ÇOK DEĞİŞKENLİ SETAR MODELİ İLE TÜRKİYE'DE DOLAR VE ALTIN FİYATLARINA DAİR BİR UYGULAMA

Dr. Ümran M. KAHRAMAN
Necmettin Erbakan Üniversitesi
uteksen@gmail.com

Öznur AYDINER
Kırklareli Üniversitesi
ozaydiner@hotmail.com

Özet

Kendinden uyarımlı eşiksel otoregresif (SETAR [Self-exciting threshold autoregressive]) model, doğrusal olmayan zaman serisi modellerinden biridir. Model, bir zaman serisinin kendi geçmiş değerlerinden etkilenerek farklı rejimlerde farklı doğrusal otoregresif süreçlere sahip olmasını ifade etmektedir. Tsay (1998), çalışmasında tek değişkenli kendinden uyarımlı eşiksel otoregresif süreci çok değişkenli yapı için genişletmiştir.

Bu çalışmada, çok değişkenli kendinden uyarımlı eşiksel otoregresif model uygulaması için TL cinsinden günlük Dolar (USD) kuru ve altın fiyatları serisi kullanılmıştır. Altın fiyatları serisi gösterge değişken olarak alınıp çok değişkenli SETAR model oluşturulmuş ve modelin performansını değerlendirmek üzere modelden öngörüler elde edilmiştir

Yapılan çalışmada, altın fiyatlarının gösterge değişken olarak alındığı çok değişkenli Dolar ve altın fiyatları modelinden elde edilen öngörüler serilerin gözlenen değerleri ile yakın bir seyir izlemektedir. Buna göre kurulan modelin öngörü yapmak için uygun olduğu söylenebilir. Elde edilen çok değişkenli SETAR modele göre, Türkiye piyasasında altın ve Dolar fiyatlarının birbirini etkilediği ve birlikte modellenebileceği sonucuna varılmıştır.

Anahtar Kelimeler: Çok değişkenli SETAR model, Eşiksel lineer olmama testi, Türkiye'de altın ve Döviz fiyatları

JEL Kodu: C32

AN APPLICATION OF THE DOLLAR AND GOLD PRICES IN TURKEY WITH MULTIVARIABLE SETAR MODEL

Abstract

Self-exciting threshold autoregressive (SETAR) model is one of the non-linear time series models. The model represents that a time series which is influenced by its own past values, has different regimes in different linear autoregressive processes. Tsay (1998) extends the univariate self-exciting threshold autoregressive process for multivariate structure in his study. In this study, daily exchange rate of dollar (USD) and gold prices series in TL are used for multivariate self-exciting threshold autoregressive model application. Gold prices series has been taken as indicator variable and multivariate SETAR model has been created. Then, predictions have been obtained from the model to evaluate performance of the model. Accordingly, the model is said to be suitable to make predictions. According to this obtained multivariate SETAR model, the prices of gold and dollar affect each other in Turkey market and they can be modelled together.

Keywords: Multivariate SETAR model, test of threshold nonlinearity, gold and currency prices in Turkey
JEL Classification: C32

1. Introduction

Threshold autoregressive model (TAR) is one of the nonlinear time series model. Threshold autoregressive models are firstly discussed by Tong (1978) and Tong and Lim (1980). Then, Tong (1990) was explained the self-exciting threshold autoregressive model (SETAR) more broadly. Original source of the model is limited loops and circular structure of time series. Also, limited asymmetric loops can be modelled (Tong, 1990).

Purpose of this study is to choose the structural parameters of SETAR model by using the Tsay's (1989; 1998) method which provides an easy application of determining the model process. To determine the threshold value from structural parameters, a test statistic based on some prediction residuals and a threshold linearity test are applied. Also, graphical tools are used for possible threshold number and values. By using these statistics, SETAR model will be created.

In this study, a multivariate SETAR model was obtained using daily gold prices in TL and data of Dollars (USD) exchange in free market and covering the period 03.01.2005-30.12.2011. Codes were created in MATLAB 7.7.0(R2008b) program for numerical calculations.

2. Multivariate Threshold Autoregressive Model

Tsay (1998) extended the univariate threshold autoregressive process for multivariate structure in his study. Let a time series with s-dimensions $\boldsymbol{X}_t = (X_{1t}, X_{2t}, \dots, X_{st})'$ is taken. $\boldsymbol{X}_t$ is considered as a vector autoregressive process. A multivariate SETAR model with k regimes is defined as below.

$$\boldsymbol{X}_t = \boldsymbol{C}_0^{(\mathrm{i})} + \sum_{j=1}^{p_\mathrm{i}} \boldsymbol{\phi}_j^{(\mathrm{i})} \boldsymbol{X}_{t-\mathrm{i}} + \boldsymbol{\varepsilon}_t^{(\mathrm{i})}, \qquad r_{\mathrm{i}-1} < Z_{t-d} \le r_\mathrm{i} \tag{1}$$

Where $\boldsymbol{C}_0^{(i)}$ are $(s \times 1)$-dimensional constant vectors and $\boldsymbol{\phi}_j^{(i)}$ are $(s \times s)$-dimensional parameter matrices for $i = 1, \dots, k$. $\boldsymbol{\varepsilon}_t^{(i)}$ vectors at i. regime satisfy the $\boldsymbol{\varepsilon}_t^{(i)} = \sum_i^{1/2} \boldsymbol{a}_t$ equality. $\boldsymbol{\Sigma}_i^{1/2}$'s are positively defined symmetric matrices and $\{\boldsymbol{a}_t\}$, is a sequence of serially uncorrelated random vectors with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{I}$, the $(s \times s)$-dimensional identity matrix. The threshold value Z_{t-d} is assumed to be stationary and this variable depends on past observations of $\boldsymbol{X}_{t-d}$. For example;

$$Z_{t-d} = \kappa' \boldsymbol{X}_{t-d} \tag{2}$$

an arrangement can be made as above. κ, is a vector $(s \times 1)$-dimensional. If κ is chosen as $\kappa = (1,0, \dots ,0)'$, threshold value becomes $Z_{t-d} = X_{1,t-d}$. If $\kappa = (\frac{1}{s}, \frac{1}{s}, \dots, \frac{1}{s})'$, the threshold value will be average of all of the elements in $\boldsymbol{X}_{t-d}$ (Chan, Wong, & Tong, 2004).

2.1. Multivariate Nonlinearity Test

The purpose of the multivariate SETAR model is to test the threshold nonlinearity of $\boldsymbol{X}_t$, under the assumption of known variables p and d for the given observations of $\{\boldsymbol{X}_t, Z_{t-d}\}$, $t = 1, \dots, n$. Null hypothesis establishes as the linear form of $\boldsymbol{X}_t$. If it is alternative hypothesis, $\boldsymbol{X}_t$ is threshold which means it is nonlinear. For this, a regression application is created by using least squares method.

$$\boldsymbol{X}'_t = \widetilde{\boldsymbol{X}}'_t \boldsymbol{\Phi} + \boldsymbol{\varepsilon}'_\mathrm{t}, \quad t = h+1, \dots, n \tag{3}$$

Here, $h\ is$ $\max(p, d)$ and $\widetilde{\boldsymbol{X}}_t = (1, X'_{t-1}, \dots, X'_{t-p})'$ are $(ps+1)$-dimensional regressors. $\boldsymbol{\Phi}$, indicates the parameter matrices. If null hypothesis is true then the predictions of least squares will be useful in equality (3). However, if alternative hypothesis is true, predictions will be biased. In this case, the residuals which are acquired by sequentially revision of equation (3), will inform. Z_{t-d} threshold variable is out of the values $S = \{Z_{h+1-d}, \dots, Z_{n-d}\}$ for equality (3). Let $Z_{(i)}$ the smallest *i th* element of *S*. $t(i)$ indicates the time index of $Z_{(i)}$ and the linear regression which is the increasing sequence of threshold value Z_{t-d}, can be written as in the equation (4).

$$\boldsymbol{X}'_{t(i)+d} = \widetilde{\boldsymbol{X}}'_{t(i)+d} \boldsymbol{\Phi} + \boldsymbol{\varepsilon}'_{\mathrm{t(i)+d}}, \quad i = 1, \dots, n-h \tag{4}$$

Dynamic structure of $\boldsymbol{X}_t$ series does not change as seen in equation (4). Because for each t correspond to $\boldsymbol{X}_t$, $\widetilde{\boldsymbol{X}}_t$ regressors do not change. Only the order of the data entered into the regression changes. Tsay (1998), develops a method which based on the prediction residuals of recursive least squares method to determine model alteration in equality (4). If the structure of series $\boldsymbol{X}_t$ is linear, recursive least squares predictions of sequantial autoregression will be consistent and prediction residuals have white noise process. In this case, prediction residuals are expected to be uncorrelated with $\widetilde{\boldsymbol{X}}_{t(i)+d}$ regressors. However if $\boldsymbol{X}_t$ series has a threshold model, prediction residuals cannot have white noise because the predictions of least squares are biased. In this situation, the prediction residuals are related to $\widetilde{\boldsymbol{X}}_{t(i)+d}$ regressors.

Let $\widehat{\boldsymbol{\Phi}}_b$, be the least squares estimates of $\boldsymbol{\Phi}$ for $i = 1, \dots, b$ in equation (4). It means that it is the sequential regression estimates which is obtained by the smallest data number with b observations of Z_{t-d}. Then,

$$\hat{\boldsymbol{\varepsilon}}_{t(b+1)+d} = \boldsymbol{X}_{t(b+1)+d} - \widehat{\boldsymbol{\Phi}}'_{\mathrm{b}}\widetilde{\boldsymbol{X}}_{t(b+1)+d} \quad (5)$$

$$\hat{\boldsymbol{e}}_{t(b+1)+d} = \hat{\boldsymbol{\varepsilon}}_{t(b+1)+d}/[1 + \widetilde{\boldsymbol{X}}'_{t(b+1)+d}\boldsymbol{V}_b\widetilde{\boldsymbol{X}}_{t(b+1)+d}]^{1/2} \quad (6)$$

Hereby, $\boldsymbol{V}_b = [\sum_{i=1}^{b}\widetilde{\boldsymbol{X}}_{t(i)+d}\widetilde{\boldsymbol{X}}'_{t(i)+d}]^{-1}$ and (5) and (6) shows prediction residuals of regression and standardized prediction residuals in equality (4). These values can be obtained from recursive least squares algorithm. After that,

$$\hat{\boldsymbol{e}}'_{t(l)+d} = \widetilde{\boldsymbol{X}}'_{t(l)+d}\boldsymbol{\Psi} + \mathbf{w}'_{\mathrm{t(l)+d}}\,,\ l = b_0 + 1, \dots, n - h \quad (7)$$

regression is handled. b_0 indicates the beginning observation number of iterative regression. b_0 is recommended to be between $3\sqrt{n}$ and $5\sqrt{n}$ (Chan, Wong, & Tong, 2004). The regression problem is in equation (7) is to test the hypothesis H_0: $\boldsymbol{\Psi} = 0$ against the alternative hypothesis H_1:$\boldsymbol{\Psi} \neq 0$. The test statistic prepared by Tsay (1998) for testing this hypothesis, is in the equation (8).

$$C(d) = [n - h - b_0 - (sp + 1)] \times \{\ln[\det(\boldsymbol{S}_0)] - \ln[\det(\boldsymbol{S}_1)]\} \quad (8)$$

Here, d indicates the delay of threshold variable Z_{t-d} and $\boldsymbol{S}_0 = \frac{1}{n-h-b_0}\sum_{l=b_0+1}^{n-h}\hat{\boldsymbol{e}}'_{t(l)+d}\hat{\boldsymbol{e}}_{t(l)+d}$ and $\boldsymbol{S}_1 = \frac{1}{n-h-b_0}\sum_{l=b_0+1}^{n-h}\widehat{\boldsymbol{w}}'_{t(l)+d}\widehat{\boldsymbol{w}}_{t(l)+d}$ are given. $\widehat{\boldsymbol{w}}'_{\boldsymbol{t}}$ are the residuals obtained from least squares in equality (7). $\boldsymbol{X}_t$ is linear under the null hypothesis. $C(d)$ has a Chi-square distribution with degree of freedom $s(ps+1)$ (Tsay, 1998).

2.2. Determining of Model Parameters

To create the $C(d)$ statistic in equation (8) p and d values must be known. Choice of d depends on p so firstly, determining p problem will be focused. To determine degree of autoregressive, values of partial autocorrelation were used in univariate time series. For multivariate systems, partial autocorrelation matrices (PAM) will be used and p value will be chose. Tiao and Box (1981) say that if data is suitable for l order vector autoregressive process, PAM structure in the l delay is final coefficient matrices. Partial autoregression matrice of a vector AR(p) process is zero for $l > p$. Elements and standard errors of partial autoregression matrices are obtained by applying the classical multivariate least squares method to autoregressive model for $l = 1,2,\dots$ (Tiao & Box, 1981).

$\phi'_1, \dots, \phi'_p$ predictions are asymptotically normal distributed at a stationary AR(p) model. The elements of partial autoregression matrices can be shown with + and – to summarize usefully. For example; when a coefficient in PAM is greater than 2 times own standard error or smaller than -2 times, it is included in matrix with sign + and -. Values in between is shown with a dot.

Also, to determine the order of autoregressive model, likelihood ratio statistic can be used. . $\boldsymbol{\phi}_l = \boldsymbol{0}$ null hypothesis versus $\boldsymbol{\phi}_l \neq \boldsymbol{0}$ alternative hypothesis is tested by equation (9).

$$\boldsymbol{S}(l) = \sum_{t=l+1}^{n}(\boldsymbol{X}_t - \widehat{\boldsymbol{\phi}}_1\boldsymbol{X}_{t-1} - \cdots - \widehat{\boldsymbol{\phi}}_p\boldsymbol{X}_{t-p}) \times (\boldsymbol{X}_t - \widehat{\boldsymbol{\phi}}_1\boldsymbol{X}_{t-1} - \cdots - \widehat{\boldsymbol{\phi}}_p\boldsymbol{X}_{t-p})' \qquad (9)$$

When $\boldsymbol{S}(l)$ is the matrix of residual sum of squares and cross products after fitting an AR(l), likelihood ration statistic can be defined as in (10).

$$U = |\boldsymbol{S}(l)|/|\boldsymbol{S}(l-1)| \qquad (10)$$

If the statistic $M(l) = -(N - \frac{1}{2} - ls)\ln(U)$, is defined for U, the statistic is asymptotically distributed as χ^2 with s^2 degrees of freedom (Bartlett, 1938). In this case, with the presence constant term $N = n - p - 1$ is the effective number of observations (Tiao & Box, 1981).

After determining p order, delay parameter given the greatest $C(d)$ statistic and provides $d \leq p$ condition, is stated.

2.3. Prediction

If p, d, k and $\mathcal{R}_k = \{r_1, \dots, r_{k-1}\}$ are known, multivariate autoregression can divided into regimes in equation (4). For the j th regime of data a general linear model can written as in (11).

$$\boldsymbol{X}_j = \boldsymbol{A}_j \boldsymbol{\phi}^{(j)} - \boldsymbol{\varepsilon}_j \tag{11}$$

In this case

$$\boldsymbol{X}_j = \left(\boldsymbol{X'}_{(\pi_{j-1}+1)+d}, \boldsymbol{X'}_{(\pi_{j-1}+2)+d}, \dots, \boldsymbol{X'}_{(\pi_j)+d}\right)' \tag{12}$$

$$\boldsymbol{\phi}^{(j)} = (\boldsymbol{c}'_0, \boldsymbol{\phi}_1'^{(j)}, \dots, \boldsymbol{\phi}_p'^{(j)}) \tag{13}$$

$$\boldsymbol{\varepsilon}_j = \left(\boldsymbol{\varepsilon'}_{(\pi_{j-1}+1)+d}, \boldsymbol{\varepsilon'}_{(\pi_{j-1}+2)+d}, \dots, \boldsymbol{\varepsilon'}_{(\pi_j)+d}\right)' \tag{14}$$

$$\boldsymbol{A}_j = \begin{pmatrix} 1 & \boldsymbol{X'}_{(\pi_{j-1}+1)+d-1} & \cdots & \boldsymbol{X'}_{(\pi_{j-1}+1)} & \cdots & \boldsymbol{X'}_{(\pi_{j-1}+1)+d-p} \\ 1 & \boldsymbol{X'}_{(\pi_{j-1}+2)+d-1} & \cdots & \boldsymbol{X'}_{(\pi_{j-1}+2)} & \cdots & \boldsymbol{X'}_{(\pi_{j-1}+2)+d-p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & \boldsymbol{X'}_{(\pi_j)+d-1} & \cdots & \boldsymbol{X'}_{(\pi_j)} & \cdots & \boldsymbol{X'}_{(\pi_j)+d-p} \end{pmatrix} \tag{15}$$

are defined. Hereby, π_j is the greatest value of j such that $\{r_{j-1} < Z_{(j)} \leq r_j\}$ is given for $j = 1, \dots, k-1$. $\pi_0 = 0$ and $\pi_k = n - p$ are defined. The number of observation at j th regime is $n_j = \pi_j - \pi_{j-1}$. Least squares prediction of $\boldsymbol{\phi}^{(j)}$can obtain from classical multivariate least squares method. It means

$$\widehat{\boldsymbol{\phi}}^{(j)} = \left(\boldsymbol{A'}_j \boldsymbol{A}_j\right)^{-1} (\boldsymbol{A'}_j \boldsymbol{X}_j) \tag{16}$$

and for j th regime of residuals variance-covariance matrices

$$\widehat{\boldsymbol{\Sigma}}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \{\widehat{\boldsymbol{\varepsilon}}_{(\pi_{j-1}+t)+d} \widehat{\boldsymbol{\varepsilon}}'_{(\pi_{j-1}+t)+d}\} \tag{17}$$

can be written (Chan, Wong, & Tong, 2004). AIC is defined as in (18) for the model in equality (1) (Tsay, 1998).

$$\text{AIC}(p, d, k, \mathcal{R}_k)=\sum_{j=1}^{k}\{n_j \ln\left|\widehat{\boldsymbol{\Sigma}}_j\right| + 2s(sp + 1)\} \quad (18)$$

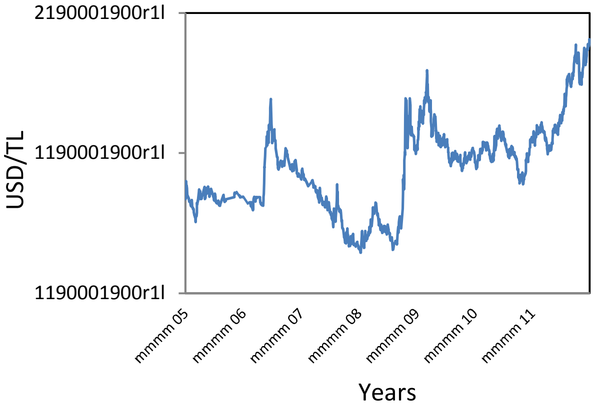
The most important problem is to determine the number of regime (k) as well as Z_t threshold variable in multivariate SETAR model. If p and d values are known, k and $\mathcal{R}_k$ parameters can search through minimizing AIC. Tsay (1998) considers that regime number is chosen 2 or 3 for convenience in calculation. Also, he suggests that according to the different percentiles of Z_{t-d} , data divided into subgroups. Applying the test statistic to these subgroups in equation (8) it can be seen whether a changing appears in the models at the subgroups. Finally, to correct the order of AR in each regime ($p_k \leq p$), AIC can be used (Chan, Wong, & Tong, 2004).

2.4. The Adequacy of the Model

Tiao and Box (1981) propose using partial autoregression matrices and likelihood ratio statistic to examine the residuals in multivariate SETAR model. In this way whether the residuals contain any model can be determined.

3. An application

For the implementation of multivariate threshold autoregressive model, daily Dollar (USD) rate (X_{1t}) and gold prices (X_{2t}) in TL series were used. Time series consisted of 1311 daily observations between the date 03.05.2005-30.12.2011 (Kahraman, 2012). Exchange rate was obtained from http://evds.tcmb.gov.tr/ (2011). The graphs given the series versus time, is shown in Figure 1.



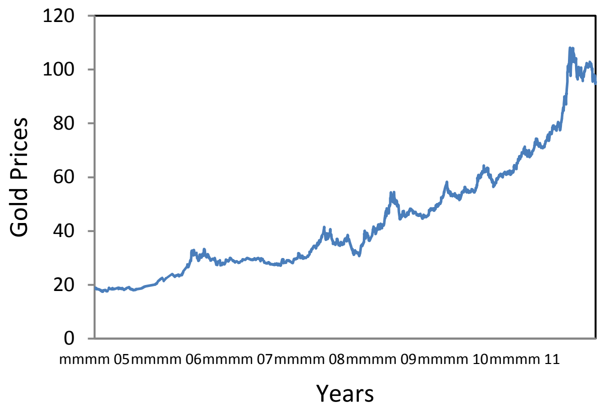


Figure 1. USD and gold TL prices

Unit root research was made for series and the result of ADF test statistic are given in Table 1. According to this, both of the series include the unit root. Return series were calculated with $\boldsymbol{R}_{\mathrm{t}} = (R_{1t}, R_{2t})'$ for both of them.

Table 1. Results of Unit Root Tests

	USD	**Altın**
ADF	***p***	***p***
Cut and trend	0.4589	0.5807
Cut and without trend	0.6123	0.3155
Without trend and cut	0.8378	0.9978

In this case, R_{1t} indicates the return values of USD series and R_{2t}indicates the return values of gold series.

$$R_{1t} = 100 * [\ln(X_{1t}) - \ln(X_{1(t-1)})$$
$$R_{2t} = 100 * [\ln(X_{2t}) - \ln(X_{2(t-1)})]$$

After series were converted, unit root structure was removed (Table 2).

Table 2. The result of ADF unit root test for return series

	USD	**Altın**
ADF	***p***	***p***
Cut and trend	0.0001	0.0001
Cut and without trend	0.0001	0.0001
Without trend and cut	0.0001	0.0001

The graphs of return series are given with Figure 2. In the study, R_{2t} was used as threshold indicator variable for multivariate SETAR model so $R_{2t} = Z_t$.

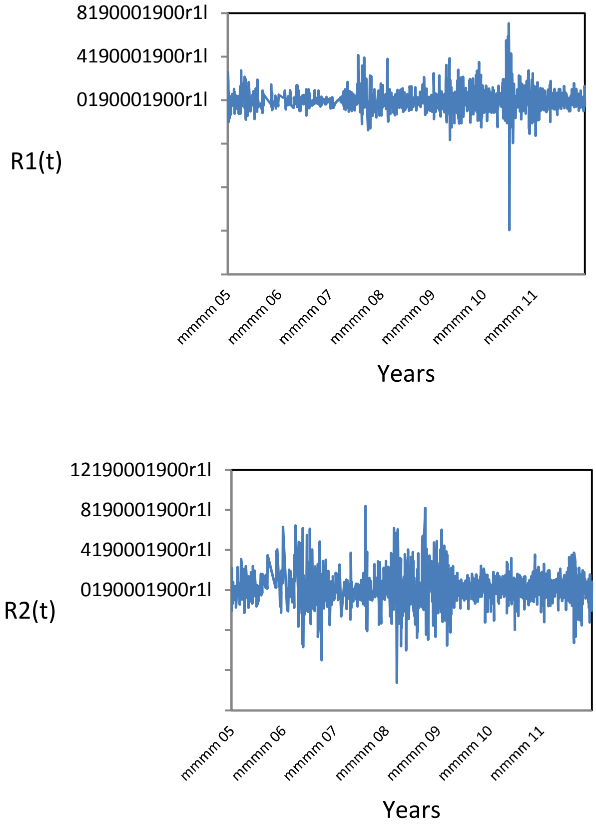


Figure 2. Return series versus years

Firstly, to determine the autoregressive degree of vector time series partial autoregression matrices (PAM) were created. Structure of PAM and likelihood ratio statistic is shown in Table

3. Also, to determine the model degree of vector autoregressive time series, AIC, SIC and likelihood ratio statistic were calculated for first 50 delay. Only the likelihood ratio statistic gave a result in determining the order of autoregression result and the other two criteria did not determine any order. Values of AIC, SIC and likelihood ratio statistic are in Table 4.

Table 3. $\boldsymbol{R}_t$ vector autoregressive series regarding PAM structures and LR statistic

Delay 1-6 $\begin{pmatrix} \cdot & \cdot \\ \cdot & - \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & + \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$

6.696 4.472 1.544 1.627 11.360 5.109

Delay 7-12 $\begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & - \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix} \begin{pmatrix} \cdot & \cdot \\ \cdot & \cdot \end{pmatrix}$

8.902 4.235 5.198 6.019 2.269 3.134

Table 4. AIC, SIC and likelihood ratio statistic values of vector time series

Delay	LR	AIC	SC	Delay	LR	AIC	SC
1	6.2629	6.7017	6.7262	26	0.2737	6.7719	7.2042
2	5.0529	6.7040	6.7448	27	3.7581	6.7751	7.2237
3	1.2872	6.7093	6.7664	28	1.9951	6.7798	7.2447
4	1.2729	6.7147	6.7881	29	2.2359	6.7843	7.2655
5	11.5321	6.7118	6.8015	30	2.2306	6.7888	7.2863
6	4.8488	6.7143	6.8203	31	1.0223	6.7943	7.3082
7	8.1614	6.7141	6.8364	32	3.3366	6.7978	7.3280
8	3.7799	6.7174	6.8560	33	4.2403	6.8006	7.3471
9	5.4669	6.7193	6.8743	34	4.9215	6.8028	7.3657
10	6.0320	6.7208	6.8921	35	1.3698	6.8080	7.3872
11	1.8717	6.7256	6.9132	36	2.4508	6.8123	7.4078
12	3.1364	6.7294	6.9334	37	1.7347	6.8172	7.4290
13	8.7154	6.7287	6.9489	38	3.1987	6.8209	7.4489
14	2.7331	6.7328	6.9694	39	1.6559	6.8258	7.4702
15	7.8556	6.7328	6.9857	40	3.6198	6.8291	7.4898
16	4.4116	6.7356	7.0047	41	1.1401	6.8345	7.5115
17	3.3968	6.7391	7.0246	42	1.4600	6.8396	7.5329
18	8.3994	6.7386	7.0404	43	2.5122	6.8438	7.5534
19	5.0241	6.7408	7.0590	44	1.7939	6.8486	7.5746
20	0.6547	6.7467	7.0811	45	1.0025	6.8541	7.5964
21	0.7458	6.7524	7.1031	46	4.8455	6.8563	7.6149
22	1.9903	6.7571	7.1242	47	1.5726	6.8613	7.6362
23	8.1776	6.7567	7.1401	48	3.9629	6.8642	7.6554
24	2.6994	6.7608	7.1605	49	3.7748	6.8673	7.6749
25	1.7237	6.7658	7.1818	50	5.3581	6.8690	7.6929

The likelihood ratio statistic for comparison of the table value is $\chi_4^2 = 9.49$. As shown in Table 3, $p = 5$ and $C(d)$ was calculated. In this case, values of d can be $d = 1,2,\dots,5$. To calculate $C(d)$ statistic, different d values and different beginning observation number values m were used and the performance of statistic was compared. The values of $C(d)$ statistic are shown in Table 5.

$C(d)$ statistic was compared with the table value $\chi_{22}^2 = 33.92$ and (p,d)=(5,2) value which gave the greatest value of $C(d)$ statistic with $m = 100$ beginning observations was used to research threshold value.

To determine the regime number and threshold values, AIC criteria was used. Firstly, the model which had two regimes and single threshold value, was studied for $k = 2$.

In this case, threshold value varied in the intervals $r \in [Q_{10}(Z_{t-d}), Q_{90}(Z_{t-d})]$. AIC values were obtained as Figure 3.

Table 5. Values of $C(d)$ statistic

m	p	d	$C(d)$
100	5	1	54.4518
		2	96.9948
		3	60.3639
		4	76.6433
		5	77.9754
125	5	1	34.3138
		2	38.1897
		3	39.0557
		4	29.9877
		5	47.7874
150	5	1	7.1343
		2	4.0483
		3	9.7988
		4	6.7465
		5	9.5608
175	5	1	7.2559
		2	3.6887
		3	10.6245
		4	7.2878
		5	9.1101

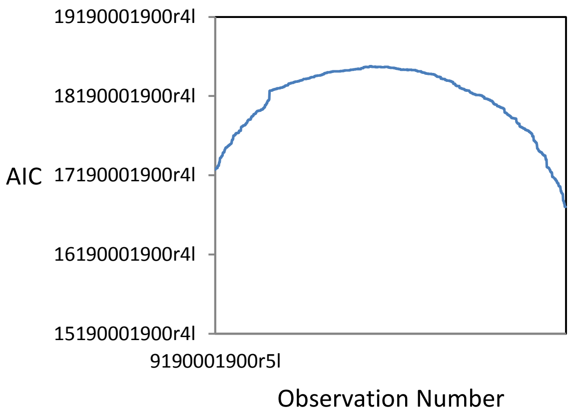


Figure 5. The change of AIC value for the case of two regimes

If the model with two regimes is discussed, AIC values decreased as the number of observation increased in Figure 6. After that, in the discussion of the model with three regimes and two threshold values, it was found that AIC values were calculated. Threshold values varied in the intervals $r_1 \in [Q_{10}(Z_{t-d}), Q_{45}(Z_{t-d})]$ and $r_2 \in [Q_{55}(Z_{t-d}), Q_{90}(Z_{t-d})]$.

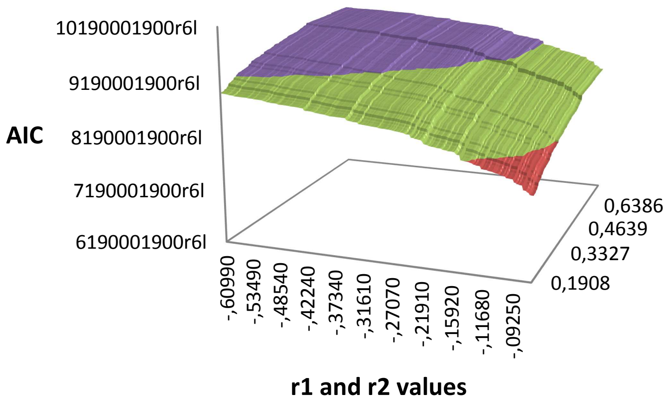


Figure 6. According to two threshold values AIC variation

When Figure 6 is examined, the threshold values can be decided according to the change of AIC. Seeing that the observations were clustered in the first regime for the data with two regimes, it was thought that it would be much better to divide them into three. According to this, $r_1 = -0.0722$ and $r_2 = 0.1908$ were chosen. Multivariate SETAR model was created like below.

$$\boldsymbol{R}_t = \begin{cases} \boldsymbol{C}_0^{(1)} + \sum_{i=1}^{5} \boldsymbol{\phi}_i^{(1)} \boldsymbol{R}_{t-i} + \boldsymbol{\varepsilon}_t^{(1)}, & Z_{t-2} \leq -0.0722 \\ \boldsymbol{C}_0^{(2)} + \sum_{i=1}^{5} \boldsymbol{\phi}_i^{(2)} \boldsymbol{R}_{t-i} + \boldsymbol{\varepsilon}_t^{(2)}, & -0.0722 < Z_{t-2} \leq 0.1908 \\ \boldsymbol{C}_0^{(3)} + \sum_{i=1}^{5} \boldsymbol{\phi}_i^{(3)} \boldsymbol{R}_{t-i} + \boldsymbol{\varepsilon}_t^{(3)}, & Z_{t-2} > 0.1908 \end{cases}$$

The optimal delay number of each regime was determined with respect to AIC criteria and the least squares predictions of parameters are given in Table 6. The numbers which are in brackets, show standard deviation of coefficients.

Table 6. The least squares predictions for three regimes model

$\widehat{\boldsymbol{C}}_0$	$\widehat{\boldsymbol{\phi}}_1$	$\widehat{\boldsymbol{\phi}}_2$	$\widehat{\boldsymbol{\phi}}_3$	$\widehat{\boldsymbol{\phi}}_4$	$\widehat{\boldsymbol{\phi}}_5$
1. rejim					
$\begin{pmatrix} -0.075 \\ (0.061) \\ 0.196 \\ (0.092) \end{pmatrix}$	$\begin{pmatrix} 0.107 & 0.011 \\ (0.046) & (0.029) \\ -0.055 & -0.092 \\ (0.070) & (0.043) \end{pmatrix}$	$\begin{pmatrix} 0.023 & 0.030 \\ (0.047) & (0.070) \\ -0.069 & 0.039 \\ (0.040) & (0.060) \end{pmatrix}$	$\begin{pmatrix} -0.089 & -0.001 \\ (0.047) & (0.028) \\ 0.031 & -0.032 \\ (0.070) & (0.043) \end{pmatrix}$	$\begin{pmatrix} -0.090 & -0.018 \\ (0.045) & (0.028) \\ -0.004 & 0.006 \\ (0.067) & (0.042) \end{pmatrix}$	$\begin{pmatrix} -0.028 & -0.009 \\ (0.039) & (0.027) \\ 0.058 & 0.130 \\ (0.059) & (0.040) \end{pmatrix}$
2. rejim					
$\begin{pmatrix} 0.019 \\ (0.120) \\ 0.293 \\ (0.128) \end{pmatrix}$	$\begin{pmatrix} -0.030 & -0.053 \\ (0.077) & (0.095) \\ -0.013 & -0.370 \\ (0.082) & (0.101) \end{pmatrix}$	$\begin{pmatrix} -0.019 & 0.940 \\ (0.107) & (1.290) \\ -0.100 & -0.017 \\ 0.114 & (1.375) \end{pmatrix}$	$\begin{pmatrix} 0.063 & -0.070 \\ (0.070) & (0.094) \\ 0.070 & -0.279 \\ (0.075) & (0.100) \end{pmatrix}$	$\begin{pmatrix} 0.307 & 0.051 \\ (0.101) & (0.083) \\ -0.207 & 0.069 \\ (0.108) & 0.089 \end{pmatrix}$	$\begin{pmatrix} 0.100 & -0.128 \\ (0.097) & (0.098) \\ 0.097 & 0.035 \\ (0.103) & (0.105) \end{pmatrix}$
3. rejim					
$\begin{pmatrix} 0.049 \\ (0.056) \\ 0.076 \\ (0.100) \end{pmatrix}$	$\begin{pmatrix} -0.025 & 0.021 \\ (0.039) & (0.022) \\ -0.034 & -0.011 \\ (0.070) & (0.039) \end{pmatrix}$	$\begin{pmatrix} 0.059 & -0.047 \\ (0.034) & (0.031) \\ -0.032 & -0.012 \\ (0.061) & (0.055) \end{pmatrix}$	$\begin{pmatrix} 0.010 & 0.008 \\ (0.040) & (0.022) \\ -0.116 & -0.001 \\ (0.071) & (0.004) \end{pmatrix}$	$\begin{pmatrix} 0.014 & 0.018 \\ (0.036) & (0.023) \\ 0.101 & -0.053 \\ (0.064) & (0.040) \end{pmatrix}$	$\begin{pmatrix} -0.111 & 0.004 \\ (0.042) & (0.023) \\ 0.002 & 0.040 \\ (0.076) & (0.041) \end{pmatrix}$

As variance-covariance matrices of each regime are shown in Table 7, AIC value of model was obtained as 159.972.

Table 7. Variance-covariance matrices of regimes and observation numbers

n	$\widehat{\boldsymbol{\Sigma}}$
585	$\begin{pmatrix} 1.142 & \\ 0.003 & 2.595 \end{pmatrix}$
131	$\begin{pmatrix} 1.362 & \\ -0.002 & 1.547 \end{pmatrix}$
589	$\begin{pmatrix} 0.813 & \\ -0.010 & 2.590 \end{pmatrix}$

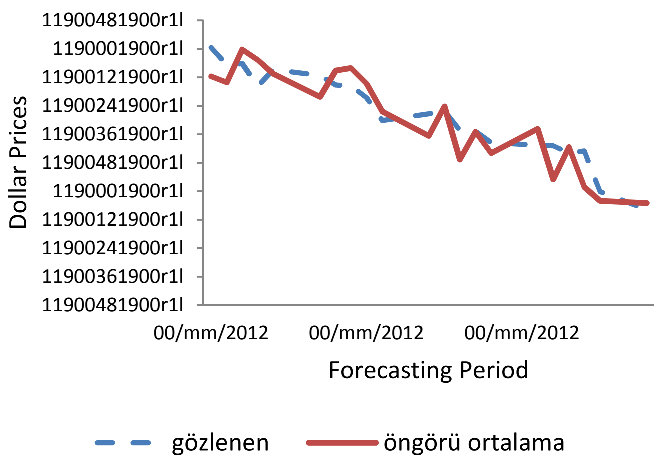
To examine the residuals of model, PAM structure of residuals and results of likelihood ratio test are given in Table 8.

Table 8. PAM structures of residuals and LR statistic

Delay 1-6	$\begin{pmatrix} . & . \\ . & . \end{pmatrix}$	$\begin{pmatrix} . & . \\ . & . \end{pmatrix}$	$\begin{pmatrix} . & . \\ . & . \end{pmatrix}$	$\begin{pmatrix} . & + \\ . & . \end{pmatrix}$	$\begin{pmatrix} . & . \\ . & . \end{pmatrix}$	$\begin{pmatrix} - & . \\ . & . \end{pmatrix}$
	2.624	0.76	3.873	7.494	1.406	7.838
Delay 7-12	$\begin{pmatrix} . & . \\ . & . \end{pmatrix}$	$\begin{pmatrix} . & . \\ . & - \end{pmatrix}$	$\begin{pmatrix} . & . \\ . & . \end{pmatrix}$	$\begin{pmatrix} . & . \\ - & . \end{pmatrix}$	$\begin{pmatrix} . & . \\ . & . \end{pmatrix}$	$\begin{pmatrix} . & . \\ . & . \end{pmatrix}$
	1.863	3.577	4.302	7.421	5.936	1.199

It can be said that residuals do not show any model structures ($\chi_4^2 = 9.49$).

Parametric bootstrap method was applied to predict from obtained multivariate SETAR model. New series were created by using residual terms of model and $B = 1000$. After that, two series were obtained by substituting into the model. Then, point predictions were made by using average of predictions. Results are in Figure 7.



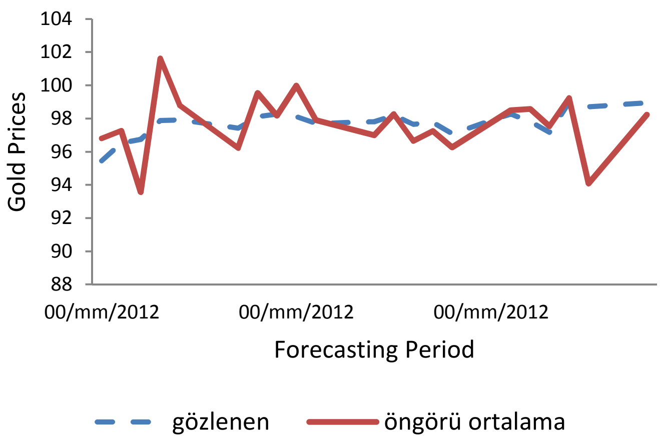


Figure 7. Prediction for gold and Dollar prices

Predictions, which were made according to model of multivariate Dollar and gold prices model in which gold prices were used as indicator value, were close to observation values of series, so it can be said that the created model is suitable to predict. According to obtained multivariate SETAR model, the results are that gold prices and Dollar prices affect each other in Turkey market and can be modelled together. In the condition, in which return of gold prices was close to zero ($-0.0722 < Z_{t-2} \leq 0.1908$), series was subject to multivariate AR(5) process. When return of gold prices was less than zero or much greater than zero, it was found that different multivariate AR(5) processes arose.

4. Results

Threshold autoregressive models which have partial linear structure stand out with a very wide range application. Threshold autoregressive model is useful especially in return series of data in the area of economic and financial due to the cyclic data structure. Due to structure of threshold, if series is not return series, it will be useful. In this study, the researchers attempted to establish multivariate SETAR model. Considering the study of Tsay (1998), the statistic which test the multivariate nonlinearity, was calculated and nonlinearity hypothesis could not be rejected for the data of daily gold and Dollar prices which are thought to be related to each other. The first order difference was required for stagnation in both series. Gold prices were taken as basis for multivariate SETAR model and according to the nonlinearity test results, it was found that concluded both of these series could be modelled by means of a common SETAR model. Another point in the application is that the beginning observation number is very effective for applying nonlinearity test. Accordingly, for these two series, especially depending on the first few months period, it can be said that nonlinearity structure is strong ($m = 100$).

It was also found that if the value (Z_{t-2}) which is return of gold series two days ago, is close to zero, a regime arises for series, however if this value moves away from zero, different regimes can arise.

The adequacy of model was determined by examining the PAM structure and likelihood ratio values of the residual terms of model. As seen in Figure 7, values obtained from forecasting period were close cruise to observation situation.

Modeling process of SETAR model, determining structure of model, examining predictions and residuals correspond to Box-Jenkins approach. Due to the convenience and flexibility of SETAR model, it is useful for analyzing data of economic.

REFERENCES

(2011). evds.tcmb.gov.tr. adresinden alınmıştır

Bartlett, M. S. (1938). Further Aspects of the Theory of Multiple Regression. *Mathematical Proceedings of the Cambridge Philosophical Society , 34*, 33-40.

Chan, W., Wong, A., & Tong, H. (2004). Some Nonlinear Threshold Autoregressive Time Series Models for Actuarial Use. *North American Actuarial Journal , 8* (4), 37-61.

Kahraman, Ü. M. (2012). A Study on Multivariate Threshold Autoregressive Models. *Unpublished Doctoral Thesis* . Konya, Türkiye: Selçuk University Institute of Science.

Tiao, G. C., & Box, G. P. (1981). Modeling Multiple Time Series with Applications. *Journal of the American Statistical Association , 76*, 802-816.

Tong, H. (1990). *Non-linear Time Series: A Dynamical System Approach* . New York: Oxford University Press.

Tong, H. (1978). On a Threshold Model. C. H. Chen (Dü.), *In Pattern Recognition and Signal Processing* içinde (s. 101-141). Amsterdam: Sijthoff and Noordhoff.

Tong, H., & Lim, K. S. (1980). Threshold Autoregression, Limit Cycles and Cyclical Data. *Journal of the Royal Statistical Society , B* (42), 245-292.

Tsay, R. S. (1998). Testing and Modeling Multivariate Threshold Models. *Journal of the American Statistical Association , 93* (443), 1188-1202.

Tsay, R. S. (1989). Testing and Modeling Threshold Autoregressive Processes. *Journal of the American Statistical Association , 84* (405), 231-240.