (1). Π can be written as $\alpha\beta'$ where α is adjustment rate (loading vector) and β is cointegration vector. One unit root and one stationary root are considered in exogenous model. The characteristic roots of coefficient matrix A are

$$|A - zI_2| = 0$$

$$z_1 = 1 \quad \text{and} \quad z_2 = \lambda < 1 \quad , \quad |\lambda| < 1$$

One unit root is derived by solving characteristic roots of coefficient matrix A. The characteristic roots have only one roots, either if $\rho = 1$, $\alpha < 1$ or $\alpha = 1$, $\rho < 1$.

Also if $z_1 = 1$ then, X_{1t} and X_{2t} are I (1), then A has full rank 2.We rewrite coefficient matrix A as in the following:

$$A = P\begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} Q$$

where P is the eigenvectors of A as columns,

$$P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad Q = P^{-1}, \quad |P| = 1$$

we choose determinant $|P| = 1$ for simplicity then, $ad - bc = 1$ so inverse of P is equal to adjoint matrix of P.

$$P^{-1} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Then we multiply A right by Q and left by P,

$$A = \begin{bmatrix} (ad - \lambda bc) & -ab(1-\lambda) \\ cd(1-\lambda) & (-cb + \lambda ad) \end{bmatrix} \tag{1}$$

This representation of A will be used in Error Correction Model.

3. ERROR CORRECTION REPRESENTATION (ECR)

Since $|P| = 1$, the determinant is $|ad - bc| = 1$ and rewrite elements of matrix in equation (1)

$$ad - \lambda bc = ad - \lambda b + \lambda b - \lambda bc = 1 + cb(1-\lambda)$$
$$-cb + \lambda ad = 1 - ad + \lambda ad = 1 - ad(1-\lambda)$$

That is

$$A = \begin{bmatrix} (ad - \lambda bc) & -ab(1-\lambda) \\ cd(1-\lambda) & (-cb + \lambda ad) \end{bmatrix} = \begin{bmatrix} 1 + cb(1-\lambda) & -ab(1-\lambda) \\ cd(1-\lambda) & 1 - ad(1-\lambda) \end{bmatrix}$$

So A can be rewritten in following

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (1-\lambda)\begin{bmatrix} b \\ d \end{bmatrix}\begin{bmatrix} c & -a \end{bmatrix} \tag{2}$$

Replacing 2 equation instead of A matrix in nonstationary VAR(1) model ($X_t = AX_{t-1} + u_t$), then equation 3 is obtained as a stationary VAR(1) model by means of error correction representation.

$$\begin{bmatrix} \Delta X_{1t} \\ \Delta X_{2t} \end{bmatrix} = (1-\lambda)\begin{bmatrix} b \\ d \end{bmatrix}\left(cX_{1t-1} - aX_{2t-1}\right) + u_t$$

$$\begin{bmatrix} \Delta X_{1t} \\ \Delta X_{2t} \end{bmatrix} = \alpha\beta' \begin{bmatrix} X_{1t-1} \\ X_{2t-1} \end{bmatrix} + u_t \tag{3}$$

α is a error correction coefficient,

$$\alpha = \begin{bmatrix} (1-\lambda)b \\ (1-\lambda)d \end{bmatrix}$$

and β is the cointegration matrix are obtained.

$$\beta = \begin{bmatrix} c \\ -a \end{bmatrix}$$

We often write equation (3.2)

$$\Delta X_t = \Pi X_{t-1} + u_t$$

Where

$$\Pi = \alpha\beta'$$

Π is also equal to $(A - I_2)$.

If the variables are cointegrated, then rank of matrix Π is reduced. That is, for the bivariate system, rank of matrix Π is 1. Hence, one of two characteristic root is different from zero and another one is $(1-\lambda)$. If λ=1, then there is no cointegration in the bivariate system for the error correction representation.

As a result, for the bivariate system, if

- Rank($\boldsymbol{\Pi}$)=0,reduced rank and no cointegration relationship in system
- Rank($\boldsymbol{\Pi}$)=1, reduced rank and cointegration relationship in system
- Rank($\boldsymbol{\Pi}$)=2,full rank, $X_t = AX_{t-1} + u_t$ is stationary.

4. ESTIMATION OF BIVARIATE COINTEGRATED VAR (1) PROCESS

Consider cointegrated VAR (1) process is as follows:

$$\Delta X_t = \Pi X_{t-1} + u_t = \alpha\beta' X_{t-1} + u_t \quad t = 1,2,\ldots \tag{4}$$

where Π is (2x2) matrix of rank r=1 (0<r<2), α and β are (2x1) with rank r=1 and u_t is 2 dimensional white noise process with mean zero and variance-covariance matrix Σ_u. Also we suppose that X_t is I (1) process and $\alpha'_\perp \beta_\perp$ is an invertible because of it is real valued scalar. $\beta_\perp$ and $\alpha_\perp$ are orthogonal complements of α and β. If r=0, then ΔX_t is stationary and if r=p=2 then X_t is stationary.

Maximum Likelihood and Least Square estimation of Π, α and β are discussed in this section.Then asymptotic distribution of this related estimator is derived.

Unrestricted LS estimator will be discussed in this section because of the lack of the information of the variance. Using normal equations, unrestricted LS estimator of Π is obtained as follows:

$$\hat{\Pi} = (\textstyle\sum_{t=1}^{T} \Delta X_t X'_{t-1})(\sum_{t=1}^{T} X_{t-1} X'_{t-1})^{-1} \quad (5)$$

we replace $\Pi \mathrm{X}_{t-1} + u_t$ instead of ΔX_t then equation 4.3. is obtained.

$$\hat{\Pi} - \Pi = (\textstyle\sum_{t=1}^{T} u_t X'_{t-1})(\sum_{t=1}^{T} X_{t-1} X'_{t-1})^{-1} \quad (6)$$

If we choose Q (2x2) such that,

$$Q = \begin{bmatrix} \beta' \\ \alpha'_{\perp} \end{bmatrix} \quad , \quad Q^{-1} = [\alpha(\beta'\alpha)^{-1} \quad \beta_{\perp}(\alpha_{\perp}{}'\beta_{\perp})^{-1}]$$

then we multiply from the left with Q and from the right with inverse of Q, it gives us

$$Q(\hat{\Pi} - \Pi)Q^{-1} = Q\left(\sum_{t=1}^{T} u_t X'_{t-1}\right) Q' Q^{-1'} \left(\sum_{t=1}^{T} X_{t-1} X'_{t-1}\right)^{-1} Q^{-1}$$

$$= \left(\sum_{t=1}^{T} v_t z'_{t-1}\right)\left(\sum_{t=1}^{T} z_{t-1} z'_{t-1}\right)^{-1}$$

where $v_t = Qu_t$ and $z_t = QX_t$.

We indicate that the first r=1 components of z_t by $z_t{}^{(1)} = \beta' \mathrm{X}_t$ which is stationary cointegration relationship and $z_t{}^{(2)} = \alpha_{\perp}{}' \mathrm{X}_t$ which is process contains unit root. So, we can write z_t with stationary and nonstationary parts.

That is,

$$Q(\hat{\Pi} - \Pi)Q^{-1}$$

$$= \left[\sum_{t=1}^{T} v_t\, z_{t-1}{}^{(1)\,\prime} \quad \sum_{t=1}^{T} v_t\, z_{t-1}{}^{(2)\,\prime}\right] \begin{bmatrix} \sum_{t=1}^{T} z_{t-1}{}^{(1)}\, z_{t-1}{}^{(1)\,\prime} & \sum_{t=1}^{T} z_{t-1}{}^{(1)}\, z_{t-1}{}^{(2)\,\prime} \\ \sum_{t=1}^{T} z_{t-1}{}^{(2)}\, z_{t-1}{}^{(1)\,\prime} & \sum_{t=1}^{T} z_{t-1}{}^{(2)}\, z_{t-1}{}^{(2)\,\prime} \end{bmatrix}^{-1}$$

Ahn & Reinsel (1990) would be helpful for details in derivation of the asymptotic distribution of $\hat{\Pi} - \Pi$. The information which is given in Lemma 1.1 will use in the other. sections.

Lemma 1:

1- $T^{-1} \sum_{t=1}^{T} z_{t-1}{}^{(1)}\, z_{t-1}{}^{(1)\,\prime} \xrightarrow{P} \Gamma_z^{(1)}$.

where $\Gamma_z^{(1)}$ is the covariance matrix of $z_t{}^{(1)}$

2- $T^{-\frac{1}{2}} vec\left(\sum_{t=1}^{T} v_t\, z_{t-1}{}^{(1)\prime}\right) \xrightarrow{d} N\left(0, \Gamma_z^{(1)} \otimes \Sigma_v\right)$

3- $T^{-1} \sum_{t=1}^{T} v_t\, z_{t-1}{}^{(2)\prime} \xrightarrow{d} \Sigma_v^{1/2} (\int_0^1 W_k dW_k')' \, \Sigma_v^{1/2} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix}$,

Where W_K denotes the standard wiener process $W_K(s)$ of dimension K .

4- $T^{-3/2} \sum_{t=1}^{T} z_{t-1}{}^{(1)}\, z_{t-1}{}^{(2)\prime} \xrightarrow{p} 0.$

5- $T^{-2} \sum_{t=1}^{T} z_{t-1}{}^{(2)}\, z_{t-1}{}^{(2)\prime} \xrightarrow{d} [0 \quad I_{K-r}]\, \Sigma_v^{\frac{1}{2}} \left(\int_0^1 W_k W_k' ds\right) \Sigma_v^{\frac{1}{2}} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix}$

4.1 Limiting Results for The LS Estimator $\widehat{\Pi}$

We consider D matrix where its elements, $T^{1/2}$ and T, are convergence rates.

$$D = \begin{bmatrix} T^{1/2} & 0 \\ 0 & T \end{bmatrix}$$

Then

$$vec[Q(\widehat{\Pi} - \Pi)Q^{-1}D]$$

$$\xrightarrow{d} \begin{bmatrix} N(0, (\Gamma_z^{(1)})^{-1} \otimes \Sigma_v) \\ vec\left\{\Sigma_v^{\frac{1}{2}} (\int_0^1 W_k W_k' ds)' \Sigma_v^{\frac{1}{2}} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix} ([0 \quad I_{K-r}] \Sigma_v^{\frac{1}{2}} (\int_0^1 W_k W_k' ds) \Sigma_v^{\frac{1}{2}} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix})^{-1}\right\} \end{bmatrix}$$

The $vec[Q(\widehat{\Pi} - \Pi)Q^{-1}D]$ is distributed as a combination of normal distribution and wiener process.

Proof:

$$Q(\widehat{\Pi} - \Pi)Q^{-1}D$$

$$= \left[T^{-1/2} \sum_{t=1}^{T} v_t\, z_{t-1}{}^{(1)\prime} \quad T^{-1} \sum_{t=1}^{T} v_t\, z_{t-1}{}^{(2)\prime}\right] D \begin{bmatrix} \sum_{t=1}^{T} z_{t-1}{}^{(1)}\, z_{t-1}{}^{(1)\prime} & \sum_{t=1}^{T} z_{t-1}{}^{(1)}\, z_{t-1}{}^{(2)\prime} \\ \sum_{t=1}^{T} z_{t-1}{}^{(2)}\, z_{t-1}{}^{(1)\prime} & \sum_{t=1}^{T} z_{t-1}{}^{(2)}\, z_{t-1}{}^{(2)\prime} \end{bmatrix}^{-1} D$$

$=$

$$\left[T^{-1/2} \sum_{t=1}^{T} v_t\, z_{t-1}{}^{(1)\prime} \quad T^{-1} \sum_{t=1}^{T} v_t\, z_{t-1}{}^{(2)\prime}\right] \begin{bmatrix} T^{-1} \sum_{t=1}^{T} z_{t-1}{}^{(1)}\, z_{t-1}{}^{(1)\prime} & T^{-1} \sum_{t=1}^{T} z_{t-1}{}^{(1)}\, z_{t-1}{}^{(2)\prime} \\ T^{-2} \sum_{t=1}^{T} z_{t-1}{}^{(2)}\, z_{t-1}{}^{(1)\prime} & T^{-2} \sum_{t=1}^{T} z_{t-1}{}^{(2)}\, z_{t-1}{}^{(2)\prime} \end{bmatrix}^{-1}$$

Using by partitioned inverse;

$$= \left[T^{-1/2}\sum_{t=1}^{T} v_t\, z_{t-1}{}^{(1)\prime} \quad T^{-1}\sum_{t=1}^{T} v_t\, z_{t-1}{}^{(2)\prime}\right]\begin{bmatrix} S_{11}^{-1} + S_{11}^{-1}S_{12}S^{*}S_{21}S_{11}^{-1} & -S_{11}^{-1}S_{12}S^{*} \\ -S^{*}S_{21}S_{11}^{-1} & S^{*} \end{bmatrix}$$

where $S^{*} = (S_{22}^{-1} - S_{21}S_{11}^{-1}S_{12})^{-1}$

By using lemma 1(1)

$$S_{11} = T^{-1}\textstyle\sum_{t=1}^{T} z_{t-1}{}^{(1)}\ z_{t-1}{}^{(1)\prime} \xrightarrow{p} \Gamma_z^{(1)}$$

The S_{11} is converging in probability to stationary process ($z_t{}^{(1)} = \beta' \mathrm{X_t}$) variance-covariance matrix $\Gamma_z^{(1)}$.

By using lemma 1(4)

$$S_{12} = S_{21}' = T^{-1}\textstyle\sum_{t=1}^{T} z_{t-1}{}^{(1)}\ z_{t-1}{}^{(2)\prime} \xrightarrow{d} o_p(T^{1/2})$$

The S_{12} is converging in distribution to zero with converging rate $T^{(-3/2)}$.

By using lemma 1(5) and the continuous mapping theorem;

$$S_{22} = T^{-2}\textstyle\sum_{t=1}^{T} z_{t-1}{}^{(2)}\ z_{t-1}{}^{(2)\prime} = O_p(1)$$

$$S_{22}^{-1} = O_p(1)$$

The inverse of S_{22} convergence to a real-valued scalar ($[0 \quad I_{K-r}]\, \Sigma_v^{\frac{1}{2}}(\int_0^1 W_k W_k' ds)\Sigma_v^{\frac{1}{2}} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix}$) with convergence rate $T^{(-2)}$.

Using rules of partitioned inverse;

$$S^{*} = S_{22}^{-1} + S_{22}^{-1}S_{21}(S_{11} - S_{12}S_{22}^{-1}S_{21})^{-1}S_{12}S_{22}^{-1}$$

$$= O_p(1) + O_p(1)O_p\left(T^{-\frac{1}{2}}\right)O_p(1)o_p\left(T^{\frac{1}{2}}\right)O_p(1)$$

$$= O_p(1)$$

Since $o_p\left(T^{\frac{1}{2}}\right)$which S_{12} is divided by $\left(T^{\frac{1}{2}}\right)$,convergences to zero, S^{*} convergences to S_{22}^{-1}.

It can be seen easily, $S_{11} - S_{12}S_{22}^{-1}S_{21}$ convergences to a scalar.

$$S_{11} - S_{12}S_{22}^{-1}S_{21} = S_{11} - o_p\left(T^{\frac{1}{2}}\right)O_p(1)o_p\left(T^{\frac{1}{2}}\right) = S_{11} + o_p(1) = O_p(1)$$

Based on continuous mapping theorem, the inverse of $S_{11} - S_{12}S_{22}^{-1}S_{21}$ also convergences to the scalar.

$$(S_{11} - S_{12}S_{22}^{-1}S_{21})^{-1} = O_p(1)$$

As a result,

$$S_{11}^{-1} + S_{11}^{-1} S_{12} S^{*} S_{21} S_{11}^{-1} = \left(\Gamma_z^{(1)}\right)^{-1} + O_p(1) o_p\left(T^{\frac{1}{2}}\right) O_p(1) o_p\left(T^{\frac{1}{2}}\right) O_p(1)$$

$$= (\Gamma_z^{(1)})^{-1} + o_p(1)$$

and

$$-S_{11}^{-1} S_{12} S^{*} = -O_p(1) o_p\left(T^{\frac{1}{2}}\right) O_p(1) = o_p(1)$$

Thus,

$$= \left[T^{-\frac{1}{2}} \sum_{t=1}^{T} v_t \, z_{t-1}{}^{(1)\,\prime} \quad T^{-1} \sum_{t=1}^{T} v_t \, z_{t-1}{}^{(2)\,\prime}\right]$$

$$\times \begin{bmatrix} (T^{-1} \sum_{t=1}^{T} z_{t-1}{}^{(1)} \, z_{t-1}{}^{(1)\,\prime})^{-1} + o_p(1) & o_p(1) \\ o_p(1) & (T^{-2} \sum_{t=1}^{T} z_{t-1}{}^{(2)} \, z_{t-1}{}^{(2)\,\prime})^{-1} + o_p(1) \end{bmatrix}$$

=

$$\left[T^{-\frac{1}{2}} \sum_{t=1}^{T} v_t \, z_{t-1}{}^{(1)\,\prime} (T^{-1} \sum_{t=1}^{T} z_{t-1}{}^{(1)} \, z_{t-1}{}^{(1)\,\prime})^{-1} \quad T^{-1} \sum_{t=1}^{T} v_t \, z_{t-1}{}^{(2)\,\prime} (T^{-2} \sum_{t=1}^{T} z_{t-1}{}^{(2)} \, z_{t-1}{}^{(2)\,\prime})^{-1}\right]$$
$$+ o_p(1)$$

Finally,

$$vec\left[Q\left(\hat{\Pi} - \Pi\right) Q^{-1} D\right]$$

$$= \begin{bmatrix} vec(T^{-\frac{1}{2}} \sum_{t=1}^{T} v_t \, z_{t-1}{}^{(1)\,\prime} (T^{-1} \sum_{t=1}^{T} z_{t-1}{}^{(1)} \, z_{t-1}{}^{(1)\,\prime})^{-1}) \\ vec(T^{-1} \sum_{t=1}^{T} v_t \, z_{t-1}{}^{(2)\,\prime} \left(T^{-2} \sum_{t=1}^{T} z_{t-1}{}^{(2)} \, z_{t-1}{}^{(2)\,\prime}\right)^{-1}) \end{bmatrix}$$

Using lemma 1(1), lemma 1(2) and lemma 1(5), the proof has been completed.

$$\xrightarrow{d} \begin{bmatrix} N(0, (\Gamma_z^{(1)})^{-1} \otimes \Sigma_v \\ vec\left\{\Sigma_v^{\frac{1}{2}} (\int_0^1 W_k W_k' ds)' \Sigma_v^{\frac{1}{2}} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix} ([0 \quad I_{K-r}] \Sigma_v^{\frac{1}{2}} (\int_0^1 W_k W_k' ds) \Sigma_v^{\frac{1}{2}} \begin{bmatrix} 0 \\ I_{K-r} \end{bmatrix})^{-1}\right\} \end{bmatrix}$$

The $vec\left[Q\left(\widehat{\Pi}-\Pi\right)Q^{-1}D\right]$ is still consisting of nonnormal elements. Choosing proper convergence rate, the nonnormal part of matrix could be normal.

The distribution of unrestricted LSE estimator $\widehat{\Pi}$ is asymptotically normal,

$$\sqrt{T}vec(\widehat{\Pi}-\Pi)\xrightarrow{d}N(0,\beta\left(\Gamma_z^{(1)}\right)^{-1}\beta'\otimes\Sigma_u)$$

And $\beta\left(\Gamma_z^{(1)}\right)^{-1}\beta'$ is estimated by using $(T^{-1}\sum_{t=1}^{T}X_{t-1}X_{t-1}')^{-1}$

4.2. Limiting Results for the MLE Estimator $\widehat{\Pi}$

When the error process is assumed to be Normal distribution, maximum likelihood estimator can be used to estimate unknown parameters. If α and $\sum_u$ are known, the maximum likelihood estimator is the same as Generalized Least Sqaure (GLS) estimator for $\hat{\beta}_{K-r}'$. The log likelihood function is given as following:

$$\ln(l)=-\frac{KT}{2}ln2\pi-\frac{T}{2}ln|\Sigma_u|-\frac{1}{2}\sum_{t=1}^{T}(\Delta y_t-\Pi X_{t-1})'\Sigma_u^{-1}(\Delta X_t-\Pi X_{t-1})$$

For maximizing log-likelihood function the following determinant should be minimized.

$$\left|T^{-1}\sum_{t=1}^{T}(\Delta y_t-\Pi X_{t-1})(\Delta y_t-\Pi X_{t-1})'\right|$$

For the general case, rank (Π) =r, it means that there are r cointegration relationship. We can write $\Pi=\alpha\beta'$, so the determinant is given by

$$\left|T^{-1}\sum_{t=1}^{T}(\Delta X_t-\alpha\beta' X_{t-1})(\Delta X_t-\alpha\beta' X_{t-1})'\right|$$

with respect to α and β. The minimum value of the determinant is attained for

$$\tilde{\beta}=[v_1 \quad ,\dots, \quad v_r]'(\sum_{t=1}^{T}X_{t-1}X_{t-1}')^{-1/2}$$

$$\tilde{\alpha}=\left(\sum_{t=1}^{T}\Delta X_tX_{t-1}'\tilde{\beta}\right)\left(\sum_{t=1}^{T}\tilde{\beta}'X_{t-1}X_{t-1}'\tilde{\beta}\right)^{-1}.$$

Where the eigenvalues $\lambda_1,\geq\lambda_2\geq\cdots\geq\lambda_K$ and the associated orthonormal eigenvectors $v_1 \quad ,\dots, \quad v_r$
is obtained from the following matrix

$$\left(\sum_{t=1}^{T}X_tX_{t-1}'\right)^{-1/2}\left(\sum_{t=1}^{T}X_{t-1}\Delta X_t'\right)\left(\sum_{t=1}^{T}X_tX_t'\right)\left(\sum_{t=1}^{T}\Delta X_tX_{t-1}'\right)\left(\sum_{t=1}^{T}X_tX_{t-1}'\right)^{-1/2}$$

And also $\breve{\Pi} = \breve{\alpha}\breve{\beta}'$ must have same asymptotic results as the unrestricted LS estimator of Π. We know that $\hat{\beta}'$ does not affect the LS estimator Π. And also, MLE estimator of α is equal to LS estimator (Lutkepohl 2005). That is given in the following asymptotic results,

$$\sqrt{T}vec(\breve{\alpha}\breve{\beta}' - \Pi) \xrightarrow{d} N\left(0, \beta \left(\Gamma_z^{(1)}\right)^{-1} \beta' \otimes \Sigma_u\right)$$

To reach unique $\hat{\beta}'$, normalized MLE estimator of . β should be obtained. $\breve{\beta} = \begin{bmatrix} I_r \\ \breve{\beta}_{K-r} \end{bmatrix}$ is normalized MLE estimator β and also the normalized estimator for MLE estimator $\breve{\alpha}$ can be obtained explicitly. $\breve{\beta}$ and $\breve{\alpha}$ estimators are given below:

$$\breve{\alpha} = \left(\sum_{t=1}^{T} \Delta X_t X_{t-1}' \breve{\beta}\right)\left(\sum_{t=1}^{T} \breve{\beta}' X_{t-1} X_{t-1}' \breve{\beta}\right)^{-1}$$

$$\breve{\beta}_{K-r}' = \left(\breve{\alpha}' \breve{\Sigma}_u^{-1} \breve{\alpha}\right)^{-1} \breve{\alpha}' \breve{\Sigma}_u^{-1} \left(\sum_{t=1}^{T} \left(\Delta y_t - \breve{\alpha} X_{t-1}^{(1)}\right) X_{t-1}^{(2)\prime}\right)\left(\left(\sum_{t=1}^{T} X_{t-1}^{(2)} X_{t-1}^{(2)\prime}\right)^{-1}\right)$$

MLE estimators of $\breve{\Pi}$, $\breve{\alpha}$ and $\breve{\beta}$ have same asymptotic properties as LS estimators of $\hat{\Pi}$,$\hat{\alpha}$ and $\hat{\beta}$. So, asymptotic properties are identical for both estimation techniques.

5. SIMULATION STUDY

In this section, finite sample properties of both estimator is considered through Monte Carlo simulation. Cointegrated bivariate model $X_t = AX_{t-1} + u_t$ is simulated with following coefficient matrix,

$$A = \begin{bmatrix} \rho & \theta \\ 0 & \alpha \end{bmatrix},$$

and variance covariance matrix of *iid* error process

$$\Sigma_u = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Simulation is performed for different ρ and α values in A matrix. . One unit root is derived by solving characteristic roots of coefficient matrix A. Characteristic roots have only one root, either if $\rho = 1$, $\alpha < 1$ or $\alpha = 1$, $\rho < 1$ .We assume cointegrated process contains one unit root.

The aim of the study is to examine the asymptotic properties of $E(\hat{\rho}) - \rho$ and $E(\hat{\alpha}) - \alpha$ firstly for constant ρ and varying α., secondly for constant α and varying ρ. α and ρ should not be greater than 1, because we consider one unit root and one stationary root in the bivariate system. In both steps, θ is the same because its value doesn't affect the stationarity of the system.

Then $E(\hat{\rho}) - \rho$ and $E(\hat{\alpha}) - \alpha$ are performed for different replications T=50, 100, 250 through Monte Carlo simulation.

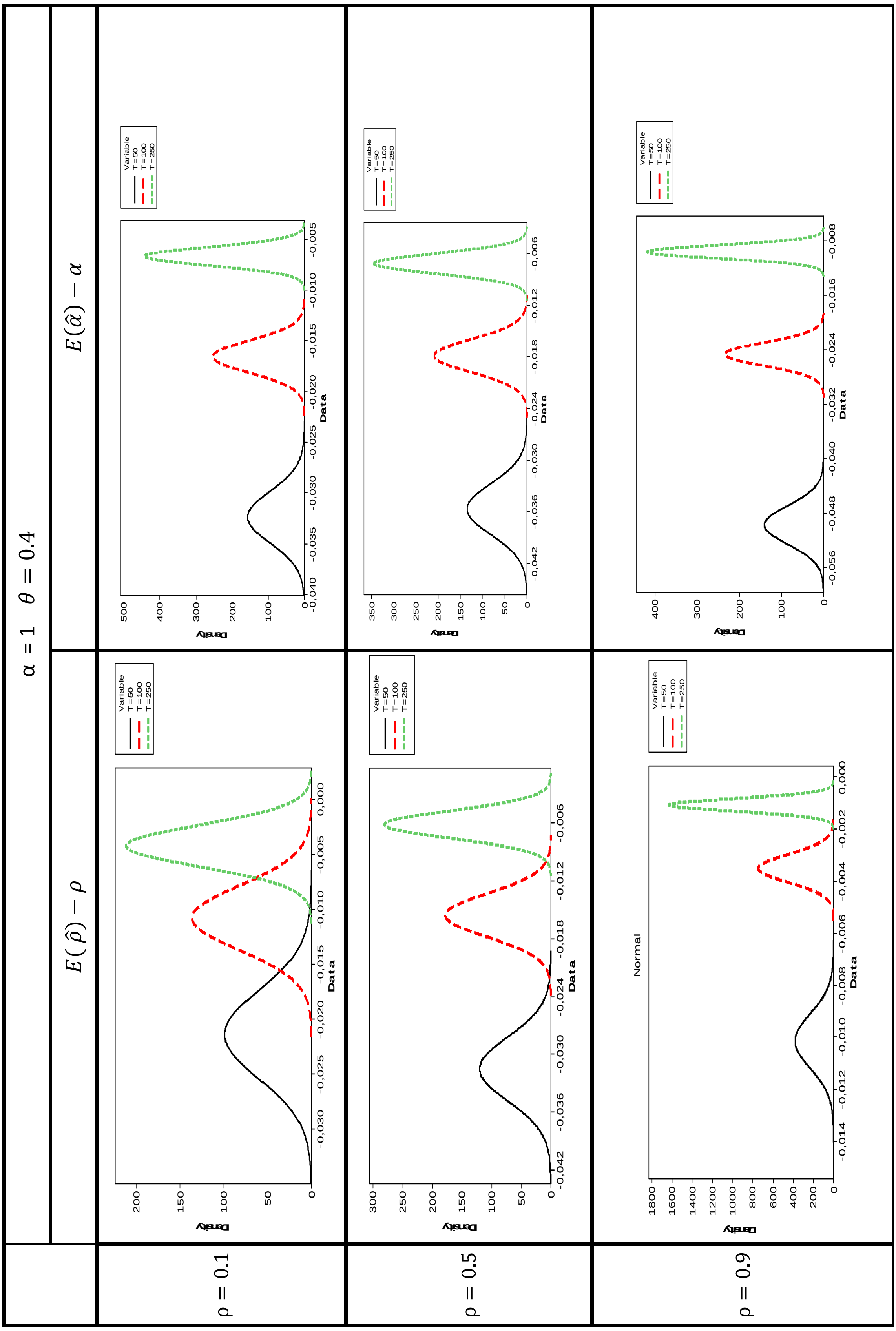

Figure 1. Histograms of $E(\hat{\rho}) - \rho$ and $E(\hat{\alpha}) - \alpha$ for $\alpha = 1$ and ρ =0.1 , 0.5 , 0.9 ; $\theta = 0.4$

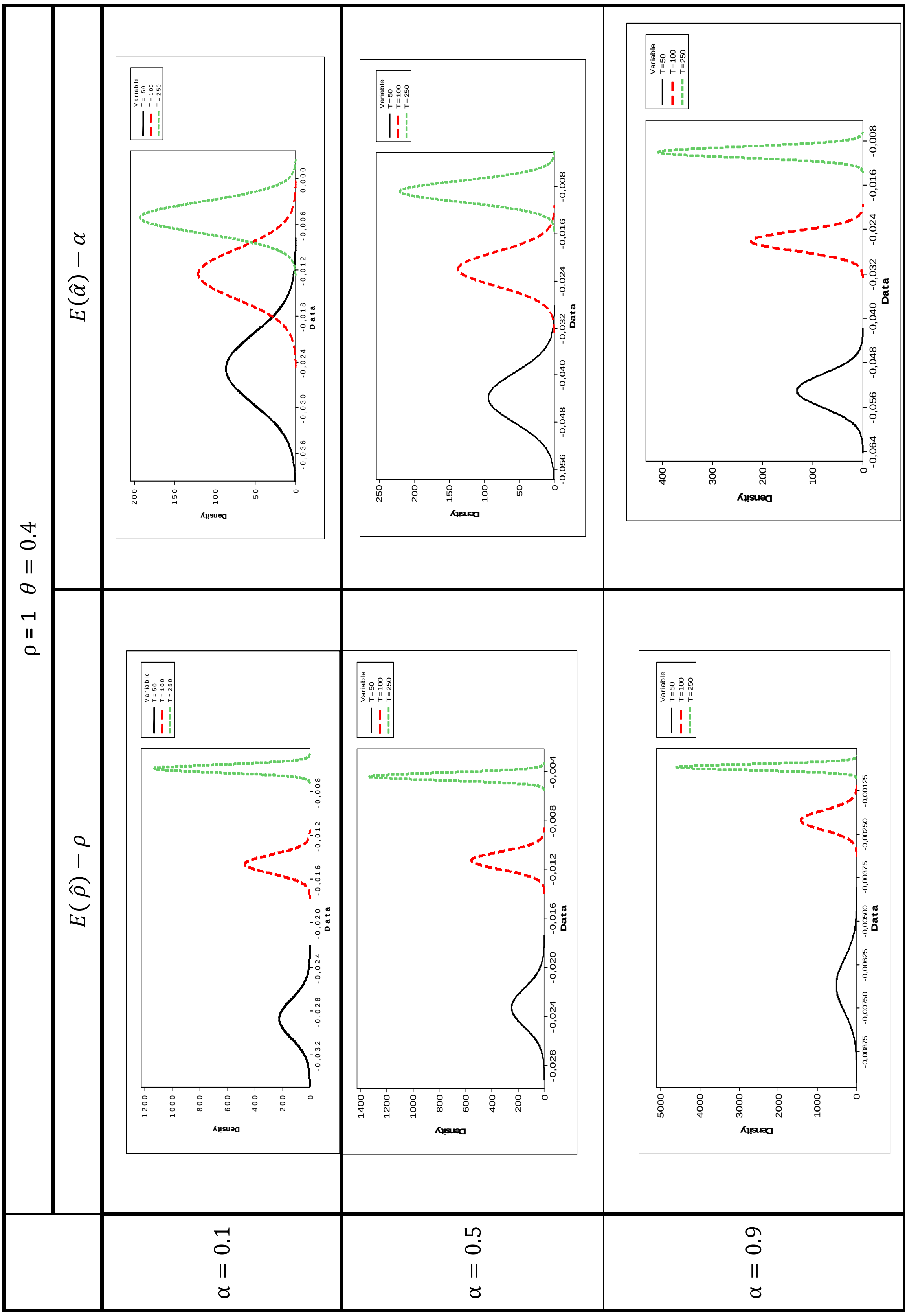

Figure 1. Histograms of $E(\hat{\rho}) - \rho$ and $E(\hat{\alpha}) - \alpha$ for $\alpha = 1$ and ρ =0.1 , 0.5 , 0.9 ; $\theta = 0$.

When histograms which are illustrated in Figure 1 are examined- $\alpha = 1$ and ρ =0.1 , 0.5 , 0.9- distributions of $E(\hat{\rho}) - \rho$ have smaller kurtosis as T increases for any case in ρ. Also, when $\alpha = 1$ is constant and ρ increases, kurtosis and bias of distribution of $E(\hat{\rho}) - \rho$ have better result, that is smaller kurtosis and narrower confidence interval ,and smaller bias for all ρ. Unlike $E(\hat{\rho}) - \rho$, when α is equal to 1 and ρ is increasing, histograms of $E(\hat{\alpha}) - \alpha$ have the almost same results for all ρ .

As shown in Figure 2, for $\rho = 1$, α =0.1, 0.5 ,0.9, $\theta = 0.4$, kurtosis of distributions of $E(\hat{\rho}) - \rho$ is decreasing considerably in contrast to kurtosis of distribution $E(\hat{\alpha}) - \alpha$ . .Also, properties of distribution $E(\hat{\alpha}) - \alpha$ is almost same as the time series length 100,250.

	α	1		
	ρ	**0.1**	**0.5**	**0.9**
T=50	***MSE*** $\hat{\alpha}$	0.0011	0.0013	0.0025
	MSE $\hat{\rho}$	0.000474	0.1358	0.6238
T=100	***MSE*** $\hat{\alpha}$	0.000277	0.00032	0.00060
	MSE $\hat{\rho}$	0.000125	0.1478	0.6344
T=250	***MSE*** $\hat{\alpha}$	0.000045	0.000052	0.000092
	MSE $\hat{\rho}$	0.000021	0.1551	0.6338

Table 1. Mean Square Error of Parameters when $\alpha = 1$

As it is shown in Table 1; for all cases, as time series length increases, mean square errors (MSE) of $\hat{\alpha}$ and $\hat{\rho}$ parameters decreases. When ρ approaches to one, MSE of parameter $\hat{\rho}$ increases remarkably comparing to α. Reversely, when α approaches to 1, this increasing rate of MSE of $\hat{\alpha}$ and $\hat{\rho}$ parameters is slower than ρ approaches to 1 as shown in Table 2.

	ρ	1		
	α	**0.1**	**0.5**	**0.9**
T=50	***MSE*** $\hat{\alpha}$	0.000642	0.001938	0.002818
	MSE $\hat{\rho}$	0.000830	0.000545	0.000047
T=100	***MSE*** $\hat{\alpha}$	0.000167	0.000492	0.000683
	MSE $\hat{\rho}$	0.000215	0.000127	0.000004
T=250	***MSE*** $\hat{\alpha}$	0.000030	0.000081	0.000102
	MSE $\hat{\rho}$	0.000035	0.000019	0.000000

Table 2. Mean Square Error of Parameters when $\rho = 1$

6. CONCLUSION

When α =1, the distributions of $E(\hat{\rho}) - \rho$ have smaller kurtosis as T increases for any case in ρ. Also, when $\alpha = 1$ is constant and ρ increases, kurtosis and bias of distribution of $E(\hat{\rho}) - \rho$ have better result, that is smaller kurtosis and narrower confidence interval ,and smaller bias for all ρ. Unlike $E(\hat{\rho}) - \rho$, when α is equal to 1 and ρ is increasing, histograms of $E(\hat{\alpha}) - \alpha$ have the almost same results for all ρ .

For $\rho = 1$, α =0.1, 0.5 ,0.9, $\theta = 0.4$, kurtosis of distributions of $E(\hat{\rho}) - \rho$ is decreasing considerably in contrast to kurtosis of distribution $E(\hat{\alpha}) - \alpha$ . .Also, properties of distribution $E(\hat{\alpha}) - \alpha$ is almost same as the time series length 100,250.

For all cases, as time series length increases, mean square errors (MSE) of $\hat{\alpha}$ and $\hat{\rho}$ parameters decreases. When ρ approaches to one, MSE of parameter $\hat{\rho}$ increases remarkably comparing to α. Reversely, when α approaches to 1, this increasing rate of MSE of $\hat{\alpha}$ and $\hat{\rho}$ parameters is slower than ρ approaches to 1.

When ρ has unit root, the MSE of parameters have better results. In existence of exogenous variables in the bivariate system, unit root case should be taken account of parameter ρ=1. Unbias and consistency results are obtained in this case.

REFERENCES

Akdi, Y (2010). *Zaman Serileri Analizi (Birim Kökler ve Kointegrasyon).* Gazi Kitabevi. Ankara

Cho, S., Hong, H., Ahn, S. (2010). Inference of Cointegrated Model with Exogenous Variables. *SIRFE Working Papers* 10-A04

Johansen, S (1988). Statistical Analysis of Cointegration Vectors. *Journal of Economic Dynamics and Control* , 231,254.

Johansen, S (1996). *Likelihood-Based Inference in Cointegrated Vector-Autoregressive Models.* Oxford University Press, Oxford.,

Lütkepohl, H (2006). *New Introduction to Multiple Time Series Analysis*. Springer, Berlin.

Seo , B (2004). Estimation and Inference in The Cointegrated System with Stationary Covariates. *Journal of the Korean Statistical Society* , 345-366.

Wei, W (2006). *Time Series Analysis: Univariate and Multivariate Methods.* Pearson, United States.