Mathematical Sciences And Applications E-Notes
Volume 1 No. 2 pp. 123–135 (2013) ©MSAEN

ANALYSIS OF TURKEY'S DOMESTIC DEBT STOCK INCREMENT WITH RR AND PCR

IBRAHIM DEMIR, GIZEM ALTUNEL, ERHAN CENE

(Communicated by Murat TOSUN)

Abstract. There are many variables that affect domestic debt. But when variables are taken into model together, it is usually seen that they have strong relationship which points out to multicollinearity problem. In a multiple linear regression model, in order to decrease the standard error of related coefficients between dependent and independent variables and in order to have more sensitive predictions, multicollinearity problem have to be solved. Some techniques overcome multicollinearity problem by using data gathering and variable elimination methods. Also in some methods biased predicting methods are used since they correct multicollinearity problem without eliminating variables but those methods give biased results.

One of the biased methods is Ridge Regression (RR) which has been widely used in case of multicollinearity. Another method is Principal Component Regression (PCR), which constitutes regression model by collecting related variables in a single variable where all of the formed variables are uncorrelated and orthogonal to each other. In this study, an application to domestic debt increment by using RR and PCR methods has been conducted.

In analysis, a multilinear regression model with 8 variables constructed between the years 1985-2010 intuitively thought to be affecting domestic debt increment. However Least Square (LS) method, RR Method and PCR method showed that some of the variables were insignificant. Thus, after the first analysis, number of independent variables were limited to four and the final analyses were conducted with four independent variables.

Analysis showed that, while RR and PCR coefficients matched with the theoretical expectations, LS coefficients gave inconsistent results. RR results indicated that the public sector borrowing requirement and the internal debt service were the only two variables that affects domestic debt increment. On the other hand PCR showed that beside those two variables the exchange rate and the government budget deficit were also effective on domestic debt increment. As a result RR and PCR applied in case of multicollinearity and in both methods significant variables met the theoretical expectations. However RR results are limited since there were only two significant variables whereas PCR results gave relatively better results in which four significant variables were found.

Date: Received: October 15, 2012; Accepted: May 24, 2013.

2000 *Mathematics Subject Classification.* 62J05, 62J07, 62P20.

Key words and phrases. Domestic Debt, Ridge Regression, Multicollinearity, Principal Component Regression.

This article is the written version of author's plenary talk delivered on September 03-07, 2012 at 1st International Euroasian Conference on Mathematical Sciences and Applications IECMSA-2012 at Prishtine, Kosovo.

1. INTRODUCTION

One of the mostly argued economic problems in Turkey is chronicle public deficit and public debts. Public deficits have been continuously appearing in Turkey. As a result of public deficits, high level of public debt arises continuously. When public debt is examined, domestic debt is particularly noteworthy. In parallel, domestic debt stock in Turkey, also shows a continuous enlargement. Dramatic aspects of this increment can be understood from investigating developments and analysing domestic debt. A lot of theoretical analyses have been conducting about Turkey's domestic debt. However, those theoretical analyses should be supported with some numerical and practical analysis. Thus, in this study not only theoretical analysis will be supported by numerical analysis, but also variables that affect domestic debt stock in economic process will be revealed.

One of the assumptions about multiple regression models states that independent variables should not be correlated. When this assumption is violated, more specifically when there is linear or nearly linear relationship between independent variables, multicollinearity problem occurs. Since such a linear relationship affects magnitudes and signs of regression coefficients, different predictions from original values may appear. Also multicollinearity may cause differences in the prediction of standard errors of regression coefficients from original values and thus leads to t-statistics values differ from original ones also. As a result of this situation multiple determination coefficient (R^2) values would be bigger than it should be [7]. So it is critical to determine if there is a multicollinearity problem and in the presence of multicollinearity proper methods should be chosen for predicting regression coefficients.

In literature, there are a lot of studies concerning domestic debt increment. While Vector Autoregressive Models (VAR) are used in some of the studies, multiple linear regression models were preferred in some others. Studies that used linear regression encountered with the multicollinearity problem and this problem is solved by reducing number of independent variables.

In contrast with other studies, in this study RR and PCR are used for modelling domestic debt increment in Turkey in case of multicollinearity. Those techniques are applied separately and then results are compared. For this purpose, first multicollinearity determining techniques will be briefly described and then RR and PCR, two methods for fixing multicollinearity, will be explained. Later, by using variables that affect domestic debt increment, RR and PCR will be performed and results will be given.

This paper will proceed as follows: Second section will be allocated for literature. While third section will mention multicollinearity, fourth section will mention LS regression, fifth section will explain RR and sixth section will describe PCR. Analyses results will take part in seventh section and paper will be concluded with the final section that deals with results and discussion.

2. LITERATURE

Vinod, H. D. [15]**,** introduces RR and research studies that have been made until his period.

Ipek, O. [9]**,** wants to predict Turkey's export numbers, between the years 1968-1995. Turkey's export, wholesale price index, consumer price index, gross domestic product (GDP) with constant producer prices, GDP per capita and real effective

exchange rate are used as independent variables and it is understood that there is multicollinearity problem. In order to solve the multicollinearity problem without eliminating correlated variables, RR technique is used for estimating parameters. As a result, it is deducted that all the independent variables have a positive effect on Turkey's export and it is mentioned that this result met theoretical expectations and RR should be used for modelling Turkey's export numbers.

Yavuz, A. [16]**,** determines Turkey's domestic stock increment by using a regression model with 8 independent variables. In his study, it is preferred to reduce number of independent variables when encountered with the multicollinearity problem and most important variables that affect stock debt increment in Turkey is specified.

Sonuvar, E. [13]**,** also wants to predict Turkey's export numbers between the years 1985-2005. First a model is obtained by using LS Method and during analysis it is understood that multicollinearity problem exists. Since multicollinearity problem exists, RR is preferred. After obtaining RR Results, it is compared with the LS results. After that, time series analysis and PCR are performed for the same data and it is deducted that, RR results and PCR results have some similarities. It is also inferred that RR results supports theoretical results.

Karakaya, E. [10]**,** analysed tourism income in Turkey's economy between the years 1985-2006 and factors affecting tourism income is tried to determine. Criteria and methods such as partial correlation coefficients, variance inflation factor and condition number criteria are used in order to determine multicollinearity and it is found that almost all of the independent variables are causing multicollinearity. In order to overcome this problem, RR and Liu parameter are applied. As a result, it is said that biased estimation techniques offer more convenient estimations from those obtain from LS methods.

El-Deren, M and Rashwan, N. I. [6]**,** state that RR can be used when multicollinearity exists, and LS results and RR results are compared over a simulated data.

3. MULTICOLLINEARITY PROBLEM

Past studies concerning domestic debt increment showed that multicollinearity problem exists. In our study, independent variables are used for modelling debt stock increment which cause multicollinearity problem too. Because of this, a brief description about multicollinearity and methods for determining and fixing multicollinearity are given in this section.

3.1. **Definition of Multicollinearity.**

One or more than one linear relationship between independent variables cause multicollinearity problem. Multicollinearity problem causes to bigger standard errors, thus leads to larger confidence intervals which decreases sensitivity of the prediction. Therefore in case of multicollinearity, one should not trust to the results of hypotheses tests.

3.2. **Results of Multicollinearity.**

In a regression analysis multicollinearity causes following problems [7]:

1. In case of perfect multicollinearity regression coefficients cannot be determined and standard errors become infinite.

2. In case of multicollinearity, variance and covariance of regression coefficients enlarge.

3. Even though multiple determination coefficient (R^2) value of the model increases, none or very few independent variables are statistically significant in terms of t-test.

4. Direction of relationship between dependent and independent variables may conflict with the theoretical and empirical expectations.

3.3. **Techniques for Determining Multicollinearity.** Multicollinearity problem can be determined by investigating correlation matrix, multiple determination coefficient, condition number, variance inflation factor, F and t-test.

Correlation matrix for the standardized independent variables is widely used and widely preferred for its simplicity. When the correlation coefficient between independent variables exceeds 0.80, one should suspect from the multicollinearity problem [12].

Another method to determine multicollinearity is to use auxiliary regressions. Auxiliary regressions are regressions in which every X_i is a dependent variable and other remaining X variables are independent variables. This technique depends on comparing multiple determination coefficient of each auxiliary regression with the main model's multiple determination coefficients. Let R_j^2 be the multiple determination coefficient of jth auxiliary regression and let R^2 be the main model multiple determination coefficient. $R_j^2 > R^2$ indicates that multicollinearity is caused by the jth variable [10].

Condition number is defined as the ratio of biggest and the smallest eigenvalue of the correlation matrix. Let θ denotes the condition number. It is said that multicollinearity is a severe problem when $\theta > 1000$, moderate problem when $100 \leq \theta < 1000$ and slight problem when $\theta < 100$ [10].

Let $X'X$ denotes the correlation matrix of the independent variables. jth diagonal element of the $(X'X)^{-1}$ is called as variance inflation factor (VIF_j). Let R_j^2 be the multiple determination coefficient of jth auxiliary regression, and then VIF_j can be alternatively written as $VIF_j = (1 - R_j^2)^{-1}$. It is usually said that a multicollinearity problem occurs when $VIF_j > 10$ [11].

Another indicator of multicollinearity is the conflict between the t-tests results and F test result. In case of multicollinearity, variables are found insignificant in t-test, while F test results indicate a significant model [1].

3.4. **Fixing Multicollinearity.** Fixing multicollinearity depends on source and degree of multicollinearity, purpose of the model and also to the importance of correlated variables.

Following methods are suggested for fixing multicollinearity:

1. Increasing sample size can lessen standard error.

2. When independent variables are in a real relationship, in some cases these related variables can be unified to a one variable in order to overcome multicollinearity.

3. All the variables in the model can be proportioned to one of the independent variables.

4. Lagged values of independent variables can be replaced by the one lagged values of dependent variable.

5. Biased estimation methods such as PCR, Shrunken Estimator, Eigenvalue Estimator, Ridge Estimator, and Liu Estimator, can be used [4].

4. LEAST SQUARE REGRESSION

LS method is the most recognized and used method for the regression analysis. It depends on minimizing the sum of squared residuals in a regression model. A regression model can be defined as

$$Y = \hat{\beta}X + \hat{\epsilon}. \tag{4.1}$$

Here, Y is a column matrix which denotes dependent variable, X denotes the matrix consisted of independent variables, $\hat{\beta}$ stand for matrix of estimation for regression coefficients β and $\hat{\epsilon}$ represents the matrix of error terms. In a multiple regression model $\hat{\beta}$ is estimated with the formula

$$\hat{\beta}_R = (X'X)^{-1}X'Y. \tag{4.2}$$

5. RIDGE REGRESSION (RR)

RR method is a biased estimation method which gives parameter estimations with the minimum variance. RR method is used for minimizing effects between independent variables and gathering decisive coefficient estimations. RR method is a similar method to LS method. RR method adds small and positive value to each diagonal elements of matrix and then estimates parameters. Regression solution can be given with

$$\hat{\beta}_R = (X'X + kI)^{-1}X'Y. \tag{5.1}$$

Here $\hat{\beta}_R$ denotes the RR coefficient, while X is the independent variables matrix, Y is the dependent variable matrix, k is the biasing constant named Ridge Parameter and I is the identity matrix. Adding constant k to the diagonal elements of correlation matrix makes condition number of the correlation matrix reduce. Since RR Solution is identical to LS Solution for $k = 0$ Ridge Estimation could be expressed as a linear transformation of the LS Estimation [3].

5.1. **Determining Ridge Parameter.** Determining optimum Ridge Parameter k is crucial since having a minimum variance and the degree of bias depend on k. Here k takes values between 0 and 1. When k gets values closer to 1, bias of estimations get bigger but variances of them become smaller.

Ridge Trace graphics are widely used for determining optimum k value. Ridge Trace graphics have Ridge Parameter (k) values in one axis and standardized regression coefficients on the other axis [8]. In these graphics, biased regression coefficients can be seen as a function of k. In case of strong linear relationship among independent variables, coefficients would change greatly for small k values, and as k gets bigger coefficients would be stabilized. Optimum k value is chosen from the stabilized area of the biased and standardized regression coefficients [3]. Another method for determining k is to look at VIF values and choose k from the area where all the VIF values come closer to 1 [2].

6. PRINCIPAL COMPONENT REGRESSION (PCR)

Another method for fixing multicollinearity problem is PCR. When independent variables are uncorrelated, correlation coefficient between these variables would be close to 0. Geometrically speaking, these uncorrelated variables are orthogonal to each other. But in multicollinearity case, highly correlated coefficients exist, thus variables are not orthogonal any more. To overcome this situation, highly correlated

variables can be gathered together and a new regression model with little number of variables than the original model can be formed. Moreover, in the new model variables are gathered in such a way that new variables are not correlated, thus they are also orthogonal. Those new variables are called Principal Components [1]. Briefly, PCR forms k dimensional space from a linear combination of variables in p dimensional space where $k < p$. These k new variables are orthogonal and uncorrelated to each other.

6.1. **Determining Number of Principal Components.** In a PCR, after finding number of components, it is crucial to decide how many component will take place in the analysis. For this purpose, lots of methods were developed. The simplest and mostly used method is to sum eigenvalues (λ) of the correlation matrix until sum exceeds 2/3 of total variance. Here λ_i denotes the ith eigenvalue of the correlation matrix. If we denote the number of new components with k, number of k and its proportion in the total variance is calculated as follows

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{i=1}^{p} \lambda_i} > \frac{2}{3}. \tag{6.1}$$

7. RR ANALYSIS OF DOMESTIC DEBT INCREMENT IN TURKEY

The government is responsible for determining expenses and balancing income and expenses in order to supply communal needs. For that, government makes a budget that shows sources and expenses. The government budget is a law or a judicial reference which estimates income and expenses and also gives permission for practising and carrying out in one - year period. In other words, public income is being collected to compensate public expenses and those two items are separately stated in the budget. If ordinary incomes cannot compensate public expenses, the budget deficit arises and new resources are needed to finance this deficit. The financial public deficit is not a desired situation. In developing economies the problems caused by the budget deficit have been widely discussed in the last decade.

Because of the impossibility of financing the public sector deficit by decreasing the expenses or increasing tax income in a short time, government tends to use other financing facilities.

The public sector deficits are tried to be covered by using three basic sources without extra taxes:

- The Central Bank Sources (Monetization),
- External Debt,
- Internal Debt.

Government benefits from the central bank's sources by borrowing. Financing of public deficit by note press is not preferable for economic stability reasons because of the direct relationship between increasing emission capacity and inflation. External debt obtains necessary sources from external markets with a cost of interest charge. Also internal debt is defined as the money that is borrowed from internal market. Internal debt is widely preferred in Turkey and also in other developing countries. Internal debt in Turkey had been specially turned into an important financing source in 1980's.

Internal borrowing becomes the most convenient source for compensating budget deficit since external debt is hard to obtain, costly and also central bank sources are not preferred for abovementioned reasons. Since internal debt management is very important in Turkey, it has to be watched carefully.

7.1. **Determining Number of Principal Components.** In this part, seven factors expected to affect Turkey's domestic debt increment (DDI) between the years 1985-2010, are taken as independent variables and a regression model with 8 variables is constructed. Data is obtained from state planning organization's web site. The internet address of the data can be seen in the bibliography [5]. Variables that take place in the model are briefly explained below.

The Public Sector Borrowing Requirement (PSB): It is described as the difference between the total public income and total public expenses except the changes in the debt. This variable usually affects magnitude of debt and especially domestic debt. It causes increment in the domestic debt stock.

Exchange Rate (ER): Value of a national currency in terms of other countries' currencies is called as the exchange rate. In some periods, government may turn the funds, which were gathered through domestic debt, to foreign currency in order to finance external debt. The magnitude of exchange rate is important in this situation. High exchange rate causes increase in domestic debt, hence supports domestic debt stock increment.

The Commercial Deposit Interest Rate (CDIR): It is an economic value that shows similarities with the internal debt interest rate and is also named as the borrowing price or the borrowing cost. Increment in the commercial interest rate increases government's domestic debt cost which causes increment in the domestic debt stock.

The Internal Debt Interest Rate (IDIR): It is the cost of the internal debt. Government may have to be in debt in order to finance interest payments. If new source of money cannot be provided, it may lead to increment in the domestic debt increment.

The Annual Inflation Rate (AIR): It expresses the range in general prices in one year period. Increment in annual inflation rate causes domestic debt stock to increase.

Government Budget Deficit (GBD): It appears when public expenses exceed public incomes. It increases domestic debt stock. In order to cover budget deficits, government loans money from domestic market thus makes a serious demand pressure on the national savings. On the internal debt situation, all or some of the private savings are transferred to public sector. While increasing financial demand at the public sector increases interest rates, deposit times starts to shorten. As deposit times get shorter, government would have to pay its coming debts at a higher interest rate and this increases domestic debt stock [14].

The Internal Debt Service (IDS): The internal debt service expresses the amount of payment (capital + interest) made in one year in a country. In other words, internal debt service represents country's domestic debt payment power or capacity. If a country gives budget deficits continuously, it could finance its debts by only taking new debts which increases domestic debt stock increment.

7.2. **Analysis.** Firstly, multiple linear regression model with seven independent variables is constructed with LS Method. While F test results find the regression

TABLE 1. Correlation Matrix for Variables

Variables	DDI	PSB	ER	CDIR	IDIR	AIR	GBD	IDS
DDI	1	0.813	0.730	-0.421	-0.251	-0.355	0.808	0.704
PSB	0.813	1	0.660	-0.426	-0.268	-0.383	0.940	0.574
ER	0.730	0.660	1	-0.764	-0.592	-0.744	0.838	0.971
CDIR	-0.421	-0.426	-0.764	1	0.906	0.923	-0.621	-0.790
IDIR	-0.251	-0.268	-0.592	0.906	1	0.935	-0.457	-0.634
AIR	-0.355	-0.383	-0.744	0.923	0.935	1	-0.588	-0.770
GBD	0.808	0.940	0.838	-0.621	-0.457	-0.588	1	0.788
IDS	0.704	0.574	0.971	-0.790	-0.634	-0.770	0.788	1

TABLE 2. Estimated LS Regression Coefficients and Confidence Intervals (CI) for 7 Independent Variables

Model	$\hat{b}$	Standard Error	$\hat{\beta}$	t	p	Lower CI	Upper CI
Constant	-17788489.144	9397702.971		-1.893	0.075	-37532330.443	1955352.155
PSB	2.649	0.626	1.856	4.233	0.001	1.334	3.963
ER	-11.266	13.241	-0.363	-0.851	0.406	-39.085	16.553
CDIR	135485.598	203392.783	0.180	0.666	0.514	-291826.784	562797.980
IDIR	-157471.649	148495.286	-0.327	-1.060	0.303	-469448.669	154505.370
AIR	313882.180	207180.564	0.487	1.515	0.147	-121388.033	749152.393
BGBD	-2.098	0.725	-1.640	-2.896	0.010	-3.620	-0.576
IDS	0.427	0.121	1.593	3.539	0.002	0.174	0.681
R^2	0.873						

model significant ($F = 17.665, \alpha = 0.05$ and $p = 0.000$), t-test results find four variables insignificant (Table-2). This situation can be seen as an indicator of multicollinearity. Even though multiple determination coefficient points out a good fit with a value of 0.873 (Table-2), existence of insignificant variables obstruct the validity of the model. Also multicollinearity is investigated with the techniques that have been mentioned in Subsection 3.3 and multicollinearity is detected. Correlation matrix showed that variables with high linear relationship exist (Table-1), which indicates a possible multicollinearity problem. After that, variance inflation factors are obtained by using auxiliary regression models and all the VIF values are found bigger than 10 (Table-3) which refers to a multicollinearity problem. Also, multiple determination coefficients of all auxiliary regression models are found over 0.90. Since those values exceed multiple determination coefficient of the main regression model ($R^2 = 0.873$), it is decided that multicollinearity exists. Condition number which is calculated by using the eigenvalues of the correlation matrix showed in Table-1, is found 446.3. Since this value far exceeds critical value of 100, it can be said that there is multicollinearity.

By using various multicollinearity detection methods, it is decided that multicollinearity exists. Thus, to solve this problem, RR and PCR are performed which were explained in Section 5 and Section 6 respectively.

Standardized RR coefficients and VIF for those coefficients are shown in Table-3 for various Ridge Parameters (k). Using this table, one can choose optimum Ridge parameter as $k = 0.1$ from the area where all the standardized RR coefficients stabilize and all the VIF values approach to 1. Also same result can be deducted

TABLE 3. Standardized RR Coefficients and Variance Inflation Factor (VIF) for 7 Independent Variables

	Standardized RR Coefficients for 7 Variable							VIF for 7 Independent Variables						
k	PSB	ER	CDIR	IDIR	AIR	GBD	IDS	PSB	ER	CDIR	IDIR	AIR	GBD	IDS
0	1.8557	-0.3632	0.1803	-0.3272	0.4867	-1.6395	1.5932	27.228	25.817	10.375	13.489	14.621	45.406	28.716
...														
0.01	1.1454	-0.0622	0.1293	-0.2395	0.4591	-0.7495	0.9912	8.9849	12.336	7.9266	8.7268	9.7706	15.051	11.533
0.02	0.8817	0.0403	0.1028	-0.1756	0.4126	-0.4206	0.7621	4.7865	7.5176	6.2979	6.1936	7.1246	7.7391	6.7202
...														
0.09	0.4811	0.1724	0.047	-0.0069	0.236	0.0641	0.3889	1.0672	1.5013	2.1268	1.6334	1.9234	1.0914	1.434
0.1	0.464	0.1763	0.044	0.0023	0.2231	0.0828	0.3716	0.977	1.325	1.901	1.455	1.699	0.945	1.282
0.2	0.376	0.1881	0.0264	0.0449	0.1462	0.1663	0.28	0.5758	0.5909	0.8221	0.67	0.701	0.3917	0.6236
0.3	0.3371	0.1856	0.0168	0.0546	0.1088	0.1891	0.2403	0.4269	0.3724	0.4758	0.4301	0.4069	0.2524	0.4079

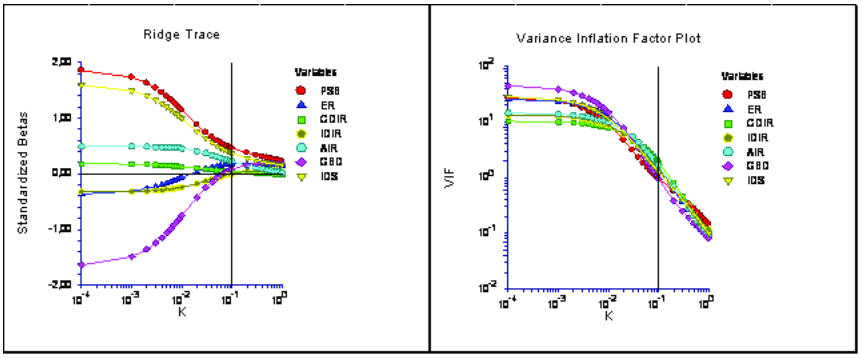

FIGURE 1. Ridge Trace and Variance Inflation Factor (VIF) Graphic for 7 Independent Variables

from Ridge Trace Graphic which serves to the same purpose with Table-3 (Figure-1).

F test results of RR which is constructed with seven independent variables indicate a significant model (F=7.4194, p=0.000292). Also all the VIF values are found smaller than 10 which points out correction in multicollinearity (Table-4). But regression model is still not valid since there are insignificant variables as a result of t-test (Table-5)

Another biased method for fixing multicollinearity is PCR. While regression model is found valid after F-test (F=7.4194, p=0.000292), some variables are found insignificant after t-test (Table-5). Even though all the VIF values are below 10 (Table-4), existence of insignificant variables make it impossible to use the model.

Due to insignificant variables, a good model cannot be obtained. By using four independent variables which are significant as result of LS method, all three models are constructed again. LS Variance Analysis of the new model whose independent variables are PSB, ER, GBD ve IDS, are found significant at $\alpha = 0.05$ level ($F = 27.081$ and $p = 0.000$). But exchange rate is found insignificant after t-test results (Table-8). So one should suspect from the multicollinearity problem, since confliction between F and t-tests can be considered as a sign of multicollinearity. Once again existence of multicollinearity problem is checked. Since highly correlated variables are still in the model (Table-1), this can be thought as the first sign of multicollinearity. Also VIF values of all independent variables are found bigger than 10 (Table-7) and condition number for the new variables are determined as

TABLE 4. Comparing VIF Values Obtained From Least Squares (LS), Ridge Regression (RR) and Principal Component Regression (PCR) for 7 Independent Variables

Variables	LS VIF VALUES	RR VIF VALUES	PCR VIF VALUES
PSB	27.2278	0.9769	0.2937
ER	25.8171	1.3249	0.0489
CDIR	10.3751	1.9012	0.1028
IDIR	13.4893	1.4545	0.2045
AIR	14.6209	1.699	0.1271
GBD	45.4064	0.9449	0.1787
IDS	28.7158	1.2822	0.0342

TABLE 5. Comparing Regression Coefficients Obtained From Least Squares (LS), Ridge Regression (RR) and Principal Component Regression (PCR) for 7 Independent Variables

Variables	LS Coeff.	RR Coeff.	PCR Coeff.	St. LS Coeff.	St. RR Coeff.	St. PCR Coeff.	LS St. Error	RR St. Error	PCR St. Error
Constant	-17788490	-12331600	-9269442						
PSB	2.64865	0.6622656	0.5331777	1.8557	0.464	0.3736	0.6257481	0.170834	0.092494
ER	-11.2657	5.468378	5.9047040	-0.3632	0.1763	0.1904	13.24132	4.323496	0.820341
CDIR	135485.6	33034.52	21639.16	0.1803	0.044	0.0288	203392.8	125492.2	28819.30
IDIR	-157471.7	1122.933	61177.44	-0.3272	0.0023	0.1271	148495.3	70281.18	26021.43
AIR	313882.2	143858.8	34895.97	0.4867	0.2231	0.0541	207180.6	101792.1	27491.52
GBD	-2.09822	0.1059664	0.4114879	-1.6395	0.0828	0.3215	0.7245712	0.150651	0.064685
IDS	0.4273954	0.0996845	0.0393587	1.5932	0.3716	0.1467	0.1207819	0.036786	0.005933
R^2	0.8729	0.736	0.7426						

TABLE 6. Standardized RR Coefficients and Variance Inflation Factor (VIF) for 4 Independent Variables

	Standardized RR Coefficients for 4 Variable				VIF for 4 Independent Variables			
k	PSB	ER	GBD	IDS	PSB	ER	GBD	IDS
0	2.0812	-0.5156	-1.9131	1.5177	25.1435	24.3425	42.4964	28.026
...								
0.01	1.3111	-0.1799	-0.942	0.8598	8.6799	11.9032	14.778	10.8553
0.02	1.0137	-0.0628	-0.5683	0.6184	4.7107	7.2962	7.7469	6.119
0.09	0.5399	0.096	0.0057	0.2716	1.0899	1.376	1.1087	1.1882
0.1	0.5185	0.102	0.0288	0.2585	1.0002	1.2039	0.9578	1.0601
0.2	0.4056	0.1303	0.1344	0.1987	0.5984	0.5144	0.3873	0.5323
0.3	0.3551	0.14	0.1662	0.1792	0.4448	0.3272	0.245	0.3695

277.96. Since all of those criteria indicate multicollinearity once again, RR and PCR can be performed.

Standardized RR coefficients and VIF for those coefficients are shown in Table-6 for various Ridge Parameters (k). Using this table, one can choose optimum Ridge parameter as $k = 0.1$ from the area where all the standardized RR coefficients stabilize and all the VIF values approach to 1. Also same result can be deducted from Ridge Trace Graphic which can be used for the same purpose with Table-6 (Figure-2).

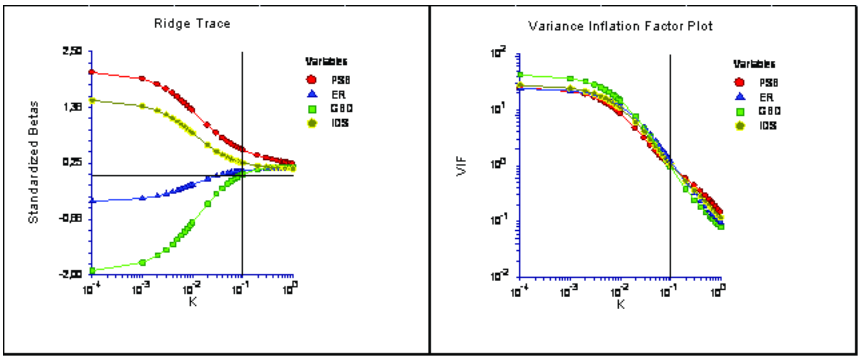

FIGURE 2. Ridge Trace and Variance Inflation Factor (VIF) Graphic for 4 Independent Variables

TABLE 7. Comparing VIF Values Obtained From Least Squares (LS), Ridge Regression (RR) and Principal Component Regression (PCR) for 4 Independent Variables

Variables	LS VIF VALUES	RR VIF VALUES	PC VIF VALUES
PSB	25.1435	1.0002	0.0641
ER	24.3425	0.9578	0.0775
GBD	42.4964	1.0601	0.0814
IDS	28.026	1.2039	0.0717

TABLE 8. Comparing Regression Coefficients Obtained From Least Squares (LS), Ridge Regression (RR) and Principal Component Regression (PCR) for 7 Independent Variables

Variables	LS Coeff.	RR Coeff.	PC Coeff.	St. LS Coeff.	St. RR Coeff.	St. PC Coeff.	LS St. Error	RR St. Error	PC St. Error
Constant	-61789.371	83556.91	-1057545						
PSB	2.970	0.7400	0.2993	2.0812	0.5185	0.209	0.6293	0.1702	4.42E-02
ER	-15.991	3.163	7.153	-0.515	0.102	0.2306	13.456	4.059	1.056294
GBD	-2.448	0.0368	0.3023	-1.9130	0.2585	0.2363	0.7336	0.1493	4.46E-02
IDS	0.4071	0.0693	0.0594	1.5176	0.0288	0.2218	0.1248	3.29E-02	8.79E-03
R^2	0.8376	0.7012	0.6859						

RR results for $k = 0.1$ shows that while F-test result ($F = 12.3228$ and $p = 0.000026$) suggest that model is significant at $\alpha = 0.05$, t-test results find variables exchange rate and budget deficit insignificant (Table-8). Thus RR model is still not very trustworthy even VIF values are below 10 (Table-7). When those two insignificant variables are removed from the model then multicollinearity problem is removed and RR is not needed any more.

When looked at the PCR, both F test ($F = 11.4657$, $\alpha = 0.05$ and $p = 0.000043$) and t-tests are found significant (Table-8). Also since all the VIF values are below 10 (Table-7), it can be said that multicollinearity problem is solved by using PCR.

LS MODEL

$$DDI = -61789.371 + 2.970 \times PSB - 15.991 \times ER - 2.448 \times GBD + 0.4071 \times IDS. \quad (7.1)$$

RR MODEL

$$DDI = 83556.91 + 0.740 \times PSB + 3.163 \times ER + 0.0368661 \times GBD + 0.069 \times IDS. \quad (7.2)$$

PCR MODEL

$$DDI = -1057545 + 0.299 \times PSB + 7.153 \times ER + 0.3023 \times GBD + 0.0594 \times IDS. \quad (7.3)$$

Results obtained from LS Regression, RR and PCR have been summarized in Table-8. It can be seen from the table that only LS Model gave variables with negative coefficient (See (3)-(5))which conflicts with the theoretical expectations given in subsection 7.1 in which all of the variables are expected to increase domestic debt increment. Despite their higher multiple determination coefficient (R^2), it can be inferred that multicollinearity cannot be solved with those models, since there are still insignificant variables in LS and RR results. However all the variables in the PCR are significant and PCR solved the multicollinearity with relatively less multiple determination coefficients.

Result And Discussion. Multicollinearity is detected between the variables that explain domestic debt stock increment. In order to make estimation with the minimum error, RR and PCR are preferred instead of LS Method in case of multicollinearity. First, all of the variables are put in the model together and LS, RR and PCR are performed. After the first analysis, insignificant variables are eliminated from the model and final results obtained. According to results, LS method did not match the theoretical expectations since signs of some variables are found negative contrary to what economy theory suggests. RR results indicate that multicollinearity problem is solved with 4 variables but since there are still insignificant variables, RR results are not satisfactory, even though it has positive coefficients for all the variables. PCR gave the best result among these models, with all of the 4 variables are significant. PCR also meets the demands of theory with positive signed coefficients. As a result, while the only two variables that affect Domestic Debt Stock Increment are Public Sector Borrowing Requirement and Internal Debt Service in RR, PCR results show that along with these variables Exchange Rate and Budget Deficit also affect Domestic Debt Stock Increment too.

References

[1] Alpar, R., Çok değişkenli istatistiksel yöntemler, Nobel,Ankara,2003.

[2] Ayyıldız, H. and Kürşat Ö., A study on sales forecasting for pharmaceutical industry in Turkey. *Journal of Economics and Administrative Sciences, University of Eskişehir Osmangazi* (2011),6(1),71-102.

[3] Büyükuysal, M.Ç., Ridge regresyon analizi ve bir uygulama, Uludağ University, The Institute of Health Sciences,(2010), Master Thesis.

[4] Canküyer, E. and Sönmez, H., The study and an application of the problem of multicollinearlty in regression analysis, *Anatolian University, Journal of Science Faculty* (1996), no.2, 173-187.

[5] State planning organization: economic and social Indicators 1950-2010, http://www.dpt.gov.tr/PortalDesign/PortalControls/WebIcerikGosterim.aspx?Enc=83D5A6FF03C7B4FC5A73E5CFAD2D9676

[6] El-Deren, M and Rashwan, N. I., Solving multicollinearity problem using ridge regression models, *Int. J. Contemp. Math. Sciences,* (2011), Vol. 6, no. 12, 585 600.

[7] Gujarati, D. N. and Porter C. D., Basic econometrics, 5th Edition, McGraw-Hill/Irwin, New York, 2008.

[8] Hoerl, A. E. and Kennard, R. W., Ridge regression: applications to non-orthogonal problems, *Technometrics*, (1970), Vol. 12, No. 1, 69-82.

[9] Ipek, O., A study on ridge regression., *V. National Econometrics and Statistics Symposium*, Çukurova University, (2001).
[10] Karakaya, E., The comparison of the efficiency and bias of ridge and liu estimators, Anadolu Univesity, Statistics Program, (2011), Master Thesis.
[11] Kaşko, Y.,The observed type I error and power of logistic regression model under multicollinearity, Ankara University, Department of Animal Science, (2007), Master Thesis.
[12] Montgomery, D.C. and Peck, E.A., Introduction to linear regression analysis, John Wiley and Sons, No: 641, New York,1991.
[13] Sonuvar, E., Ridge regression analysis of Turkey's export, Yildiz Technical University, Graduate School of Natural and Applied Science, (2007), Master Thesis,.
[14] Şen, H. and Sağbaş, I., Bütçe açıkları: teori ve Türkiye uygulaması, Seçkin Yayınevi, 2004.
[15] Vinod, H. D. (1978), A survey of ridge regression and related techniques for improvements over ordinary least squares, *The Review of Economics and Statistics,* (1978), Vol. 60, no. 1, 121-131.
[16] Yavuz, A., Türkiyede iç borç stoğundaki değişimin analizine yönelik bir regresyon analizi çalşması, *Süleyman Demirel University The Journal of Faculty of Economics and Administrative Sciences,* (2003), Vol. 8, Issue: 1, 339-356.

Yildiz Technical University, Department of Statistics, Istanbul TURKEY
E-mail address: idemir@gmail.com