

Analysis of The Factors Affecting The Profitability of Banks In Turkey By Lasso Regression

Gizem Nur Bayrakçı ^{a*}

^aBanking Regulation And Supervision Agency (BRSA), İstanbul, Turkey

Abstract

In the article examining the factors affecting the profitability indicators of the banking sector, which has a large share in the financial sector in Turkey, the analysis of the factors affecting the return on assets and return on equity ratios of the top 10 deposit banks in 2020, ranked according to their asset size in 2020, was made with the lasso regression method. As a result of the analysis, separate models were established for return on assets and return on equity. Among the independent variables, only the deposit variable was excluded from the model and it was seen that the effect on profitability was not as much as the variables of securities portfolio, equity, NPL ratio, asset share, inflation and gross domestic product.

Keyword

Profitability in Banking, Return on Assets, Return on Equity, LASSO regression method

1. INTRODUCTION

Banks, whose main task in the financial sector is to carry out financial intermediation transactions between fund suppliers and fund demanders, have an important share in the financial sector. Banks working for profit, like many financial institutions, play an important role in the smooth running of the economy in the transactions they make for this purpose, while ensuring the stable functioning of the financial structure. Interpreting the profitability indicators of banks that provide information about the state of the financial structure in Turkey and in the world; It provides general information about the financial sector.

As of the end of 2020, the share of banks in the financial sector in Turkey is over 90 percent. For this reason, the profitability indicators of the banks examined in Turkey should not be seen only as a sectoral size, but rather as a sector in which the changes in the country's economy can be monitored due to its large share in the financial sector.

Among the profitability indicators in the article, return on equity (ROE) and return on assets (ROA) variables were selected as dependent variables, and the effects of the variables of securities portfolio, deposit, equity,

* Corresponding author.

Contact: Gizem Nur Bayrakçı ✉ gizemnurbayrakci@gmail.com

To cite this article: Bayrakçı G. N. (2022). Analysis of The Factors Affecting The Profitability of Banks In Turkey By Lasso Regression. *AYBU Business Journal*, 2(1), 46-57.

The opinions presented in the study belong to the author and do not bind the Banking Regulation and Supervision Agency or its employees. Responsibility belongs to the author.



NPL ratio, asset share, inflation and gross domestic product were examined. LASSO regression method was used in the analysis part of the article. When we look at the literature, it has been seen that more than one study has been done on the profitability of banks, it has been observed that panel data analysis and regression methods are used more in the analyzes, while LASSO regression method is not used as widely as panel data analysis.

In the article, the data of the top 10 deposit banks with the largest asset size as of December 2020, from the banks in Turkey between the years 2011-2020 were obtained. The application results were obtained and interpreted using the R programming language.

According to Article 3 of the Banking Law No. 5411, banks; It is divided into 3 as Deposit Banks, Development and Investment Banks and Participation Banks. Banks are supervised and regulated by the Banking Regulation and Supervision Agency (BRSA).

Financial statements are the tables in which the transactions of commercial enterprises in a certain period are monitored. Financial statements; It consists of balance sheet, income statement, cash flow statement and statement of changes in equity. The data to be used for the article are taken from the income statements and balance sheets of banks.

The balance sheet is the tables that show the financial situation of the banks in a certain period. In the banks balance sheet, the assets and resources of the bank are accounted for in accordance with their fair values under active and passive accounts ([Akgüç, 2011](#)). Active accounts include the bank's receivables, passive accounts include equity and the bank's debts, which consist of foreign resources. Bank balance sheet is based on the principle that active and passive accounts are equal to each other ([Takan, 2001](#)). Loans in active accounts are among the items with the largest share in the banks balance sheet. Foreign resources (syndication loan, etc.) are mainly used as a source of funds in passive accounts, apart from the shareholders' equity ([Tunay, 2005](#)).

Income statement, on the other hand, is the financial statement that shows the income of the bank within an operating period (usually 1 year) and the expenses, costs and profits or losses they have obtained as a result of their activities. The income statement, which expresses the accounting equivalent of net income, shows the summary of the bank's performance ([Takan, 2001](#)). The main items that make up the income statement are interest income-expenses, non-interest income-expenses and net profit for the period. Interest income-expenses, the interest paid by the banks on their deposits and debts and the interest income they receive from the loans they extend, the income part of the non-interest income-expenses; net fees and commissions, foreign exchange income, expense from money market and capital market transactions; personel expenses, provisions, foreign exchange expenses consist of tax and depreciation expenses, and the net profit for the period consists of the sum of the items that have been counted in general terms. Net profit for the period is one of the most important profitability indicators showing the profit obtained as a result of the activities of the bank in an operating period.

Financial statements of banks are disclosed to the public quarterly. Financial statements disclosed to the public cause competition among banks. The resulting competition enables banks to diversify their products, see the effects of efficient use of resources on their performance, and evaluate their own performance within the sector.

Profitability is an indicator of the success of an enterprise in carrying out its activities. The main purpose of banks and financial institutions is to make a profit, as in other enterprises. Making a profit is very important in order to sustain the existence of banks, increase their market value and meet the expectations of their shareholders. Profitability in banking, on the other hand, expresses the profits of the banks in a certain period as a ratio, not as an amount ([İskenderoğlu et al., 2012](#)). Profitability rates are calculated on the basis of the financial statements (income and balance sheet) announced by the banks. Some ratios are calculated by taking related items and some ratios are calculated by bringing together the relevant amounts. ([Akguc, 2011](#)). Profitability indicators used as dependent variable in the article; return on assets (ROA) and return on equity (ROE).

Return on assets is an indicator of how effectively the assets used by the bank in its activities are used and how much return is obtained on the assets. It is accepted as an efficiency criterion that informs the

shareholders of the banks about the management efficiency of the bank they are a shareholder of (Şıklar, 2004). The formula for return on assets:

$$ROA = \frac{\text{Net Profit For The Period}}{\text{Average Total Assets}} \quad (1)$$

Return on equity shows the degree of profitability of the capital in the bank, it is a performance criterion that measures the return on the investment they make, in other words, maximizing the profit, which is the main goal of the bank's shareholders (Şıklar, 2004). The formula for return on equity:

$$ROE = \frac{\text{Net Profit For The Period}}{\text{Average Total Equity}} \quad (2)$$

2. LITERATURE

Yıldırım et al. (2021), applied LASSO (Least Absolute Shrinkage and Selection Operator) regression, LARS regression and multiple linear regression models from penalized regression models to the dataset of 442 patients diagnosed with diabetes. As a result of the analysis of LASSO regression and LARS regression, it was seen that the same variables were found to be significant and had a higher R2 and mean square error than the multiple linear regression model.

With the LASSO method seen in the Financial Stability Report published by the CBRT in May 2020, it was investigated whether the variables on the credit default swap (CDS) were significant in the model. In the research, it was seen that the effect of exchange rate, exchange rate changes, GDP ratio, VIX index, indebtedness rate of financial institutions, foreign exchange reserves, loan interest rate, foreign exchange reserves, dollar and local currency interest difference on CDS. Since the downward movement of the risk premium is known to have a positive effect on profitability, it has been taken into account in the study while scanning the literature.

While it is seen that multiple regression method and panel data analysis method are used when examining the data explaining profitability in banking in the studies in the literature, it is seen that LASSO regression is not used as widely as panel data analysis in economic data. When we look at the studies on LASSO regression, it is seen that this method is mostly used in the analysis of data related to health sciences and in studies on machine learning, but it has not been encountered as often as in other study areas in the analysis of economic data. The use of the LASSO regression method in the article offers an alternative to the other analyzes in terms of using the LASSO regression method in economic data.

3. MATERIAL AND METHOD

The data used in the article; It is the year-end data included in the financial statements of banks announced to the public as of 2011 - 2020. The banks subject to the analysis are the top 10 deposit banks with the largest asset size as of the end of 2020. The data of the variables were obtained from the financial statements of the banks and the information on the websites of the CBRT, BRSA, TURKSTAT and The Bank Association of Turkey.

In the application part of the article, the LASSO regression method, which has a lower estimation error compared to other regression methods and makes the selection of the variable at the same time while analyzing the independent variables that explain the dependent variables, is seen to be more successful than the least square method, ridge regression method and panel regression method. Evaluation was made in the conclusion part. The R program was used while performing the LASSO regression method analysis.

4. LASSO REGRESSION METHOD

LASSO (Least Absolute Shrinkage and Selection Operator) regression method is a regression method that reduces the non-significant coefficients in the model towards zero and provides meaningful variables. The method was first introduced by the statistician “Robert Tibshirani” in 1996, by using punishment in model selection in the field of geophysics, and was later developed.

In the LASSO method, by using the penalty term in the L1 norm, the regression coefficients are reduced. Among the shrinking variables, there are also independent variables with a zero coefficient. These variables with zero coefficients are excluded from the model and the independent variables that explain the dependent variable in the model remain. In the LASSO method, by excluding the variable that is not suitable for the model, variable selection is made at the same time. Thanks to the excluded variable or variables, it becomes easier to interpret the model as the variables that better explain the model remain in the model. (Kayanani et al.,2019)

In the regression analysis, while all variables remained in the model, forward selection, backward selection and stepwise selection methods were used before the LASSO method was used. Since these methods do not take into account the relationship of the variables with each other, they cause misleading estimations. In the LASSO regression method analysis, the significant variables explaining the dependent variable remain in the model. With this feature, it has been seen that the LASSO method, which helps in the estimation of the regression coefficient and the selection of the compatible variable to be included in the model, is more successful than the ridge regression method when the number of observations is larger than the number of independent variables and the independent variables are related to each other. (Zou and Hastie, 2005).

In the LASSO method, the LASS estimator is reduced by the penalization method to obtain a smaller mean of variance and error squares. The variables to be included in the model are standardized and the transformation is made so that the mean is 0 and the variance is 1.

Penalized function for linear regression method:

$$Q_N = \left(y_i - \sum_j x_{ij} \beta_j \right)^T \left(y_i - \sum_j x_{ij} \beta_j \right) + \lambda \sum_{j=1}^p |\beta_j| \tag{3}$$

When it is desired to make parameter estimation and variable selection from this objective function,

$$Q_N = -\ln L(\theta) + \lambda \sum |\beta_j|$$

The likelihood function in the form is minimized and used for parameter estimation and variable selection.

$\sum_{j=1}^p |\beta_j|$: L1 norm penalty term

λ : setting parameter ($\lambda > 0$)

θ : parameter vector

$\ln L(.)$: log-likelihood function

The LASSO method equation was developed with the aim of minimizing the mean squared error by forcing the sum of the absolute values of the regression coefficients to be less than a fixed value. The λ value is the parameter that adjusts the shrinkage of the coefficient estimates. If $\lambda=0$, the equation becomes the same as the classical regression method model estimator, the EKK method, while the model turns into a multiple

regression model. (If λ is large, it means that all units are penalized, and small means that penalization is not done much.)

LASSO regression method is also included in machine learning algorithms. In machine learning, an inference is made from the data consumed and prediction results are obtained. A high λ causes over-learning, i.e. memorizing the data on the training data in the used model, resulting in success only on that data, while a λ of 1 causes all variables to be excluded and the bias to increase due to the low coefficient values. In this case, incomplete learning, that is, the model used does not fit the training data and the model is not suitable for new data. In overlearning, the model has high variance and low side, while incomplete learning has low variance and high side. LASSO regression method used to obtain the highest efficiency with the least variable; It increases the quality of the trained data by finding the best λ value describing the model and making coefficient estimations with the best λ found, and ensures that the coefficients of the most significant variables remain in the model.

5. ANALYSIS

5.1. Variables Used in the Scope of Analysis and Their Definitions

The variables used within the scope of the analysis belong to the top 10 deposit banks, which have the highest share in the banking sector as of the 4th quarter of 2020, among the banks in Turkey. The data on the variables were obtained from the financial statements of the banks, which were announced to the public, and from institutions such as TURKSTAT and CBRT between the years 2011-2020. Variables taken from banks balance sheets are called microeconomic variables, and economic variables that are not determined by banks that are not included in banks' balance sheets but affect the performance of financial institutions are called macroeconomic variables.

Table 1: Banks included in the analysis

The Name Of The Bank			
1	T.C. Ziraat Bankası A.Ş.	6	Yapı ve Kredi Bankası A.Ş.
2	Türkiye İş Bankası A.Ş.	7	Türkiye Vakıflar Bankası T.A.O.
3	Türkiye Garanti Bankası A.Ş.	8	QNB Finansbank A.Ş.
4	Akbank T.A.Ş.	9	Denizbank A.Ş.
5	Türkiye Halk Bankası A.Ş.	10	Türk Ekonomi Bankası A.Ş.

Table 2: Variables included in the study

		Variables	Ratios Taken For Analysis	
Dependent Variables	Micro Variables	Return on Assets	(Net profit for the period / Average total assets) (ROA)	
		Return on Equity	(Net profit for the period / Average total equity) (ROE)	
Securities Potfolio		(Securities portfolio / Total assets) (SP)		
Deposit		(Total deposits / Total assets) (DEP)		
Equities		(Total equity / Total assets) (EQU)		
Asset share		(Bank's assets / Total assets) (AS)		
NPL ratio		(Non-performing loans (Gross) / Total loans) (NPL)		
Independent Variables		Macro Variables	Inflation	Inflation (INF)
			Gross Domestic Product	Gross domestic product quarterly growth rates (GDP)

The dependent variables used in the application are the performance indicators of the banks.

- Return on Assets (ROA)
- Return on equity (ROE).

Among the independent variables used in the analysis, microeconomic variables specific to the bank are equity, deposits, NPL ratio, securities portfolio and asset share, while macroeconomic variables are inflation

and gross domestic product. Microeconomic variables taken from banks' balance sheets were made suitable for analysis by dividing each variable by total assets due to the size of the amounts in the financial statements. Macroeconomic variables were included in the analysis by taking annual percentage rates.

While the data obtained from various sources were obtained, the dependent variables of microeconomic variables, return on assets and return on equity, were used by annualizing according to the end of the year, while the amounts valid as of the end of the year were taken into account in the independent variables. The descriptions of the variables used in the application are as follows:

Return on Assets (ROA): It is a ratio showing the profit from the bank's assets. It is one of the important profitability indicators as it gives information about the effective use of the assets of the banks.

Return on Equity (ROE): It is a ratio showing the profit obtained from the equity of the bank. It is one of the important profitability indicators as it gives information about the efficient use of banks' equity and is also used to measure the return on capital of the bank's shareholders.

Assets: These are the total assets of the bank.

Equity: It is included in the bank's resources and refers to the resources provided by the bank's partners. It is a source that shows the soundness of the bank in times of crisis.

Deposit: These are the funds that the bank receives from the savers.

Credit: It is the purchasing power provided by the bank to real or legal persons, provided that they are bought back.

Non-performing loans: Loans are classified as non-performing loans if the loan payments given to real or legal persons exceed 90 days. While the ratio of non-performing loans gives information about the solvency of the persons or institutions to which the loan is given, it also gives information about the asset quality of the bank.

NPL ratio: It shows the rate of return of the loans provided by the banks. (Saritaş et al.: 2016)

Securities Portfolio: It consists of government debt securities, and other papers (bonds, bills...).

Inflation: It expresses the increase (decrease) in the general level of prices.

Gross Domestic Product (GDP): It is the market equivalent of the products produced within the borders of the country in a certain time period.

While analyzing, the explanatory variables of return on assets and return on equity, which are profitability indicators, were examined within the scope of the model. LASSO regression method was applied for the created model and the best variables explaining the dependent variables, return on assets and return on equity, were obtained in the analysis. The interpretation of the variables was made according to the results of the analysis.

5.2. LASSO Regression Analysis

During the application, the dependent variables were modeled separately. Before starting the application, in order to obtain unbiased and consistent results, the distribution tables of the variables explaining both dependent variables, correlation analysis table, and assumption tables were examined separately in accordance with the LASSO regression method. The change distribution of the microeconomic and macroeconomic independent variables specific to banks used in practice by years is as follows.

Models:

$$ROA = y_1 + \beta_1 SP + \beta_2 DEP + EQU_{0z} + \beta_4 AS + \beta_5 NPL + \beta_6 INF + \beta_7 GDP + \varepsilon_{it}$$

$$ROE = y_1 + \beta_1 SP + \beta_2 DEP + \beta_3 WQU + \beta_4 AS + \beta_5 NPL + \beta_6 INF + \beta_7 GDP + \varepsilon_{it}$$

Descriptive statistics:**Table 3:** Descriptive statistics of variables in the LASSO regression method

	ROA	ROE	SP	DEP	EQU	AS	NPL	INF	GDP
Minimum	0,06	0,23	0,06	0,48	0,06	0,03	-0,03	0,25	0,13
Maximum	0,16	0,54	0,43	0,68	0,15	0,19	0,07	0,45	0,22
1. quarter	0,11	0,34	0,12	0,55	0,10	0,05	0,03	0,29	0,14
3. quarter	0,13	0,42	0,19	0,60	0,12	0,14	0,06	0,35	0,17
Mean	0,12	0,38	0,16	0,57	0,11	0,10	0,04	0,32	0,16
Median	0,12	0,38	0,15	0,57	0,11	0,11	0,05	0,31	0,16
Variance	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Standard Deviation	0,02	0,06	0,06	0,04	0,02	0,04	0,03	0,06	0,02
Skewness	-0,31	0,22	1,56	0,41	-0,10	-0,19	-1,20	0,85	0,64
Kurtosis	0,22	0,83	4,34	-0,35	-0,08	-1,17	0,73	0,00	-0,17

Table 4: Correlation matrix for the variables in the LASSO regression method

	ROA	ROE	SP	DEP	EQU	AS	NPL	INF	GDP
ROA	100%								
ROE	89%	100%							
SP	26%	28%	100%						
DEP	-10%	1%	11%	100%					
EQU	56%	15%	7%	-19%	100%				
AS	24%	17%	63%	-11%	17%	100%			
NPL	-13%	-14%	4%	6%	-12%	0%	100%		
INF	-24%	-14%	-19%	3%	-15%	-3%	-29%	100%	
GDP	13%	-5%	4%	9%	52%	-3%	-2%	33%	100%

In the correlation analysis table; It is seen that the highest positive relationship among the variables is between ROA and ROE variables with 89%, and the lowest negative relationship is between inflation variables with -24%. When the relationship of ROA, which is one of the dependent variables subject to the analysis, is examined, it is seen that it has the highest positive relationship with the equity variable with a rate of 56% and the lowest relationship with the inflation variable with a rate of -24%. When the relationship between ROE, which is the other dependent variable, is examined, it is seen that it has the highest positive relationship with the securities portfolio with a rate of 28%, the lowest relationship with the NPL ratio and inflation variables with a rate of 14%.

5.3. Assumptions

In practice, the variables affecting the profitability of deposit banks, which were among the top 10 banks in Turkey as of 2020, were analyzed using the LASSO regression method. The ratios obtained from the financial statements of the year-end data between the years 2011-2020 were used. For the LASSO regression method; It was seen that the data did not provide the assumptions. In order to eliminate the problem in the data, the square root transformation was applied to the dependent variables, return on assets (ROA) and return on equity (ROE), and the independent variables, the inflation variable, and the analysis continued. The test values obtained as a result of the analyzes are as follows.

5.3.1. Assumptions for LASSO Regression Analysis

For LASSO regression analysis assumptions, normality, autocorrelation and varying variance assumptions, which are the assumptions used to determine the suitability of the data for analysis in the regression equation, were examined separately for the dependent variables return on assets and return on equity. The null hypotheses of the assumptions to be tested for return on assets and return on equity are as follows:

H_0 : There isn't autocorrelation problem.

H_0 : The model provides the assumption of normality.

H_0 : There isn't problem of changing variance.

Table 5: Assumption test values for return on assets

ROA									
Assumptions	Test Name								
Autocorrelation Test	VIF	mdc	mev	oz	akpay	tdo	enf	gdp	
		1,98	1,14	2,11	1,99	2,19	1,78	1,21	
Normality Test	Shapiro-Wilk Testi	0,19							
Heteroskedasticity	Breusch-Pagan Testi	0,21							

Table 6: Assumption test values for return on equity

ROE									
Varsayım	Test Name								
Autocorrelation Test	VIF	mdc	mev	oz	akpay	tdo	enf	gdp	
		1,98	1,14	2,11	1,99	2,19	1,78	1,21	
Normality Test	Shapiro-Wilk Testi	0,07							
Heteroskedasticity	Breusch-Pagan Testi	0,11							

In the regression analysis, the assumption, also known as the autocorrelation or multicollinearity problem, was tested with the VIF values, which are the variance inflation factors. The fact that the VIF values used to measure the relationships of the independent variables with each other are greater than 10 indicates the existence of an autocorrelation problem. Looking at the test results, it was seen that the test values of the variables were less than 10, so there was no autocorrelation problem. When the normality test and the varying variance tests were examined, it was seen that their values were greater than the significance value of $\alpha=0.05$. The test values being greater than $\alpha=0.05$ significance level means that the null hypotheses cannot be rejected. Considering the undeniable hypotheses and the VIF coefficient, it is seen that the variables are normally distributed and they do not have the problem of varying variance and autocorrelation.

5.3.2. Model selection and estimation results for the LASSO regression method

In the LASSO regression method, the most suitable variables for the model are found by the penalization method as a result of the analysis. The selection of models to be found separately for ROA and ROE variables is as follows.

Table 7: LASSO regression method estimation results

Variables	Coefficients of ROA	Coefficients of ROE
Intercept	0,07	0,28
SP	0,09	0,26
DEP	0,00	0,00
EQU	0,67	1,03
INF	-0,05	-0,05
GDP	-0,12	0,80
NPL	-0,04	-0,01
AS	-0,09	-0,47
R Square	0,90	0,96
Best Lambda Value	0,001	0,005

For ROA and ROE dependent variables according to the LASSO regression method, when the independent variables used in the application explain the model, it is seen that the independent variables in the model explain the ROA model by 90% and the ROE model by 96%. The most suitable lambda values for the models are 0.001 for ROA and 0.005 for ROE. While choosing the model, a selection was made by considering the a priori expectations (all the independent variables used in the application to be meaningfully included in the model) and the most appropriate lambda values.

In the LASSO regression method, the coefficients of the variables that do not explain the model as significantly as other variables are reduced to 0 and removed from the model. Considering the compatibility of the independent variables with the model according to the estimation results of the models established for profitability, it is seen that the coefficient of the deposit variable, which is one of the explanatory variables for the models of ROA and ROE variables, is 0. In this case, it is seen that the deposit variable does not explain the model as well as the other variables, and the model continues with the SP, EQU, INF, GDP, NPL and AS variables other than deposits.

When the variable coefficients for two different models are examined, it is seen that while the relationship between the dependent variables and the dependent variables is the same, except for the GDP variable, only the GDP variable has a negative relationship with the ROA variable and a positive significant relationship with the ROE variable.

In the ROA model, it is seen that the variables other than the DEP variable remained in the model. While the relationship between SP and EQU variables and ROA is positive, the relationship between INF, GDP, NPL and SP variables and ROA is negative and statistically significant. Looking at the coefficients in the model, it is seen that a one-unit change in the independent variables of EQU and GDP are the variables that will affect ROA the most positively and negatively, respectively.

In the ROE model, it is seen that the variables other than the DEP variable remain in the model. There is a positive relationship between SP, EQU and GDP variables and ROE, and a negative and statistically significant relationship between INF, and SP variables and ROE. Looking at the coefficients in the model, it is seen that a one-unit change in the independent variables of EQU and SP are the variables that will affect ROE the most positively and negatively, respectively.

While there is a positive relationship between the SP variable, which expresses the financial assets of the banks, and ROA and ROE, it is seen that the ROE variable is more affected by the SP variable and will be more positively affected by the changes that will occur in the securities. As of 2011, due to the decrease in the demand for GDDS (Government Debt Securities), the decrease in the share of SP in the balance sheet, together with the increase in the demand for credit usage and the financing of the public finance deficit with money market instruments such as bonds and bills, have a positive contribution to profitability.

When the DEP variable, which expresses the deposits of the banks, is examined, it is seen that the ROA and ROE variables are not included in the models. Considering the DEP variable, when the pressure of the interest given to the deposit on the profitability of the bank is taken into account, it is seen that the

profitability affects negatively and when it is considered together with the other independent variables in the model, the DEP variable is the variable that is excluded from the variables that explain the model.

It is seen that the EQU variable, which expresses the equity of the banks, is the independent variable that affects the ROA and ROE variables positively. It is seen that the positive effect seen in the balance sheet has a positive effect on profitability, with the reflection of the positive effect on the income statement due to the fact that the strong capital structure increases the usage area of the equity capital within the legal restrictions in the transactions made by the bank.

There is a negative relationship between the NPL variable, which expresses the NPL ratio of the banks, and the ROA and ROE variables. The increase in the NPL variable indicates an increase in cases where banks do not take their loans on time or do not take them at all, considering the significant ratio of loans extended to customers in the balance sheet. It is seen that it has a negative effect on profitability as it significantly affects the asset quality of banks and causes the asset-liability balance to deteriorate.

It is seen that there is a negative relationship between the AS variable, which expresses the share of banks in the banking sector, and the ROA and ROE variables in the article. The fact that the AS variable has a negative effect on profitability shows that the profitability of banks varies depending on the transactions they make and the effective use of these transactions, not economies of scale.

It is seen that the inflation variable, which is one of the macroeconomic variables that affect the profitability of banks, is negatively related to ROA and ROE variables, and it affects both dependent variables negatively at the same rate. It is considered that the negative effect of the INF variable is due to the negative effects of the increase in the general level of prices, such as the increase in borrowing costs, financial uncertainty and dollarization, and the deterioration of the asset-liability balance on the profitability of the banks.

There is a negative relationship between the GDP variable, another macroeconomic variable that affects the profitability of banks, and the ROA variable, and there is a positive relationship with the ROE variable. Looking at the GDP variable, the changes in the balance sheets of banks with the increase of the GDP variable are not the same in the ROA and ROE models, the ROE variable is more sensitive than the macroeconomic variable GDP, the positive contribution of growth while increasing the macroeconomic activities and stability makes a positive contribution to the ROE variable and positively affects the profitability. It is seen that negative effects such as increasing cost and decreasing demand have a negative effect on the ROA variable and have an effect on the profitability not increasing as expected.

6. CONCLUSION AND EVALUATION

In the article, the factors affecting the profitability of the top 10 deposit banks, which are in the banking sector, which has a large share in the financial system in Turkey as of December 2020, and selected according to their asset size as of December 2020, between 2011 and 2020 are examined. Securities portfolio, deposit, equity, NPL ratio, asset share, inflation and gross domestic product data were used as independent variables for profitability, which was examined through the return on assets and return on equity criteria.

While obtaining the results, separate models were established for the dependent variables, return on assets and return on equity, and while analyzing the model to see the effects of the independent variables on the dependent variables, the LASSO regression method, which excludes the variables with the penalization method and leaves the best independent variables in the model, was used.

In the LASSO regression method, the deposit variable was excluded from the models. In terms of the relationship of the coefficients, it is concluded that the GDP variable is negatively related to the return on assets, positively related to the equity variable, and the coefficients of the other independent variables with different values are in the same direction in both models.

In the LASSO regression method, it was concluded that the EQU and GDP variables affected the most in the ROA model, the EQU and GDP variables affected the most in the ROE model, and the SP the least.

In the LASSO regression method, the variables that reduce the significance of the model were examined and it was concluded that the absence of deposit variable in the model for the return on assets and return on equity variables would help in obtaining more meaningful estimations.

When the profitability in banking data is compared with the panel data analyzes in the literature, it is considered that almost all of the independent variables in the model are found to be significant, since no dummy variable is added to reduce the degree of freedom in the LASSO regression method and therefore there is no variable that restricts the acceptance region of the model.

As general evaluation; When the profitability indicators of banks are examined, it is seen that having a strong equity structure, solid asset structure and decrease in NPL have a positive effect on profitability. However, it is observed that the increase in gross domestic product and inflation, which are macroeconomic variables, creates a risk environment and accordingly causes an increase in uncertainty, thus having a negative effect on profitability. In this situation, where profitability is adversely affected due to the decrease in demands as a result of prudence, banks should take action especially considering the problems that may arise in macroeconomic conditions. In this case, it is considered that the operational studies carried out by the internal systems of the banks to protect and move forward the current situation against the risks that may arise will contribute positively to the profitability of the banks.

REFERENCES

- Akgüç, Ö. 2011. Banka finansal tablolarının analizi [Analysis of bank financial statements]. *Arayış Basım ve Yayıncılık*. İstanbul.
- Bankacılık Düzenleme ve Denetleme Kurumu (Banking Regulation And Supervision Agency) Retrieved from <https://www.bddk.org.tr/>
- Çıtak, L., Erdoğan, S. ve Özmerdivanlı, A. 2011. Türk Mevduat bankalarında faiz ve faiz dışı gelirler ve giderlerin kârlılık üzerindeki etkisi [The effect of interest and non-interest incomes and expenses on profitability in Turkish deposit banks]. 15. Finans Sempozyumu, Malatya.
- Derici Y., Çiftçi D., Türker A. 2021. Etkili değişkenlerin cezalı regresyon yöntemleri ile belirlenmesi: Diyabet veri kümesi üzerine bir uygulama [Determination of effective variables with penalized regression methods: An application on diabetes dataset], Mersin University
- Fonti V, Belitser E. 2017. Feature Selection Using Lasso. Research Paper In Business Analytics.(https://beta.vu.nl/nl/Images/werkstuk-fonti_tcm235-836234.pdf.)
- İbiş, C., Çatıkkaş, Ö. V. Ç., & Çelikdemir, N. 2018. Banka Muhasebesi İlkeler ve Uygulamalar [Bank Accounting Principles and Practices]. *İstanbul: Türkiye Bankalar Birliği Yayınları*.
- İskenderoğlu, Ö., Karadeniz, E., Atioğlu E. 2012. Türk Bankacılık Sektöründe Büyüme, Büyüklük ve Sermaye Yapısı Kararlarının Karlılığa Etkisinin Analizi [Analysis of the Effects of Growth, Size and Capital Structure Decisions on Profitability in the Turkish Banking Sector], *Eskişehir Osmangazi Üniversitesi İ.İ.B.F. Dergisi*, Nisan 7(1), 291-311
- Kayanan M, Wijekoon P. 2019. Performance of Lasso And Elastic Net Estimators in Misspecified Linear Regression Model. *Ceylon Journal Of Science*.;48(3):293-299.
- Kocakafa T., 2020 Overfitting Retrieved from <https://www.veribilimiokulu.com/overfitting/> Erişim tarihi:24.06.2021
- Küçük, A. (2019). Doğrusal regresyonda Ridge, Liu ve LASSO tahmin edicileri üzerine bir çalışma [A study on Ridge, Liu and LASSO estimators in linear regression].
- Sarıtaş, H., Kangallı Uyar, S. G., & Gökçe, A. (2016). Banka karlılığı ile finansal oranlar ve makroekonomik değişkenler arasındaki ilişkilerin sistem dinamik panel veri modeli ile analizi: Türkiye araştırması. [The Analysis of Banks' Profitability with Financial Ratios and Macroeconomic Variables Based on

System Dynamic Panel Data Model: Research on Turkey]. *Eskişehir Osmangazi Üniversitesi İİBF Dergisi*, Nisan 2016, 11(1), 87- 108

Soundness during the Indonesian Banking Crisis. *Review of Quantitative Finance and Accounting*, 19: 247–260.

Şıklar, İ. 2004. Finansal Ekonomi [Financial Economics], Eskişehir, Anadolu Üniversitesi Açıköğretim Fakültesi Yayını.

Takan, M. 2001. Bankacılık teori uygulama ve yönetim [Banking theory application and management]. Adana: Nobel Yayın

TCMB Elektronik Veri Dağıtım Sistemi Retrieved from <http://evds.tcmb.gov.tr/>

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288.

Trading Economics Retrieved from <https://tradingeconomics.com/turkey/gdp-growth> Erişim tarihi:2.06.2021

Tunay, B. 2005. Finansal sistem: yapısı, işleyişi, yönetimi ve ekonomisi [Financial system: structure, functioning, management and economy]. İstanbul : Birsen Yayınevi.

Türkiye Bankalar Birliği (The Banks Association of Turkey) Retrieved from <http://www.tbb.org.tr>

Verbeek, M. 2004, A guide to modern econometrics (2. Baskı). England: John Wiley & Sons, Ltd.

Yazıcı, Mehmet. 2017. Bankacılığa Giriş [Introduction to Banking]

Zou, H. ve T., Hastie. (2016)."Regularization and Variable Selection via The Elastic Net", *Journal of Royal Statistical Society Series B*, 2005/67, (2), ss. 301-320.