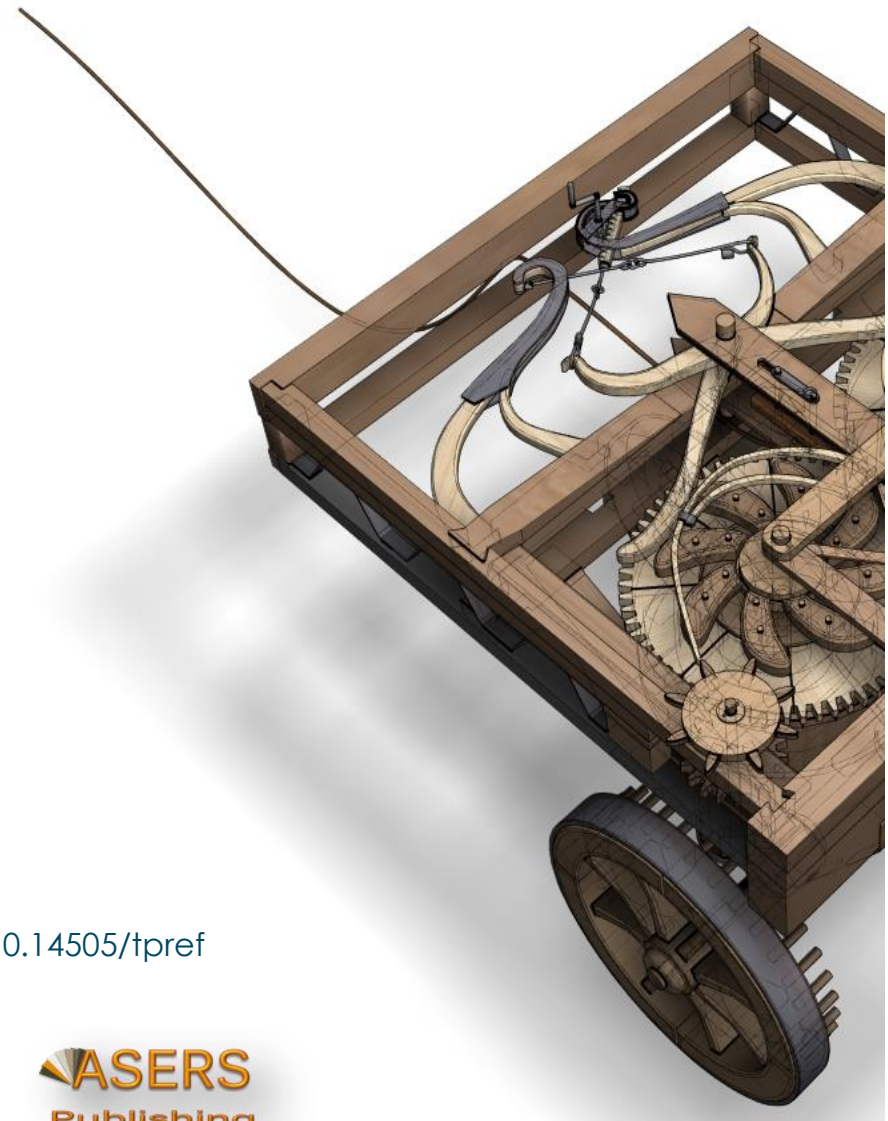
Theoretical and Practical Research in Economic Fields



Quarterly
Volume XV
Issue 4(32)
Winter 2024

ISSN: 2068 – 7710
Journal DOI: https://doi.org/10.14505/tpref

ASERS
Publishing

Volume XV
Issue 4(32)
Fall 2024

ASERS Publishing
ISSN 2068 – 7710
Journal's Issue DOI:
https://doi.org/10.14505/tpref.v15.3(31).00

Table of Contents

Volume XV
Issue 3(31)
Fall 2024

ASERS Publishing
http://www.asers.eu/asers-publishing
ISSN 2068 – 7710
Journal's Issue DOI:
https://doi.org/10.14505/tpref.v15.3(31).00

Call for Papers
Spring Issue
Theoretical and Practical Research in Economic Fields

Many economists today are concerned by the proliferation of journals and the concomitant labyrinth of research to be conquered in order to reach the specific information they require. To combat this tendency, **Theoretical and Practical Research in Economic Fields** has been conceived and designed outside the realm of the traditional economics journal. It consists of concise communications that provide a means of rapid and efficient dissemination of new results, models, and methods in all fields of economic research.

Theoretical and Practical Research in Economic Fields publishes original articles in all branches of economics – theoretical and practical, abstract, and applied, providing wide-ranging coverage across the subject area.

Journal promotes research that aim at the unification of the theoretical-quantitative and the empirical-quantitative approach to economic problems and that are penetrated by constructive and rigorous thinking. It explores a unique range of topics from the frontier of theoretical developments in many new and important areas, to research on current and applied economic problems, to methodologically innovative, theoretical, and applied studies in economics. The interaction between practical work and economic policy is an important feature of the journal.

Theoretical and Practical Research in Economic Fields is indexed in SCOPUS, RePEC, ProQuest, Cabell Directories and CEEOL databases.

The primary aim of the Journal has been and remains the provision of a forum for the dissemination of a variety of international issues, practical research, and other matters of interest to researchers and practitioners in a diversity of subject areas linked to the broad theme of economic sciences.

At the same time, the journal encourages the interdisciplinary approach within the economic sciences, this being a challenge for all researchers.

The advisory board of the journal includes distinguished scholars who have fruitfully straddled disciplinary boundaries in their academic research.

All the papers will be first considered by the Editors for general relevance, originality, and significance. If accepted for review, papers will then be subject to double blind peer review.

Deadline for submission of proposals:	10th February 2024
Expected publication date:	30th March 2024
Website:	http://journals.aserspublishing.eu/tpref
E-mail:	tpref@aserspublishing.eu

To prepare your paper for submission, please see full author guidelines in the following file: https://journals.aserspublishing.eu/tpref/Template_for_Authors_TPREF.docx on our site.

Volume XV, Issue 4(32), Winter 2024

DOI: https://doi.org/10.14505/tpref.v15.4(32).16

Return on Equity in Albanian Banks: A Data-Driven Analysis Using XGBoost

Olsi XHOXHI
Department of Economics, Faculty of Economy
University "Ismail Qemali", Vlore, Albania
ORCID: 0009-0002-1138-1331
olsi.xhoxhi@univlora.edu.al

Grigor DEDE
Department of Finance and Accounting, Faculty of Economy
University "Ismail Qemali", Vlore, Albania
ORCID: 0009-0009-0012-6862
grigor.dede@univlora.edu.al

Zamira SINAJ
Department of Business, Faculty of Economy
University "Ismail Qemali", Vlore, Albania
ORCID: 0000-0003-2231-6842
zamira.sinaj@univlora.edu.al

Article info: Received 20 August 2024; Received in revised form 18 September 2024; Accepted for publication 8 November 2024; Published 30 December 2024.

Abstract: This paper estimates the determinants of bank profitability in Albania with respect to return on equity. This study will employ annual reports from eleven banks in Albania covering the period 2014 to 2023, supplemented by extra sources from the Albanian Association of Banks and the Bank of Albania. Data analyzed at a great level of detail using Python. It evaluates an integrated conceptual framework of the influence of financial metrics, institutional infrastructure, and customer service channels on RoE. An XGBoost regression model is applied to predict RoE since this algorithm has advanced ensemble learning capabilities in refining predictions by iterative error corrections. The model is fine-tuned for Mean Squared Error minimization and performs very well on the training data, as evidenced by an R^2 score of 0.9999. The model slightly overfits, as can be indicated by the lower performance metrics on test data: MSE of 31.8768 and an R^2 score of 0.6024. Concretely, number of outlets, investments in securities, and debit card issuance are the variables most influencing RoE, as can be inferred from the importance analysis. This paper contributes to the literature with useful insights into financial performance and operational dynamics in Albanian banks, evidencing variables that play a significant role in profitability and, therefore, supply scope for further research and model improvement.

Keywords: bank profitability; feature importance; machine learning; XGBoost.

JEL Classification: E44; G21; G28; C10.

Introduction

Forecasts of bank profitability are very important in any effective management of the financial institution and strategic planning, more so for those banks with specialized services or niche markets. Accurate profitability forecasts can help such institutions greatly make relevant business decisions, manage their resources efficiently, and develop a competitive edge over other banks. Traditional ways used in making predictions hardly capture the complex patterns and interactions that exist in financial data, hence less accurate results.

In the sphere of artificial intelligence, machine learning offers advanced techniques for the above-mentioned problems. Using algorithms and statistical models that learn from experience, machine learning offers superior predictive capabilities. More importantly, one of the main reasons that makes XGBoost an important technique is its capacity for dealing with large datasets and creating complex patterns; hence, in bank profitability

forecasting, it becomes very effective. The following paper presents the application of XGBoost in predicting profitability for specialty banks. Doing so, it fills the literature gap by proving its effectiveness in this context. The objective of this study is to provide novel insights into profitability forecasting by elaborating on model adaptability to changing market conditions and making more accurate predictions.

1. Literature Review

This chapter reviews the existing literature on the banking sector, highlighting the key research findings and theoretical perspectives that shape our understanding of its dynamics. We will explore major contributions to the field and identify gaps that our study aims to address.

Menicucci and Paolucci (2016) in their research have used panel data estimation with a representative sample of 28 large European banks between 2006 to 2015 to evaluate the effect of internal bank-specific factors on European bank profitability. The authors further support that bank size, being measured by the value of total assets, is the most important driver of profitability. This shows that with scales of operation, bigger banks have economies of scale and product diversification. Besides this, the study has also found from the resultants that in relation to profitability, asset quality exerts negativity, while the deposit ratio of the bank shows positive and impactful results for ROE. The capital strength of banks through the equity/total assets ratio secures profitability, in the sense that high equity reduces the costs of external financing, but net loans to total assets show no significant signs with accounting for profitability variations. The results offer rich insights for bank management, regulators, policy makers, shareholders, and other stakeholders on how stability and competition can be better preserved within the European banking market. The study also provides directions relevant for future investigation of the problem, in particular, the analysis of other factors, such as taxation or service quality, and an exploration into differences between banks in the different size and profitability groups.

In their study, Titko *et al.* extend the debate on performance management and performance drivers of the banking sector within the Baltic States. In their findings, it is noted that although Lithuanian data supported a positive relationship between bank size and its degree of profitability, Latvian data did not support such a hypothesis. Their findings from the research on the association between operational efficiency and bank profitability were very mixed for Lithuania and rather inconclusive for Latvia. The study also does not find any support for the notion that the advanced level of infrastructure and e-banking services is actually an explanation of high profitability, at least in the case of Latvia. Data was limited, and statistics from the local authorities were not consistent with the statistics provided by the European Central Bank. This, therefore, made the authors recommend that future analyses be based on individual banks, mostly in small banking sectors, with an extension of the research period if more concrete results are to be expected.

According to Aspal *et al.* (2019), Capital Adequacy is another internal factor unique to banks, apart from the asset factor. Asset Quality (AQ), Management Efficiency (ME), Earnings Quality (EQ), and liquidity are some of the critical factors of differences in financial performance by private sector banks in India. Findings demonstrated the negative relationship of Capital Adequacy with bank performance that is statistically non-significant. Asset quality was negatively and significantly related to profitability. This means that, if an asset quality lacks or a high level of non-performing assets exists, it will detract from high performance. Management efficiency, as indicated in the expenditure-to-income ratio, negatively relates to profitability to the bare minimum, meaning that it is very important for the bank to cut back on its expenditures in order to maximize the profits. On the other hand, it is noteworthy that Earnings Quality and Liquidity management had an overall positive and significant effect on the financial performance: banks with good earnings quality and good liquidity management are likely to achieve a more significant profit. Lastly, macroeconomic factors like GDP and inflation perform negatively related to bank performance, but their overall impact on the private banking sector financial performance was considered inconclusive.

Mashamba and Chikutuma (2023) identify the drivers of bank profitability in Zimbabwe, one of the countries that has experienced a myriad of economic challenge over the last ten years. Based on panel data of 11 commercial banks from 2011 to 2020, they found out that bank-specific variables such as non-interest income, liquidity, cost efficiency, capital adequacy, and bank stability - worthwhile and turning negative - all drive profitability to levels of statistical significance. The industry factor of bank concentration, however, has a negative effect on profitability. Surprisingly, though, the macroeconomic variables - GDP and inflation - are high in Zimbabwe, with very low growth rates and high inflation, and yet they do not seem to affect bank profitability. In addition, it appears that increasing regulatory capital reduces the positive impact of bank stability on profitability, suggesting that the possibility for risk-taking, and therefore profits, is reduced with higher capital requirements for banks. Other findings also include that there is no significant moderating effect of FinTech on bank performance,

thus proving that the impact of FinTech on the banks' competitiveness in non-interest income activities is limited. In its entirety, tha auhors draw a conclusion that it is the internal factors, controlled by bank managers and regulators, that are really important for bank profitability, not the external factors related to economic factors. They, therefore, provide policy recommendations toward the improvement of bank profitability and the strengthening of the banking sector in Zimbabwe.

According to Muchtar *et al.* (2021), the study investigates the risk determinants of banks in ASEAN countries moderated by bank scale, whereby Random Effects Model will be used to undertake analysis that involves panel data. Their study is on banks in ASEAN-4 and also Indonesia, that is, Indonesia, Malaysia, Philippines, and Thailand, examining such risk measures as Standard Deviation (STD) and Value at Risk (VaR). The results further indicate that, although of the ASEAN-4, the CAR is negatively related to STD, it has found no association with Bank Risk in Indonesia. Meanwhile, the NPLs do not often alter Bank Risk for ASEAN-4, but do have a significant effect on VaR in Indonesia. The LDR results in a negative effect on VaR in both regions. MP has a negative effect on STD and a positive effect on VaR in the ASEAN-4. However, it has no effect in Indonesia. Exchange Rate has a negative impact on STD in the ASEAN-4, and it has a positive effect in Indonesia. Interest Rates do not have an influence on neither STD nor VaR in either of the regions. TI has a negative influence on STD but has a positive effect on VaR in the ASEAN-4. In Indonesia, it negatively influences Bank Risk. Finally, bank scale or BUKU Bank (BB) has a negative impact on STD in ASEAN-4 and has no significant impact on Bank Risk in Indonesia.

In the research, Keka *et al.* (2023) have highlighted what determines the profitability of commercial banks in Kosovo and Albania. According to them, this research is very important for policy makers, regulators, and bank management. Determinants considered in this study include the number of employees, interest rates on loans, non-performing loans, total loans, among others. In carrying out this study, quarterly secondary data from the years 2010-2020 were sourced, totalling 400 observations. The research employs multiple linear regression and ordinary least square analysis so as to establish any kind of interactions between dependent and independent variables that would help in establishing what drives bank profitability. All the statistical analyses were done with the aid of specialized software such as STATA and SPSS. Results show that loan interest rate, total loans, and non-performing loans are the major variables affecting profitability, while total loans and number of employees are the major determinants of ROE. These findings, according to Keka et al., add some valuable lessons for fine-tuning bank management and policy making toward improved profitability and stability.

According to Chand *et al.* (2024), while exploring the interplay of globalization, institutional quality, and global uncertainties - the COVID-19 pandemic - in the banking sector of Fiji during the period of 2000 to 2021, all of the above affect the profitability of banks in small island economies. They have pointed out that bank-specific indicators are not sufficient to understand profitability, and they have established, for example, that while elements such as net interest margin and capital adequacy have a positive effect on performance, other challenges such as non-performing loans and global crises are factors that diminish it. Their findings favor a strong regulatory environment and political stability as ways of enhancing the resilience of the financial sector and propose exploring opportunities in remittances and economic growth. They also embrace the unique challenges that small economies like Fiji face and put emphasis on the need for an integrated approach to profitability and risk analysis. Notwithstanding the limitations of a small sample size and only being confined to the deposit and lending institutions, the study provides significant insights and areas that can be used as extensions in future studies to fill data gaps.

The relative value importance indicator obtained from the random forest model, Almaskati (2022) use it to obtain an assessment of comparative significance with respect to the various determinants affecting bank risk and profitability. From this study, they conclude that bank-specific factors are the main drivers of profitability, whereas country factors have a more significant impact on determining risk. They identify market power and size as key in profiling both profitability and risk. It also depicts the high contribution of a nation's financial development and regulatory quality in driving a bank's risk. This analysis also shows that the risk profile of a bank is driven by a number of variables with almost equal intensity while profitability is driven by very few dominant variables while others are relatively insignificant. Further research in the future into how sudden or structural changes in values of these key variables affect banks' profitability and risk should bring out insights into their effect and help regulators and policymakers zero in on what is most important for stability in the banking system.

In their empirical study, Mirovic *et al.* (2024) regarded the literature gap with the determinants of bank profitability in the Eurozone through an analysis using quarterly data from 2015 to 2020. Their study was conducted by descriptive statistics, panel unit root tests, and cross-section dependence tests with the application of static and dynamic panel models like pooled least squares, random effects, fixed effects, and generalized

method of moments. The Hausman test showed that the random-effects model was appropriate for the estimation of the determinants of profitability. In their result, they found that both bank-specific and macroeconomic factors had significant effects on profitability: non-performing loans and a cost-to-income ratio increase reduced ROA and ROE, while net interest margins, net income from fees and commissions, and trading assets boosted these measures of profitability. Moreover, ROA and ROE improved with enhanced GDP growth and were eroded by rising inflation, unemployment, and debt. Their findings are consistent with prior studies and underline the fact that profitability is required for shareholders' returns, customers' benefits, financial stability, regulatory compliance, and investor confidence. In this regard, this paper contributes to literature since most of the studies do not use all these macroeconomic variables. One of the known limitations is not considering economic cycles in the analysis, and further research is suggested in evaluating the consistency across different periods in an economic cycle: pre-pandemic crisis, during, and post-pandemic crisis.

In the case when studying the profitability of the Macedonian banking system during the period from 2007 to 2022, Kosumi and Xharku (2024) developed a model where the return on assets stood as a dependent variable. According to their results, factors such as the size of the banking sector, credit risk, liquidity, income diversification, and non-performing loans are the negative drivers of profitability. On the other side, capital adequacy, operational efficiency, GDP, and interest rates are positive determinants for the same. In detail, Macedonian banks should focus on asset management to improve profitability, increase non-interest income in order to avoid credit risk and non-performing loans, and maintain favorable liquidity ratios. Second-tier banks should adopt international expansion, improvement in the management of the loan portfolio, and investing in technology for improved crisis management. Diversifying income sources, hence boosting non-interest revenue, remains another way through which the banks could survive the crisis and strategically navigate the legal environment. This may serve as valuable information for policymakers, regulators, and bank management within North Macedonia. Future research will focus on covering a larger set of economic, legal, industry, and bank-specific factors that explain the profitability of banks.

According to Raiter (2021), most financial institutions are challenged in dealing with a wide variety of risks that affect it, especially on credit risk. Most banks do not really understand the risks of an individual loan or investment and even the total credit risk inherent in their portfolios. Credit risk is still the most crucial financial risk for most institutions. Effective risk management would, therefore, seek to keep the potential impact of these hazards within acceptable levels across society, including risks associated with the environment, technology, people, organizations, and politics. This paper examined the determinants of credit risk by commercial banks worldwide based on WDI and Bankscope data on 106 commercial banks worldwide. This analysis brought out that, in addition to currency rates and regulatory capital, variables like inflation, interest rate, and unemployment rates have a very strong impact on credit risk - one would be considering that a high value of these variables implies a high credit risk for banks. Conversely, bank efficiency, bank size, and the GDP growth rate seemed to bear a strong negative impact on credit risk, showing that credit risk decreases when times are economically good. It was also found that the private sector banks, in general, exhibited lower credit risk as compared to the public sector banks.

In their paper, Islam and Rana (2019) investigate the various bank-specific and macroeconomic factors impacting bank profitability by taking data from 23 commercial banks in Bangladesh for the period 2013-2017. The data has been obtained from annual reports of the respective individual banks, Bangladesh Bureau of Statistics, and different publications of the Bangladesh Bank. They hold their regression analysis through a fixed effect panel data model. In the case of this study, three profitability measures will be used: ROA, ROE, and NIM. For ROA, the results indicate that earning variables, including TIN and NII, and structure of assets, DPST, positively affect ROA, while the quality of assets, NPL, negatively affects it. For ROE, earning variables, TIN and NII, and strength of capital, CAP, all have a significant positive effect, while the quality of assets has a remarkable negative effect. In the case of NIM, earning variables are significantly related to capital strength and liquidity. It can be seen that among the macro variables, the GDP growth rate, inflation rate, and interest rate are found to be insignificant for influencing profitability. These findings would enable investors, policymakers, management, and other stakeholders to make informed decisions for future improvisations in the performance of the financial organizations.

Aktas *et al.* (2015) point out that banks normally maintain capital in excess of the legal requirement, thus implying that there could be other factors besides the regulatory requirement for capital that influence their capital structure. Their paper estimated two models for explaining the determinants of capital adequacy ratio in the Southeast European region. Model 1 includes bank-specific variables such as size, profitability measured by ROA, leverage, liquidity, net interest margin, and risk. The second model includes a considerable number of

environmental factors, such as the economic growth rate, inflation, real interest rate, stock market volatility of the Eurozone, deposit insurance coverage, and governance indicators. They have evidence from the analysis using Feasible GLS that size, ROA, leverage, liquidity, NIM, and risk significantly affect CAR, with the different signs indicating whether it is a positive or negative effect. The other factors playing important roles are environmental factors of economic growth rate, stock market volatility in the Eurozone, deposit insurance, and governance. This study consequently reveals that determinants of CAR in developed economies can be relevant in explaining CAR for banks within the SEE region, which, with the exception of Greece, are transition economies undergoing various economic and banking reforms.

Assfaw (2020) have estimated the determinant variables for capital structure in private commercial banks of Ethiopia based on data from 2010–2018. They ran a Clustered Robust random effect regression model which portrayed that on average, bank assets were financed by debt in an average percentage share of 86.53%. With an increase in bank size, there was a rise in leverage measures hence supporting the trade-off theory but contradicting the pecking order theory. In the second model, profitability negatively influenced the levels of debt, consistent with the pecking order theory but contrary to the trade-off theory. The tax variable positively influenced leverage, thus agreeing with the static trade-off theory. Tangibility of assets had a negative effect on leverage and thus conflicted with both the agency and the static trade-off theories. Earnings volatility positively affected leverage consistent with pecking order theory but not with trade-off theory. Growth opportunities, real GDP, and inflation had a minimal impact. The study, therefore, is to advise Ethiopian banks to put into consideration both factors, that is, microeconomic and macroeconomic factors in optimizing their capital structure and point out areas for future research.

Serwadda (2018) attempted to gauge the influence of intrinsic drivers on the profitability of commercial banks in Hungary from the period 2000–2015. In doing so, they applied a balanced panel data set of 26 banks with 416 observations, where they checked for profitability using return on average assets as the dependent variable against explanatory variables such as non-performing loans, overhead costs, bank size, net interest margin, liquidity risk, and capital adequacy ratio. Their findings showed that bank-specific factors have a great influence on the profitability of a bank, especially bank size and asset quality. On the contrary, those which weighed negatively in determining profit included non-performing loans, high overhead costs, and liquidity levels. Net interest margin and capital adequacy ratio had a positive but insignificant effect, and the recommendations were for banks to pay more attention to reducing and managing overhead costs, enhancing asset quality, and optimizing liquidity levels in order to realize higher profitability. It also recommended that bank managers keep tight reins on credit and liquidity risks, as well as strive to diversify income against the backdrop of effective cost controls.

2. Research Methodology

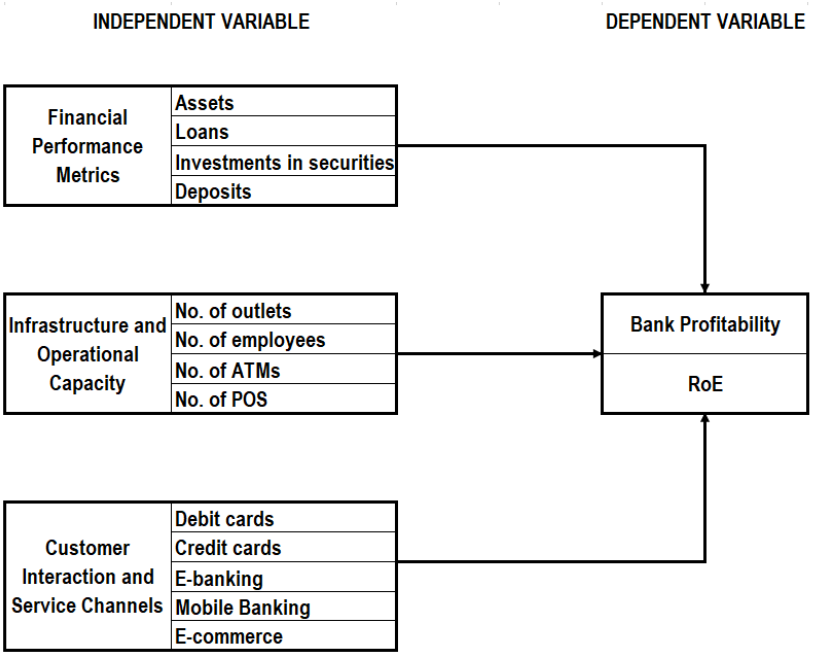
2.1. Data Source and Analysis

In this study, data were acquired from annual reports of eleven Albanian banks, covering the period 2014 – 2023. These banks include Banka Amerikane e Investimeve, Banka Kombëtare Tregtare, Credins Bank, Fibank Albania, Intesa Sanpaolo Bank of Albania, ProCredit Bank, Raiffeisen Bank Shqipëri, OTP Bank Albania, Banka e Tiranes, Union Bank, and Banka e Bashkuar e Shqipërisë. Downloaded directly from the official websites of the banks and supplemented by data from the Albanian Association of Banks and the Bank of Albania, these reports were subjected to analysis by means of Python, a very versatile and powerful programming language, oriented to the area of data treatment and statistical analysis.

2.2. Conceptual Framework

The purpose of this research is to evaluate Return on Equity's complex relation with other independent variables in order to find the driving factors of bank profitability. This analysis is based on the assumption that financial performance measures, such as assets, loans, investments in securities, and deposits, make major contributions to RoE while reflecting the bank's operational efficiency and market position. It is expected that the infrastructure and operational capacity variables will influence RoE, as such variables outlets, employees, ATMs, and POS terminals are bound to condition the bank's ability to effectively manage and leverage the bank's resources. The variables of customer interaction and service channels are also expected to have an influence on RoE, since these are related to debit cards, credit cards, e-banking, mobile banking, and e-commerce, which deepen relationships and increase revenue sources. This is visualized in the conceptual model of this study, which lays out these relationships and follows through with a clear depiction of how each of these variables contributes to the overall profitability of banking institutions.

Figure 1. Conceptual framework



Source: Author's calculations

2.3. Econometric Model Framework

For this research, we will use an XGBoost model since it is one of the best and most efficient algorithms for regression tasks. Since it is a boosting algorithm, XGBoost combines many weak models in order to develop a strong model for prediction. It makes iterated predictions for continuous variables by minimizing the loss function, and therefore in regression problems, it gives continuous variable predictions. Base learners are decision trees that are refined sequentially to correct errors made by previous models; therefore, a highly accurate model results from key ensemble learning principles.

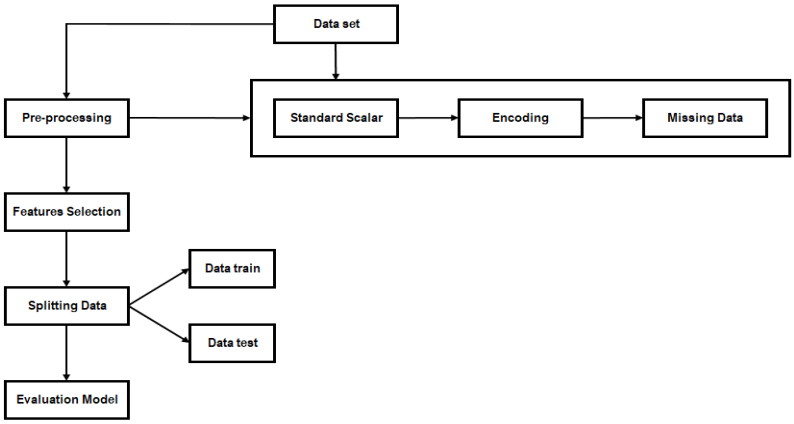
In the regression tasks, XGBoost is ready to minimize many different loss functions: Mean Squared Error and Mean Absolute Error; Huber Loss; Quantile Loss; Log-Cosh Loss; Poisson Loss. In this research, MSE will be used since it is a measure of average squared differences between predicted and actual values.

The training process initializes the model with basic predictions, and then decision trees are added sequentially, updating the predictions based on what each tree contributes. The other core hyperparameter is the learning rate, which tells by how much each new tree shall eventually impact the final prediction, the smaller the rate, the greater the number of trees needed for optimum performance. Careful data preparation is necessary when incorporating the XGBoost regression in a framework of an econometric model. All the features must be pre-engineered in such a way that they represent correctly the economic relationships that are to be studied. Normalization/standardization of features might end up as necessary to be more efficient for the algorithm. Parameter Tuning: Grid or random search methods are used to find the best combination of hyperparameters, which can be used in the XGBoost model. Train and test sets have to be validated to prevent overfitting situations and also to allow validation independent of train sets.

Metrics to use for assessing goodness of fit of the XGBoost regression model include MSE, MAE, or R-squared. Besides, XGBoost also provides feature importance estimates, explaining which variables are most influential for the predictions.

In Figure 2, the process of data preparation and model evaluation is illustrated. This figure outlines the key stages involved, beginning with the dataset and encompassing pre-processing steps such as standard scaling, encoding, and handling missing data. It then progresses through feature selection, data training, and splitting the data into training and test sets. Finally, the figure includes the model evaluation phase, providing a comprehensive view of the entire workflow.

Figure 2. Data Preparation and Model Evaluation Process



Source: Author's calculations

3. Research Results

The current chapter expounds on the financial performance, institutional infrastructure, and service capabilities of the institutions under study. The key financial metrics are first analyzed, followed by an overview of the infrastructure and operational capacity, and finally the analysis of customer interaction and service channels. The following three tables present information and its interpretation to provide nuanced insight into the performance of these institutions, their operational scale, and capacity to reach clients through the various service channels. This will be insightful in assessing strengths of the institutions and in identifying areas requiring improvement.

In Table 1, the financial performance indicators of these institutions are presented, which reveal a very high level of variability in assets, loans, security investments, and deposits. Mean values are ALL92.95 billion for assets, ALL34.28 billion for loans, ALL26.20 billion for investments in securities, and ALL76.67 billion for deposits. The extreme range of the asset values from ALL21,141 to ALL491.47 billion, and the variation of loans and investments, give a good example on the different scales and financial strategies the institutions employ. The wide range of deposits also portrays the disparity in financial capacity and deployment of resources by the banks.

Table 1. Descriptive Statistics for Financial Performance Metrics

	Assets	Loans	Investments in securities	Deposits
count	105	105	105	105
mean	92,946,330	34,276,150	26,204,990	76,670,990
std	124,233,200	39,937,910	37,941,650	103,438,500
min	21,141	16,847	0	34,235
25%	1,030,211	334,541	138,967	899,849
50%	38,216,890	22,157,380	7,540,288	27,022,120
75%	130,819,000	53,002,250	39,545,180	107,079,100
max	491,473,200	138,066,000	160,292,100	415,973,100

Source: Author's calculations

Table 2. Descriptive Statistics for Infrastructure and Operational Capacity

	No. of outlets	No. of employees	No. of ATMs	No. of POS
count	105	105	105	105
mean	37	571	69	1,021
std	25	400	55	1,933
min	5	74	0	0
25%	14	252	30	0
50%	32	447	59	213
75%	57	857	74	992
max	93	1,406	230	9,631

Source: Author's calculations

Table 2 shows the institutional infrastructure and capacity to handle activities. The average bank operates 37 outlets staffed by an average of 571 people, manages 69 ATMs as well as 1,021 POS terminals. The data show high variation in particular with regard to the number of employees and the number of POS terminals, from 74 employees to 1,406, and from no POS terminals at all to 9,631. This variability may hint at differences in scale of operations and investments in technology that have consequences for the ability to deliver services and manage customer interactions appropriately.

Table 3 outlines the metrics for customer interaction and service channels, with a focus on debit and credit cards. The average institution that issues debit cards are 95,514, while that of credit cards is 9,898. The debit card issuance ranges from 0 to 364,702 and that of credit card numbers from 0 to 73,719 show various strategies in customer engagement and provision of service. It is the very wide range of these figures that highlights a very different way of dealing with customer relationships and underlines the fact that individual customer relations strategies are an important means to meeting customer needs.

Table 3. Descriptive Statistics for Customer Interaction and Service Channels

	Debit cards	Credit cards
count	105	105
mean	95,514	9,898
std	101,002	16,253
min	0	0
25%	26,372	668
50%	60,504	3,814
75%	121,180	6,513
max	364,702	73,719

Source: Author's calculations

In the table below, we have results for both training and testing. According to this result of the training set, the accuracy of the XGBoost regression model will be at a high rate. The Mean Squared Error values are 0.045172, with a Mean Absolute Error of 0.139053. It is shown by these metrics that the average error which the model makes in the prediction of training data is low. The mean squared error brings out fine-squared differences between the predicted and true values, while the mean absolute error indicates a small average absolute error. The R^2 Score was very high at 0.999915, as was also the Explained Variance Score at 0.999915; this establishes that the model explains almost all variability in the training data and is a very good fit.

On the test set, model performance is such that it looks very promising, but there is an observable difference against the training set. The MSE is 31.876789, with an MAE of 4.185301, considerably higher than what was observed in the training set. This could be an indication that predictions on unseen data are not very accurate; probably a case of overfitting, where the model goes very well with the training data but poorly on new, unseen examples. The R^2 Score of 0.602384 and the Explained Variance Score of 0.779401 already show that generalizability to new data can be much improved, although it captures a fair share of the variance in the test data.

Table 4. Model Metrics

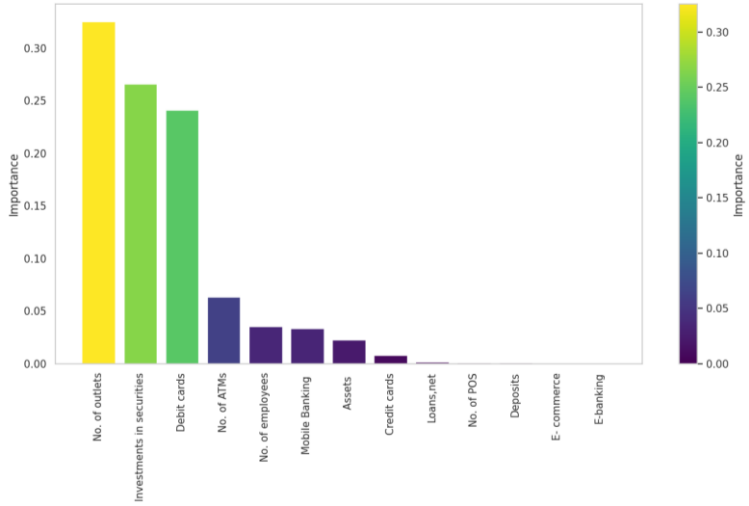
Metrics	Train	Test
MSE	0.0452	31.8768
MAE	0.1391	4.1853
RMSE	0.2125	5.6460
R2 Score	0.9999	0.6024
Explained Variance Score	0.9999	0.7794

Source: Author's calculations

According to the important analysis of the features in Figure 3, number of outlets was the most influential variable with respect to predictive power, accounting for about 32.5 percent. Next in a row comes investments in securities, contributing to about 26.6 percent of the predictive capability of the model. Debit cards are also important and represent about 24.1 percent of the importance. Comparatively, lower impacts are by features such as the number of ATMs with 6.3%, followed by the number of employees with 3.5%. Mobile banking, assets, and credit cards contribute even less, of 3.4%, 2.3%, and 0.8%, respectively. Loans, net, contribute 0.15%, number of

POS with 0.12%, deposits alone contribute minimally; e-commerce and e-banking do so less, with e-commerce at 0.0175% and e-banking at 0%. This distribution underlines how physically and financially transaction-related features are most primary in influence on other variables within the model.

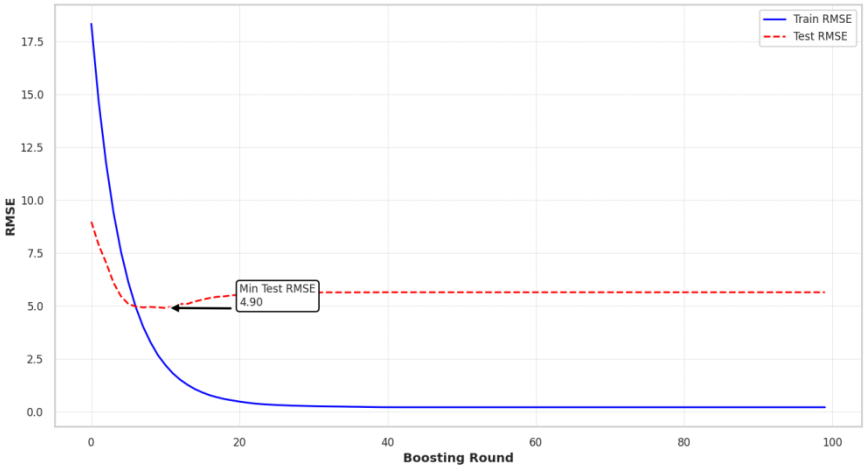
Figure 3. Feature Importances



Source: Author's calculations

Results from your XGBoost model are rather great at improving early on, and RMSE on both validation sets is decreasing rather fast during initial iterations. It learns and fits the data quite well; RMSE values decrease from 18.33297 and 8.98150 at iteration 0 down to 2.66945 and 4.93107 at iteration 9. Progress slows a bit in the middle stages; the RMSE values do decrease but at a slower rate. By iteration 20, the model's performance has already stabilized, most especially in training set, which holds an RMSE to about 0.21254. It indicates convergence, though the model shows some signs of overfitting since the testing set is still higher, which means it worked better on the training data than on the validation set. The model learned data quite well overall, and performance metrics are stable.

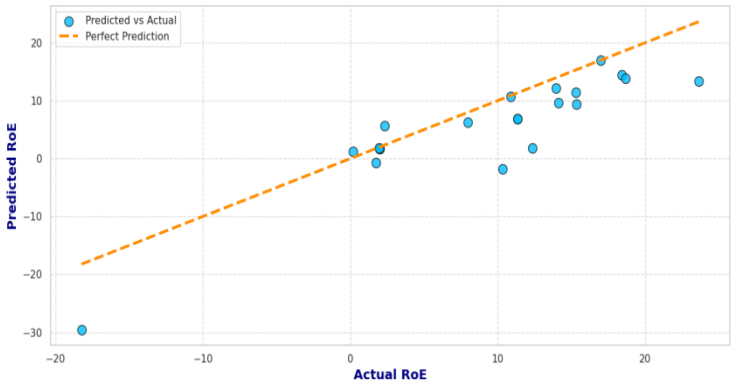
Figure 4. Training and Testing RMSE



Source: Author's calculations

In Figure 5, the comparison between predicted and actual Return on Equity (RoE) for the test set is depicted. This visualization provides insights into the model's prediction accuracy relative to actual RoE outcomes.

Figure 5. Predicted vs Actual Roe (Test Set)



Source: Author's calculations

4. Discussions

This study affords valuable results on the implication of XGBoost in bank profitability prediction, but more specifically, the implications for Albanian banks, which are concentrated in specialized activities. The results demonstrate the potential that machine learning algorithms, especially the XGBoost algorithm, can hold in improving predictions in matters related to profitability over traditional methods of such predictions.

4.1. Implication of the Results

The application of the XGBoost model turned out to have very high predictive accuracy that can be derived from the low Mean Squared Error and Mean Absolute Error, along with a high R-squared score. This clearly indicates that the model has been able to capture the complex relationship between independent variables and Return on Equity, thereby providing reliable forecasts of bank profitability.

This study contributes significantly to showing how XGBoost can be used with large datasets to bring out complex patterns that the traditional models might have missed. The capability of XGBoost to handle vast volumes of data, combined with the inclusion of a variety of facilities - ranging from financial performance indicators to institutional infrastructure, such as customer service channels - clearly demonstrates the strength of this model in the production of accurate predictions. This is of particular relevance to the instance of the specialty banks with their unique offerings and niche markets, in which traditional models are often incapable of accurate forecasts.

This also indicates the importance of inclusion of as wide range of variables as possible in profitability forecasting. It is in allowing such diverse factors as operational ability, customer interface metrics, and financial performance indicators to combine that the XGBoost model can give a picture on what could appear as a determinant of bank profitability. Such a multifaceted approach gives an enriched understanding of how various elements interact with each other and contribute to financial performance, plus assists banks in making more strategic decisions with more informed judgment.

4.2. Implications for Management of Banks

The results indicate that the use of advanced techniques of machine learning, as an example, XGBoost would enable the bank managers and decision-makers to have more precise findings and a view of action into the profitable aspect. In case profitability in future can be predicted precisely, then the banks shall be in a better position to allocate its resources, optimize operations, and develop strategies to maximize a competitive advantage. Moreover, awareness of the relative importance of various factors that influence profitability might allow managers to focus individual efforts on those initiatives that have the potential for having the greatest impact on financial performance.

The paper further resonates with the need for constant data monitoring and model update. Most probably, the factors impacting profitability may change as market conditions and internal dynamics change. This would mean the need for model predictions to respond to minimal considerations. For such a reason, constant updating of the XGBoost model with new data and recalibration of parameters shall keep the forecasts relevant and accurate to the point.

Conclusions and Further Research

Conclusions

The current research paper was able to confirm that the XGBoost model is effective in predicting the profitability of commercial banks, mainly specialization banks in Albania. The better performance of the XGBoost model compared to the other classical forecasting models, tested here in this paper, underlines that the XGBoost model has the potential to be a useful approach for financial institutions that aim to profit from predictions.

The approached in-depth analysis during this study presented several important results:

- High accuracy of XGBoost: The model of XGBoost showed a high share of predictions, relativised by errors, and proved to be considerably high in the R-squared value. This, therefore, proves the adequacy of the model for effective forecasting of bank profitability.
- Diversified variables: Diversifications of the variables included in the model have given a vast view of the profitability determinants, signifying the venture of working in a holistic manner for prediction.
- Value for Bank Management: The ability to come up with good profit projections is of overriding value to bank management in operational decisions and strategic planning.

Further Research

Though this study made major achievements in the profitability forecast, there are various aspects that would need more research:

- Other Machine Learning Algorithms: Further work may investigate what other machine learning algorithms, such as neural networks or support vector machines, would really compare with in terms of the performance of the XGBoost algorithms. This will provide a general view of which algorithms are likely to deliver the best performance and remain useful in the context of forecasting bank profits.
- Longitudinal Analysis: The extension of time of study and the incorporation of more longitudinal data can show changes in the performance of predicting profitability models over different points in time and under different economic conditions. Adaptations of XGBoost in Other Banking Sectors with Recognizable Features: Research on how other banking sectors or different regions adopt XGBoost into their systems may show its flexibility and adaptability.
- Integration of Macro-Economic Factors: Given that you will include in the model other additional independent variables, macroeconomic factors like the economic cycle or worldwide trends in finance can positively make the model more precision and accurate.
- Advancement and its Impact: In order to give the field a more prophetic view, an evaluation of the performance of big data analytics together with the impact of innovations in AI, it's a big plus as far as technology development is concerned.

This study, in short, demonstrates that XGBoost may make a significant difference in the accuracy of profitability forecasts for specialty banks. Further research in some of the indicated areas would clearly be necessary for better refinement in the use of such machine learning techniques, leading to even more robust and actionable insights into banking.

Credit Authorship Contribution Statement

Authors have contributed equally to this research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Use of Generative AI and AI-Assisted Technologies

The authors declare that they have not used generative AI and AI-assisted technologies during the preparation of this work.

References

[1] Aktas, Rafet, Bilge Bakin, and Gokhan Celik. 2015. The Determinants of Banks' Capital Adequacy Ratio: Some Evidence from South Eastern European Countries. *Journal of Economics and Behavioral Studies*, 7 (1(J)): 79–88. DOI: https://doi.org/10.22610/jebs.v7i1(j).565.

[2] Almaskati, Nawaf. 2022. The Determinants of Bank Profitability and Risk: A Random Forest Approach. *Cogent Economics & Finance*, 10 (1): 1–15. DOI: https://doi.org/10.1080/23322039.2021.2021479.

[3] Arbenita Kosumi, and Lutfi Zharku. 2024. Bank-Specific, Economic and Legal Determinants of Profitability in the Republic of North Macedonia. *Access to Justice in Eastern Europe*, 7(2): 212–33. DOI:https://doi.org/10.33327/ajee-18-7.2-a000209.

[4] Assfaw, Abdu Mohammed. 2020. The Determinants of Capital Structure in Ethiopian Private Commercial Banks: A Panel Data Approach. *Journal of Economics, Business, & Accountancy Ventura*, 23 (1). DOI:https://doi.org/10.14414/jebav.v23i1.2223

[5] Chand, Shasnil, Ronald Kumar, Peter Stauvermann, and Muhammad Shahbaz. 2024. Determinants of Bank Profitability: Do Institutions, Globalization, and Global Uncertainty Matter for Banks in Island Economies? The Case of Fiji. *Journal of Risk and Financial Management,* 17(6): 218–40. DOI:https://doi.org/10.3390/jrfm17060218

[6] Elisa, Menicucci, and Paolucci Guido. 2016. Factors Affecting Bank Profitability in Europe: An Empirical Investigation. *African Journal of Business Management*, 10 (17): 410–20. DOI:https://doi.org/10.5897/ajbm2016.8081

[7] Islam, Shafiqul, and Md Shohel Rana. 2019. Determinants of Bank Profitability: Evidence from Commercial Banks of Bangladesh. *Journal of Asian Business Strategy*, 9 (2): 174–83. DOI:https://doi.org/10.18488/journal.1006.2019.92.174.183.

[8] Kumar Aspal, Parvesh, Sanjeev Dhawan, and Afroze Nazneen. 2019. International Journal of Economics and Financial Issues Significance of Bank Specific and Macroeconomic Determinants on Performance of Indian Private Sector Banks. *International Journal of Economics and Financial Issues,* 9 (2): 168–74. DOI:https://doi.org/10.32479/ijefi.7727.

[9] Mirović, Vera, Branimir Kalaš, Nada Milenković, Jelena Andrašić, and Miloš Đaković. 2024. Modelling Profitability Determinants in the Banking Sector: The Case of the Eurozone. *Mathematics*, 12 (6): 897. DOI:https://doi.org/10.3390/math12060897.

[10] Muchtar, Susy, Adler Haymans Manurung, and Farah Margaretha. 2021. Determinant of Bank Risk with Bank Scale as Moderating Variable in ASEAN. *International Journal of Science and Society*, 3 (3): 150–69. DOI: https://doi.org/10.54783/ijsoc.v3i3.362

[11] Qehaja-Keka, Vese, Skender Ahmeti, and Muhamet Aliu. 2023. Bank Profitability Determinants: Evidence from Kosovo and Albania. *Journal of Liberty and International Affairs*, 9 (2): 297-11. DOI:https://doi.org/10.47305/JLIA2392370qk

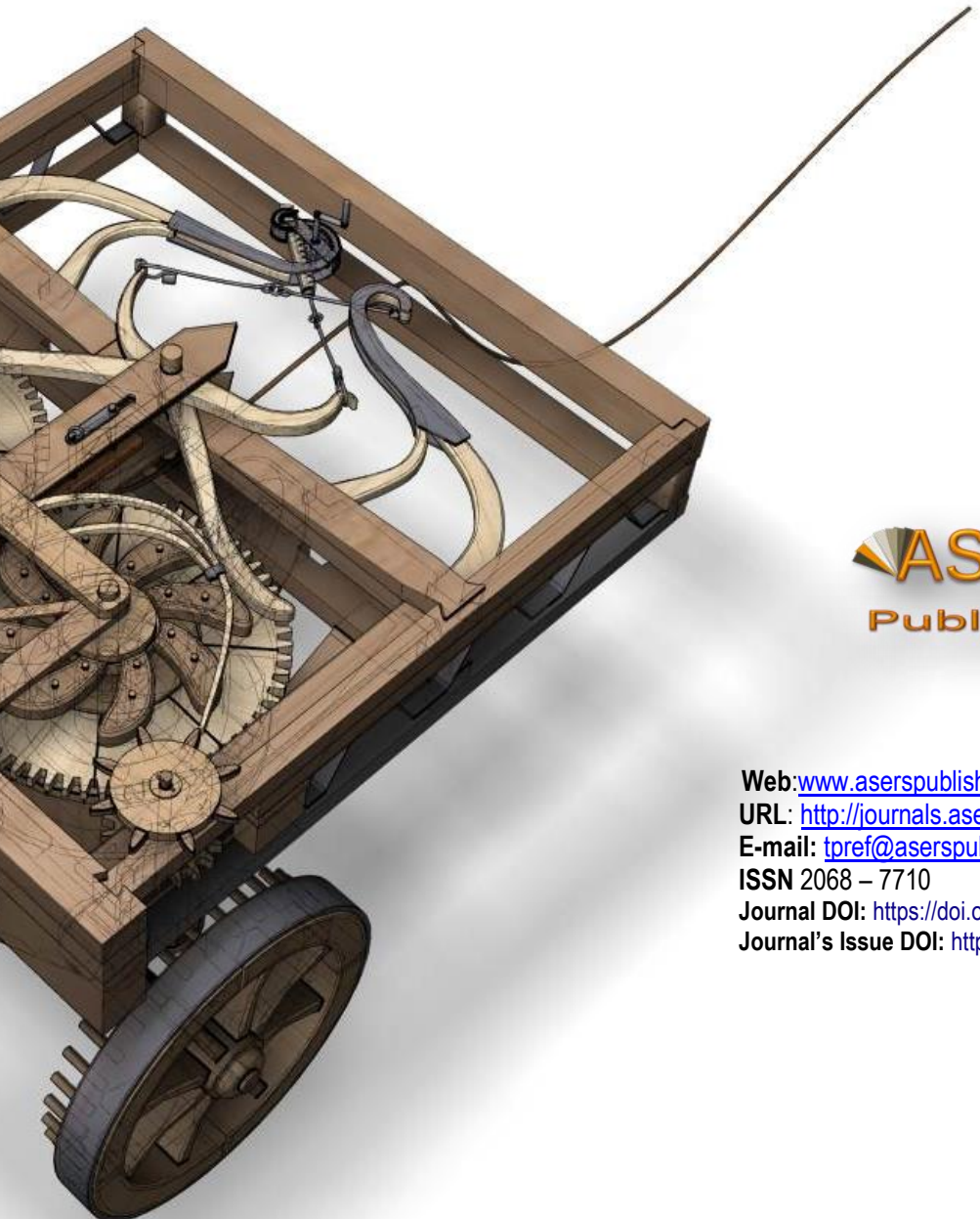
[12] Raiter, Omri. 2021. Macro-Economic and Bank-Specific Determinants of Credit Risk in Commercial Banks. *Empirical Quests for Management Essences,* 1 (1): 36-50. Available at: https://researchberg.com/index.php/eqme/article/view/28

[13] Serwadda, Isah. 2018. Determinants of Commercial Banks' Profitability. Evidence from Hungary. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 66 (5): 1325–35. DOI:https://doi.org/10.11118/actaun201866051325

[14] Tafirei Mashamba, and Chisinga N Chikutuma. 2023. Determinants of Bank Profitability: Evidence from the Emerging Economy. *Corporate and Business Strategy Review*, 4 (4, special issue): 310–23. DOI:https://doi.org/10.22495/cbsrv4i4siart12

[15] Titko, Jelena, Viktorija Skvarciany, and Daiva Jurevičienė. 2015. Drivers of Bank Profitability: Case of Latvia and Lithuania. *Intellectual Economics*, 9 (2): 120–29. DOI: https://doi.org/10.1016/j.intele.2016.02.003

ASERS



ASERS Publishing

Web:www.aserspublishing.eu
URL: http://journals.aserspublishing.eu/tpref
E-mail: tpref@aserspublishing.eu
ISSN 2068 – 7710
Journal DOI: https://doi.org/10.14505/tpref
Journal's Issue DOI: https://doi.org/10.14505/tpref.v15.4(32).00